Material recovery facility processing refund protocol

Sampling strategy

Version 01

November 2017

www.epa.nsw.gov.au

Published by:

NSW Environment Protection Authority
59 Goulburn Street, Sydney NSW 2000
PO Box A290, Sydney South NSW 1232
Phone: +61 2 9995 5000 (switchboard)
Phone: 131 555 (NSW only – environment information and publications requests)
Fax: +61 2 9995 5999
TTY users: phone 133 677, then ask for 131 555
Speak and listen users: phone 1300 555 727, then ask for 131 555
Email: info@epa.nsw.gov.au
Website: www.epa.nsw.gov.au

Report pollution and environmental incidents
Environment Line: 131 555 (NSW only) or info@epa.nsw.gov.au
See also www.epa.nsw.gov.au

ISBN 978 1 925688 61 0
EPA 2017P0390
November 2017

CONTENTS

1 ABOUT THIS DOCUMENT

This Material Recovery Facility (MRF) Sampling Strategy sets out the theory and methodology for the development of the associated MRF Sampling Plan, which is then implemented to generate the data for the development of the eligible container factors, as referred to in the MRF Protocol (July 2017). It also sets out the validation requirements for those MRFs using direct counting of containers.

Under the Container Deposit Scheme (CDS), the Scheme Coordinator is required to pay MRF operators a refund for eligible containers that are processed by a MRF for reuse or recycling. Due to the practical difficulties in directly counting each individual container at most MRFs, the MRF Protocol allows for a statistical determination of the number of containers processed. The basis of this determination is a state-wide 'eligible container factor', developed by the EPA and applied by the Scheme Coordinator to the weights of output material types claimed by a MRF operator. The MRF Protocol governance structure is illustrated in Figure 1 below.

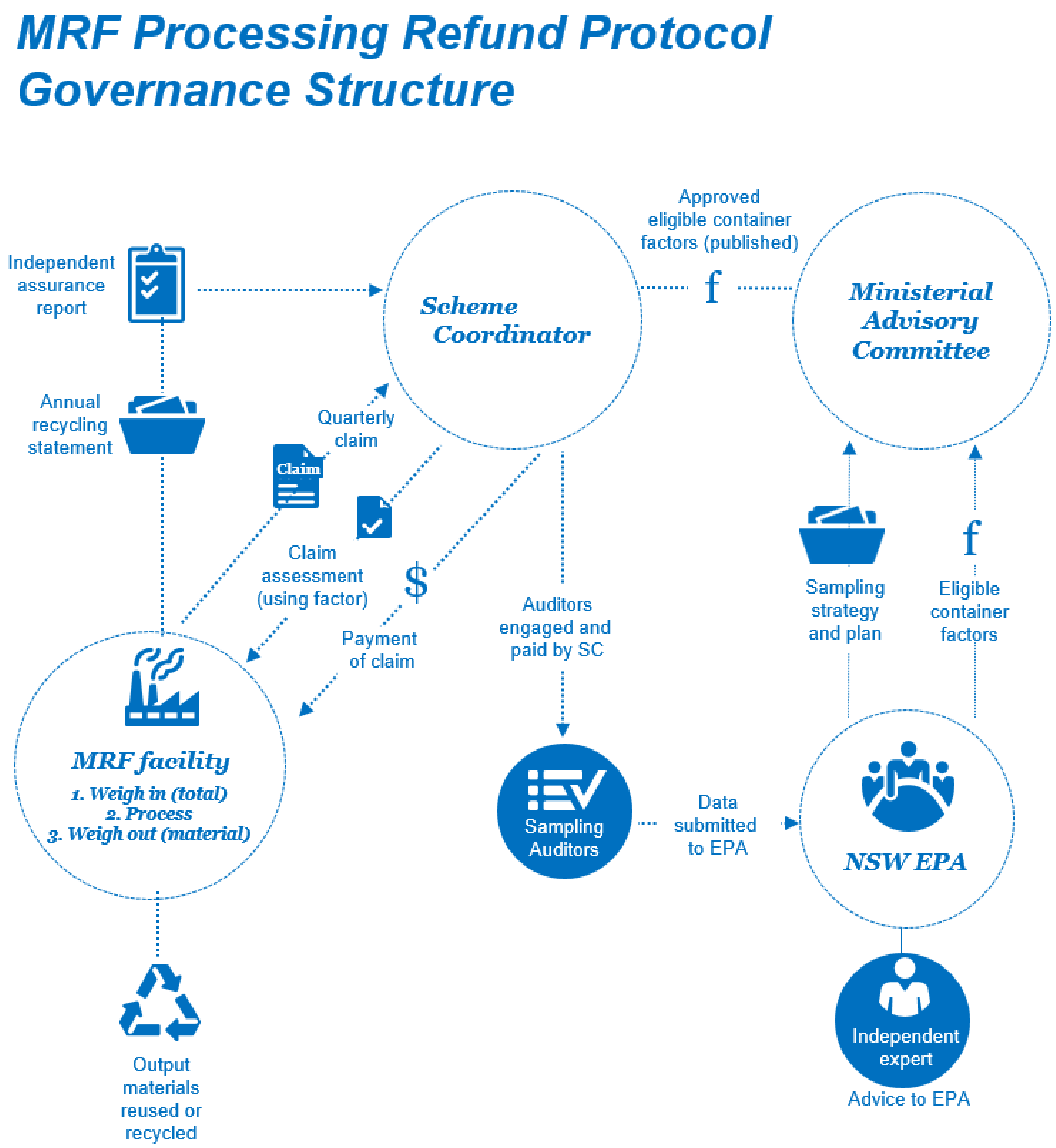


Figure 1. MRF Protocol Governance structure

1.1 Definitions

Terms within this Sampling Strategy have the same meaning as defined in the *Waste Avoidance and Resource Recovery Act 2001* (the Act), the *Waste Avoidance and Resource Recovery (Container Deposit Scheme) Regulation 2017* (the Regulation) and the MRF Protocol.

Reference to "MRF operator" in this Sampling Strategy does not include bottle crushing service operators or alternative waste treatment plant operators.

Abbreviations

NSW	New South Wales
EPA	NSW Environment Protection Authority

CDS	NSW Container Deposit Scheme
MRF	Material Recovery Facility
MRF Protocol	Material Recovery Facility Processing Refund Protocol
LPB	Liquid paperboard
HDPE	High-density polyethylene
PET	Polyethylene terephthalate
LGA	Local government area

1.2 Aims of this Sampling Strategy

Refunds to MRF operators are expected to remain a significant proportion of total refunds under the CDS for the foreseeable future. The Sampling Strategy therefore aims to ensure that:

1. the methodologies used to determine those refunds are suitably transparent for the range of relevant stakeholders;
2. the sampling carried out to support the development of the estimate factors is efficient and cost-effective;
3. the methodologies have sufficient checks and balances to ensure protection of the financial and reputational integrity of the CDS;
4. the approaches used can and will adapt as more data is collected and more is learnt;
5. the methodologies appropriately balance risks while maximising legitimate returns to MRF operators; and
6. the overall objectives of the CDS are supported, including reducing and dealing with waste generated by beverage product packaging and promoting the recovery, reuse and recycling of empty beverage containers.

2 Summary

State-wide average

The MRF Protocol and the Sampling Strategy support the development and use of a state-wide factor for each output material type to determine the processing refund amounts to be paid to MRF operators. Using a state-wide average, rather than requiring quarterly sampling at each MRF, means that sampling is more efficient, cost effective and equitable. Sampling will be centrally coordinated by the Scheme Coordinator, and the cost of sampling will be spread proportionally amongst MRF operators.

Sampling at MRFs

Sampling on outputs from MRFs, rather than on inputs to MRFs or at the kerbside, helps to promote one of the primary objectives of the CDS: promote the recovery, reuse and recycling of empty beverage containers. Due to practicalities in sampling material that has been bailed, this Sampling Strategy adopts a process sampling model where samples are taken after processing but prior to bailing at the MRF.

The data from MRF sampling will be used to determine the eligible container factors for aluminium, PET, HDPE, mixed plastics (excluding segregated PET and HDPE), and mixed plastics (including PET and HDPE). The data will also be used to calculate the levels of contamination in processed glass which will feed into the development of the eligible container factor for glass.

Sampling at kerbside

Sampling trials commissioned by the EPA have shown that the number of eligible glass containers cannot be effectively sampled at a MRF due to breakage. Given this, the number of eligible containers per tonne of glass will instead be determined from kerbside sampling and cross referenced with sampling at MRFs to determine the amount of contamination in processed glass. Aluminium and plastic (separated into PET, HDPE and other plastics) should also be included in kerbside sampling to validate the surveys conducted at MRFs.

Material types excluded from sampling

Liquid paperboard has been excluded from the Sampling Plan as it generally remains in the paper recycling stream, where it is a contaminant, and is therefore not reused or recycled. Steel has been excluded due to the extremely small number of eligible beverage containers in the steel stream. MRF operators who want to claim a processing refund for these material types may use the direct counting method.

Stratified random sampling

The Sampling Strategy employs the statistical technique of stratified random sampling which is intended to improve the chance of obtaining a representative sample more efficiently. Stratification will be by material type, by size of MRF, and over time. Data for lower priority strata (eg small MRFs and coverage of weekdays) are designed to be populated over several quarters, while higher priority strata (eg large MRFs) would be fully sampled each quarter.

Sample size and frequency

The sample unit size for process samples to be taken at a MRF is one cubic metre. This was based on a trial sampling audit conducted   Due to the difference in container sizes, the effective sample count will be different for each material type, so this has been allowed for when deciding on the number of samples to take of different material types.

Adaptive strategy

The initial Sampling Plan is constructed largely on the basis of informed assumptions. It is anticipated that this plan will evolve significantly as data are collected over the first year of the scheme. At the same time, other information about the scheme such as the total number of beverage containers placed into circulation, as well as the number of containers redeemed at collection points will also become available. The Sampling Strategy will adapt to include this information.

Efficiencies

Well-established survey sampling methods should lead to gains in the efficiency of sample design over time, as better quality information about the critical sources of population variability become available. The accuracy of the estimates should also improve as the data accumulate and more tailored estimates that build on the data may become possible.

Sampling Plan

The Sampling Plan associated with this Sampling Strategy outlines where and when sampling should occur. The details of the initial Sampling Plan, and all future Sampling Plans, will be referred to the Ministerial Advisory Committee for approval and provided to the Scheme Coordinator so that they can engage an auditor to undertake the sampling.

Sampling Plans will not be published. In order to get a representative sample, it is important that each audit is only advised to the MRF as close as possible to the audit date to reduce the opportunity for 'gaming' of the sampling process.

Stocktake to determine baseline

While the CDS commences on a specified date, recyclable material will be received and processed at MRFs before, during and after this date. An important component of this Sampling Strategy is recognising and addressing the fact that there will be material stored at MRFs that was received and processed prior to the scheme commencement date. To address this, MRF operators must complete a stock on hand assessment on 30 November 2017 and the Scheme Coordinator may engage an independent surveyor to conduct site visits at MRFs to determine the quantity material stockpiled prior to scheme commencement to assist in determining a baseline.

3 SURVEY DESIGN

The purpose of conducting sampling is to obtain data that can be used to calculate an estimate of the average number of eligible containers per tonne of recyclable material processed over a specified period. These estimates need to be sufficiently accurate to ensure that the integrity of the CDS, in terms of its financial operations and fairness to participants, is not compromised.

This will allow an unbiased estimate of the corresponding population average to be calculated for all recycled material in NSW over the period a claim is made under the CDS, along with unbiased estimates for sub-populations such as different material types and different sized MRFs.

3.1 SURVEY DESIGN CONSIDERATIONS

The dynamics of the processes undertaken at MRFs must be considered in developing the survey design. As the flow rate of material through a MRF is being measured, questions extend beyond the number of random samples required. These include physical issues such as where in the process samples are taken, as well as the physical sample sizes. The frequency with which samples are taken is also an important aspect of survey design.

Sampling Plans must consider systematic changes in the process over time if they are to be representative. These factors contribute to the variance in what is being measured, which in turn affects the number of repeated samples that will need to be taken.

Sampling must also be efficient, in terms of delivering the most accurate estimates for a given sample size or cost of sampling. To that end, the sample should be stratified or grouped according to the contribution to the overall variance in the number of eligible containers per tonne.

Two obvious stratification options are based on the throughput of the MRF and on the different streams of recyclable material processed. Sampling frequency will also impact on the efficiency of Sampling Plans. This includes short-term variation, as well as hourly, daily, weekly and seasonal variation in material flows.

3.1.1 Sampling design phases

A good sample design requires a good quantitative sense of the variation in the target populations and the flow of inputs and potentially the methods of processing those inputs. However, prior to the start of the scheme there is very limited useful information available. Moreover, changes to the target populations and the way materials are processed as people adapt to the scheme mean that past samples will no longer be representative of the current state of the scheme.

As a consequence, the Sampling Strategy has three key phases, each with more than one objective:

1. First there is an initial sample design phase, where the potential sources of variation in the data used to create the estimates for the target population need to be identified. At this time, sampling strata and sample sizes are the product of judicious and generally conservative guesses.
2. This is followed by a second phase where the survey data is processed to generate the estimates and the sampling design is optimised to reflect the variation in the actual data.
3. This leads on to a third phase where the main objective is the detection of, and response to, longer term changes in the data used to generate the estimates for the target populations. A second objective of this third stage is the identification of samples that are exceptions and also where the process may be operating outside of acceptable bounds.

3.1.2 Sampling Strategy in context

This Sampling Strategy is primarily directed at addressing the unknowns in the above process, as opposed to detailing the technical aspects of how samples should be taken. It should also be expected that in each stage there will be trade-offs to consider such as the costs of additional accuracy versus the costs of errors. The strategy needs to provide sufficient understanding of these issues to allow administrative oversight of the sampling and estimation procedures adopted.

The Sampling Strategy is also directed toward the integration of information prior to and after the introduction of the CDS in December 2017. The importance of this is highlighted when looking at the composition of recyclable materials:

$$Recyclates = (Total\ Eligible\ Beverage\ Containers - Redeemed\ Beverage\ Containers - Losses) + \text{Other Recyclable Materials}$$

Estimating the net flow of materials (the quantity in brackets above) is the object of the MRF survey. In this context, the only unknown after the scheme is in operation are the losses, as for example to landfill and litter. As a consequence, MRF survey estimates should not be considered in isolation as there are clearly checks and balances as well as the opportunity to inform the MRF Sampling Plans and to support the audit process.

Conclusion: A database of material flows from MRFs and supporting kerbside sampling should be maintained with a view to improving the accuracy and efficiency of future survey designs and validating MRF survey estimates. Further, the MRF survey data should inform future audit process in a timely way.

3.2 Estimating the number of eligible containers

From the perspective of maintaining the financial integrity of the scheme, a single NSW-wide estimate of the average number of eligible containers per tonne of recyclable material for which a claim is made is robust. As claims are made on a quarterly basis, the estimate should be representative of the materials processed for recycling over the same period. The level of accuracy of the estimate should be sufficient to ensure the financial risks and the costs of sampling and estimation are acceptable to stakeholders in the CDS.

It is important to keep in mind that the estimate of the average number of eligible containers per tonne and the error bounds about that estimate are both statistics that are subject to error. This error takes two forms:

- Bias due, for example, to the expected throughput of the sample not being the same as that of the population of interest. The wrong population may have been sampled, or a sub-sample of the population that has, for example, a different expected throughput.
- Sampling error due to variability in the samples being measured within the population of interest.

To address the first source of error, a representative sample should be taken from the population of interest. To address the second source of error, an adequate sample size should be selected. These two solutions are interrelated.

3.2.1 Population of interest

The actual population values of interest are the numbers of eligible containers in each tonne making up the total tonnage claimed against the scheme over a three-month period, as opposed to the numbers of eligible containers processed over the same period by each facility in the population of MRFs. While the total number of eligible containers, defined by multiplying the average number of eligible containers per facility by the number of MRFs, is the same as that obtained by multiplying the average per tonne by the total tonnage, the sampling strategies for the two estimates are quite different. The former requires a representative sample of facilities, the latter a representative sample of the tonnes claimed.

Recyclable materials are separated (to varying degrees) at a MRF. Hence samples can be taken at MRFs and estimates will be by material type. These samples should be relatively homogenous, when compared to flow of material into a MRF, thus allowing for a more efficient sampling design in terms of producing a reliable estimate of the average number of eligible containers per tonne of material claimed across all MRFs. This efficiency gain comes largely from recognising that some materials are more important than others in terms of their contribution to the overall claim. Greater accuracy is needed for materials with a large number of eligible containers per tonne while less accuracy is needed for those with a small number of eligible containers per tonne.

However, when developing the MRF Protocol it was recognised that a single estimate of the number of eligible containers per tonne could create inequities in claims made by MRFs with different product mixes, which in part would be passed on to local communities. It could also create changes in the incentives to process different materials with unwanted consequences. Accurate measures of the numbers of eligible containers by each material type were therefore considered to be desirable.

Some MRF operators have also expressed an interest in having their own estimates of eligible containers per tonne of recyclable material. This may be because this information would allow them to evaluate their current recycling technology or to assess the opportunity to invest in technology that could, for example, introduce less waste contamination.

While these are all valid issues for consideration as they could allow the CDS to better and more efficiently meet its objectives, there are also two important constraints:

- the cost of the survey, including the direct cost of the response burden by the MRF operators; and
- the need to ensure that reliable estimates are available at the time the first claims are made against the scheme.

The sample sizes required for reliable estimates by material type and by MRF are an issue with respect to costs, which may in some cases exceed the benefits, as well as one that complicates the logistics of collection. It is important to recognise that the MRF sampling process will continue over the life of the scheme. As the scheme progresses and better information will be available about the composition of eligible material being processed, estimating averages for MRFs with, for example, different sizes or technologies may enhance the performance of scheme. The information collected from the survey and the CDS more broadly may also bring to light other opportunities to improve the scheme.

Conclusion: The primary objective of MRF sampling in the first year should be to determine state-wide eligible container factors for each material type.

3.2.2 Representative samples

A representative sample is a subset of a population that has the same characteristics as the overall population. Here, a representative sample corresponds to having the same average number of eligible containers per tonne of recyclable material. Conversely, a non-representative sample will have different characteristics when compared to the overall population, being representative of another population with characteristics unlike that of the population of interest. Given this definition, the most obvious question is how to draw a representative sample when the characteristics of the population are not known. One answer is that a random sample can be expected to be representative if each member of the population has an equal chance of being selected into that sample.

With over 30 MRFs spread across NSW, and a limited number of auditors taking samples over a three-month period, there is a reasonable possibility that some groups within the population of interest will be missed, either by chance or systematically. How important this may be would depend on how different the characteristics in these subpopulations are, and how big a contribution they make to the overall mean number of eligible containers per tonne.

3.2.3 Stratification

Stratified random sampling is a technique that is intended to improve the chance of obtaining a representative sample. The population of interest is divided into groups, each containing members with similar characteristics, and independent samples are taken from each group. Two such strata have been introduced already, the first corresponding to sampling by material type, and the second corresponding to MRF size, as measured by throughput. The sizes of MRFs vary greatly, from small regional MRFs to large metropolitan MRFs.

It is reasonable to assume that, as throughput increases, the contribution to the variability of throughput also increases. Consequently, sampling on the basis of throughput is a proxy for sampling based on the contribution of each MRF to the variability of throughput about the overall average of interest (the average number of eligible containers per tonne of recyclate). In contrast, sampling on the basis of material type effectively takes into account systematic differences, over the sample period, in:

- the source of the input materials; and
- differences in the handling of those materials.

Within the three-month sampling period, the flow of materials into a MRF for processing will vary. For example, the location of metropolitan collection of recycling varies on a daily basis within each council area. Seasonal factors will influence beverage consumption, which in turn will alter the composition of recyclable material, as will major holidays and weather conditions. Consequently, spreading the sample over time is the third level of stratification. Over the three-month sampling period, the sample should be relatively evenly spread across say months in the first instance, weeks in the second and thirdly days of the week.

While having a number of different sub-populations adds to the burden of design, the gains will be worthwhile if the strata create a more representative sample. In particular, sampling from a large number of strata can easily create an issue of sample size within any one stratum. Again, the MRF sampling process is ongoing, so stratification can be prioritised and then spread out between successive quarters. Low priority strata, as for example small MRFs and weekdays, could be designed to be filled over several quarters, while high priority strata such as large MRFs would be fully sampled each quarter.

Conclusion: A stratified random sampling design should be adopted with prioritised strata based on material type as well as on throughput within and across quarterly sample periods.

3.2.4 Accuracy and sample size

There is a direct relationship between the expected accuracy of the estimate of the population mean number of eligible containers per tonne and sample size. However, this relationship depends on three factors:

1. the sampling distribution of this estimate;
2. the sample sizes within each of the sampling strata; and
3. the variances of the population counts (numbers of eligible containers per tonne) within each of the sampling strata.

In order to ensure that every tonne of processed recyclate has an equal chance of selection, sample sizes within each stratum should be proportional to the total tonnage processed by MRFs in that stratum. The estimate of the population mean is then the corresponding sample mean, and the central limit theorem provides reasonable assurance that the sampling distribution of this sample mean is normal, even if the distribution of population values is not, so long as the samples are of sufficient size. For a normal sampling distribution, the formula for the minimum sample size, N, that will generate a 95 per cent confidence interval for the population mean that is within δ per cent of the sample mean is:

$$N = \left(\frac{1.96}{\delta}\right)^2 \sigma^2$$

Here σ^2 denotes the population variance, i.e. the sum of the variances within each stratum, weighted by the total tonnage processed by MRFs in that stratum. Note that the minimum sample size increases at an increasing rate as:

- the acceptable percentage error becomes smaller; and
- the standard deviation σ increases.

An important point is that the relationship between minimum sample size and accuracy is based on the true population variance, something that there is limited reliable information about prior to the start of the CDS and for which there will be increasing sample information about through successive surveys. The initial sample design can be made on the basis of sampling theory, information about the composition of recyclable material prior to the introduction of the CDS, and experience. However, there is no way to assure that a given sample size will deliver the desired level of accuracy: errors may be larger, or the sample size greater than what was needed.

High levels of accuracy can impose high costs, not only in terms of the direct costs of sampling but in terms of the burden of response on MRFs. It is useful to think about the need for accuracy in terms of the costs of error and the costs of sampling to the stakeholders in the scheme. For example, a five per cent error margin at a 95 per cent level of confidence implies that: there is no more than a five per cent chance that the payment will either exceed the correct claim by five per cent or be less than the correct claim by five per cent.

This raises the obvious question of what is more important - the five per cent error margin or the five per cent chance that the correct payment will either be larger or smaller than the estimated payment by this margin. If the error margin is at the edge of what the scheme can sustain, then the risk of failure is the probability that the true payment is either over or under the actual (i.e. estimated) payment by more than the margin. If a ten per cent error is twice as bad as a five per cent error there is a better way to look at the problem: what is the acceptable average absolute error, i.e. the amount of either overpayment or underpayment. This is a simplified error margin, as for example, a five per cent error margin at a 95 per cent level of confidence may equate to an absolute error margin of 2 per cent.

It is not the intent of the Sampling Strategy to say how risk should be considered. Either of the above alternatives are sufficient to determine sample sizes as they both characterise an error distribution. However, so long as risks are likely to remain proportional to the level of error, the trade-off between accuracy and cost will become increasing unfavourable as accuracy requirements are set higher.

Conclusion: The tolerance for estimation error should be set in the context of risk associated with payments made against the CDS and the costs of sampling. An acceptable margin of error in payments can be then used to determine a sample size to meet that expected level of accuracy.

3.3 MRF Sampling

3.3.1 Sampling options

There are three basic sampling options when estimating the number of eligible beverage containers in recyclable material at a MRF:

- input sampling, either prior to collection (kerbside) or at the point of delivery to a MRF;
- process sampling of the flow of materials after they have been segregated; and
- output sampling of materials that are to be transported to recyclers.

As MRF outputs (in units of tonnes) are the target population, output sampling would, in the absence of other considerations, be the preferred approach. However, the cost of sampling outputs is prohibitive due to their size and the degree of compaction of the materials. Furthermore, breakage and other damage would make it impossible to accurately differentiate between eligible and non-eligible materials.

Input sampling provides the best opportunity to accurately differentiate between eligible and non-eligible materials, as the process of separating materials damages containers to the point where they are difficult to recognise, particularly glass. However, ensuring that input sampling is representative of the flow of outputs can be problematic if the source of materials is diverse, as for example household, commercial, construction and industrial waste. Even household waste from different geographic areas may vary substantially. This would add to the number of samples needed to be taken and make scheduling difficult. Further, the process of segregating materials generates waste and some eligible materials will enter landfill, as opposed to being sent to recycling.

Process sampling is a common quality control practice in manufacturing, particularly where it can be difficult to identify potential defects in a final product or where some components are critical to product quality. Sampling strategies are dictated in large part by the manufacturing process to ensure, for example, that tolerances at different stages of the process are set through engineering standards. Sampling strategies may also need to consider how the logistics of sampling may affect the efficiency of the production process. While the sampling objectives here are not about quality assessment and control, the principles are the same.

3.3.2 Recommendations from trial audits

The EPA commissioned trial sampling audits at a number of MRFs in NSW. A number of recommendations regarding sampling practices came out of these trials, including:

- process sampling is appropriate for all materials, excluding glass due to breakage;
- the number of eligible containers per unit of weight sampled should be directly counted, and then converted to a per tonne factor;
- conducting MRF sampling as close to the end of the processing line (the bailer) as possible to ensure the sample is representative of the final output;
- a sample unit of one cubic metre is, from a physical perspective, efficient; and
- ten to 20 samples a day can be taken, weighed and sorted in a day.

The trials included input sampling of glass, counting eligible containers on the basis of whole bottle and bottle necks. Based on a comparison of the results with kerbside audits of glass and other material, roughly 80 per cent of eligible glass containers were too damaged to be identified. As a result, it is recommended that glass be sampled at the kerbside. This would raise the same concerns as input sampling at a MRF but would be more representative of the actual percentage of recycled glass that are eligible containers.

The question of how to conduct kerbside sampling of glass is considered in Section 5.

Conclusion: MRF process sampling should be conducted to determine the eligible container factors for all output material types, other than glass. Glass should be sampled at the kerbside, with process sampling at MRFs to determine contamination levels in processed glass.

3.3.3 Physical sample sizes and sampling frequency

The recommendation from the trial sampling audits was to use one cubic metre sample units for all material types except glass. Glass contamination samples should be taken in 10 litre sample units. All samples should be taken at the MRF after processing and prior to bailing.

In adopting this method, the effective sample count will be different for each material type (i.e. one cubic metre would contain quite a large number of aluminium containers and relatively few HDPE containers), and this will need to be factored in when determining the number of samples to take of different material types. The question of how many samples to take and when to take them therefore needs to be considered further.

Conclusion: Samples should be taken in one cubic metre units, except glass contamination samples which should be taken in 10 litre units. All samples should be taken after processing and prior to bailing at the MRF.

3.3.3.1 Number of samples v sample size

There is trade-off between the number of samples taken and the physical size of a sample. The idea that a few large samples might be as good as a lot of small samples seems intuitive but this is not necessarily the case. This is because the flow being measured has a composition that can change over time. The fact that time becomes important is best illustrated with an example.

Example

The eligible container count per 5kg obtained from a 50kg sample can be thought of as the average of the counts obtained from ten 5kg samples taken over a relatively short period of time. Had ten samples been taken in that period, the number of eligible containers will have varied, but the standard error of the single per 5kg count for the 50kg sample obtained could be calculated by merging the ten 5kg samples as the standard error of the mean of the counts from the ten 5kg samples taken in that period.

However, what if the ten 5kg samples has been spread out over a longer period of time that better represents the capacity for the composition to change over time? The variation in the number of eligible containers could then be much larger, due for example to differences in when the material was sourced and from where it was sourced. This implies that the actual standard error of a “well spread” per 5kg count will be larger than the calculated standard error of the per 5kg count based on our single 50kg sample.

Another way of expressing this is to say that the correct standard error of the average of the ten individual 5kg counts making up the 50kg sample is actually greater than the standard error of the average of ten effectively independent 5kg counts taken across a much longer interval of time. This phenomenon is usually referred to as the decrease in the effective sample size due (in this case) to temporal clustering of container counts. Effective sample size is a useful way to think about how a sample deviates from the ideal case where sample values are independent and hence reflect the true level of population variability.

The question of how to stratify revolves around the frequency of sampling at a given site, and the rates of change in the composition of materials. If council trucks are coming in from different suburbs during the day, there may be significant variation within a day. If they come from different suburbs on different days there may be significant variation from day to day and so forth.

The practical problem is that the number of potentially important strata is large and information is not available on how to prioritise them in terms of maximising between stratum variability. However, the further these are apart in space and time, the less correlated the sampling units from them will tend to be, or in other words the more the between stratum variability. This gives rise to a relatively simple and conservative strategy that corresponds to implicit stratification across space and time: spread the sample out as much as possible in terms of visits to individual sites, between revisits to a specific site and over the day during a visit.

However, there are practical issues with this strategy, particularly with respect to spreading out the sample within a day or even across days within a week. For example, taking single samples in sequence from separate material lines is likely to be disruptive to MRF operations and excessively time consuming for audit staff. It is therefore likely to be more practical to address the issue of short term sampling variation through a separate fit for purpose Sampling Plan and maintain the focus of stratification for the MRF Sampling Plan on a longer time scale.

Conclusion: The sample should be broadly stratified by MRF size and by quarter (prioritising by month and by week). Over time, small fit-for-purpose Sampling Plans should be considered to examine whether clustering of samples within or between days significantly reduces effective sample sizes.

3.4 ADAPTIVE SAMPLING AND VALIDATION

As noted previously, the initial Sampling Plan is constructed largely on the basis of informed assumptions. This plan will evolve significantly as data are collected over the first year of the scheme. At the same time, other information about the scheme such as the total number of beverage container placed into circulation, as well as the number of containers redeemed at redemption centres will also become available. The Sampling Strategy should adapt to include this information.

3.4.1 Adaptive strategy objectives

An adaptive strategy should have a clear set of objectives, and as well as being capable of responding to unanticipated events. An important (but not an exhaustive) list of objectives is:

- improving the efficiency of the sample design and the accuracy of the estimates of eligible containers;
- validating the estimates of eligible containers per tonne using external information sources;
- stabilising the process of data collection and aggregation to allow consistent comparisons over time;
- monitoring changes in the composition of eligible containers as the scheme progresses;
- supporting the auditing of claims; and

- responding to stakeholder requests for MRF or region specific estimates of counts of eligible containers.

The objective of this initial MRF Sampling Strategy is not to suggest specific ways of meeting these objectives. However, there are a few points that can be made. Timing is an issue for revising the sample design to include more up-to-date information from the previous quarter's sample. The window between when the estimates need to be provided and the next round of sampling needs to commence is short. Data from the current sample should be made available in close to real time to allow revisions to be done immediately after or even prior to the completion of the current sampling run. In the early stages of the scheme, this would also allow verification that the sampling is progressing as planned and also allow preliminary estimates to be generated if required.

3.4.2 Efficiencies

Well-established survey sampling methods should lead to gains in the efficiency of sample design over time, as better quality information about the critical sources of population variability become available. The accuracy of the estimates should also improve as the data accumulate and more tailored estimates that build on the data may become possible. Central coordination of data collection across the different elements of the CDS will facilitate validation.

However, changes in the way beverage manufacturers, consumers, councils and community groups respond to the scheme will impact on the composition of recycled materials. The estimates of eligible containers per tonne will not be stationary over time and this will affect the way in which current and past data can be integrated. This represents a reasonably complex statistical problem. It is also critical to meeting the last two objectives that were set out above.

Lastly, the issue of legacy needs to be addressed. Processes need to be documented and systems maintained after their initial design.

Conclusion: The objectives outlined above should be set as a guide to an adaptive Sampling Strategy. Implementing an adaptive strategy is a complex statistical problem and individuals or an organisation with appropriate expertise and experience to carry out such a strategy should be engaged and legacy arrangements put in place.

4 MRF Sampling

4.1 Methodology

Starting from a point without information about the mean and variance of a population, the question of sample design can be addressed with an informed guess at an adequate sample size based on a desired level of precision of the estimate. The confidence in achieving that goal will be low, since the actual achieved precision at the start of the sampling process may be greater or less depending on the accuracy of this informed guess. However, after taking the initial sample this guess can be replaced by a sample size derived from an empirical estimate of the population mean and variance. As information about the population continues to accumulate, the sample design will improve and our confidence in the expected level of precision of the estimates will increase.

4.1.1 Determining initial sample size

There are three stages in the methodology employed to determine an initial sample size:

1. Data from past kerbside audits of recycling bins is used to derive a mean and a variance of a hypothetical sample taken at a MRF prior to the start of the CDS. Household recycling accounts for about 80 per cent of the inputs into a MRF.
2. Assumptions regarding the proportion of containers processed through redemption centres, referred to as the CDS redemption rate, are used to adjust the mean and variance for a post CDS sample.
3. Sample sizes are determined empirically using a Monte Carlo simulation of the MRF sampling process. This is a standard approach for the type of problem under consideration. The methodology is outlined without reference to a specific material type, but is applicable to all the materials to be sampled at a MRF.

A conservative approach (leading to larger sample requirements) is generally employed in making a guess of the minimum sample size. The number of eligible beverage containers per tonne of recyclate is expressed a way that reflects the prior information available:

$$N_{EBC} = \frac{W_{BC} * p_{EBC}}{w_{EBC}}$$

where N_{EBC} is the number of eligible beverage containers per tonne of recyclate, W_{BC} is the weight of recyclate made up of beverage containers, p_{EBC} is the proportion by weight of eligible beverage containers in all beverage containers and w_{EBC} is the average weight of an eligible beverage container. These parameters can all be expressed on a per tonne of recyclate basis. In a given sample taken at a MRF these parameters are treated as random variables that contribute to the overall variation in the number of eligible containers per tonne.

Kerbside recycling audits commissioned by the EPA in 2015-16 provide pre-CDS estimates and supporting assumptions for each of these parameters by material type. The parameters from the kerbside audit are:

- proportion of eligible beverage containers by weight – 54 per cent;
- weight of beverage per tonne of recyclate – 0.73 tonnes; and
- average eligible beverage container weight – 0.2 kg.

The household variation in the proportion of eligible beverage containers could not be obtained as the household samples were combined in a truck and then counted. However, the status (eligible vs. non-eligible) of a container in a bin can be treated as binary random event and a recycling bin as a sample of containers.

The standard error of a proportion estimated from a sample of n independent binary events, each with the same probability of occurring, is $\sqrt{p(1-p)/n}$, where n is the sample size (i.e. the number of containers). The reported average number of containers in a household bin was about 11 per week, of which roughly 8 were materials other than glass. Materials in a household bin would vary systematically as well as randomly due to the influence of household demographics and preferences. The problem is aggregating the household level data to a level that more closely matches an input sample taken at a MRF. The reported average weight of containers in a household bin was 0.8kg, excluding glass. Assuming a physical sample of 35kg, the effective kerbside number of beverage containers needs to be inflated by a factor of about 44, i.e. about $n = 300$ containers, would be observed in the input sample at the MRF.[1] Taking p to be 0.5 this would give a standard deviation of about 2.9 percentage points, or relative standard error of about 6 per cent, for the proportion of eligible containers.

The weight of all containers per tonne of recyclate was based on assumptions regarding non-beverage containers as well as contamination in MRF throughput. If the average weight of all containers were known, the estimate could also be treated as a proportion. As this data was not collected, the distribution and variance of this parameter must be assumed. It is assumed the distribution is normal and that the relative standard error is substantially larger than for eligible containers. A relative standard error between 15 and 25 per cent was used based on material type.

The average weight of an eligible container derived from the kerbside audits can be assumed to be quite accurate, given the large number of households sampled (over 6,800 which would be over 46,800 eligible containers). The household level variation in material weights was again not available from the kerbside data. However, an estimate of the proportion of containers for each material type is available from the kerbside data and a rough estimate of the variability in weight can be constructed by:

- taking the weights and proportions as fixed;
- treating a bin as a random sample of six containers; and
- assuming that the probability of a container being of a particular material type is equal to the observed proportion of containers of this material type.

The weights and proportions are shown in Table 1. A random sample of 10,000 average container weights per bin was generated numerically. The relative standard error was 42 per cent. With an effective MRF sample size of 240 (6 times 40), the estimated relative standard error of the weight of an eligible beverage container would be 2.7 per cent.

Table 1. The average weight and proportion of eligible beverage containers by material type for the 2015-16 kerbside audit

Material	Weight	Proportion
Aluminium	0.014	13.3%
Pet	0.028	26.3%
HDPE	0.048	1.8%
Glass	0.224	30.3%
Mixed Plastic	0.035	28.3%

4.1.2 The Impact of the CDS

The mean redemption rate, the proportion of containers processed through collection points, is another important assumption. The redemption of containers will reduce the input of eligible containers to MRFs. This will reduce the proportion of total beverage containers, the proportion of eligible beverage containers and the effective size of a MRF sample.

The parameters can be adjusted from the pre-CDS assumptions directly as a consequence of the assumed redemption rate. For example, a redemption rate of 50 per cent across all material types

[1] For a specific material, like aluminium, the number of containers in a MRF sample will be much higher.

would reduce the weight of beverage containers and the effective sample size by about 6 per cent. The percentage of eligible beverage containers would fall by 17 per cent.

Redemption rates will vary between households and this variation is likely to reflect the demographic characteristics of a region, due for example, to the mix of high and low-density housing. Individuals consuming beverages away from home will have different propensities to redeem and/or recycle depending on the relative ease of redemption at the point of consumption. This will directly add to the sampling variation of weight of all beverage containers and the proportion eligible containers, which can be viewed as an inflation of the variance or reduction in the effective MRF sample size.[2]

To get some idea of how important this may be, the decision to opt in or out of the scheme is treated as a binary process. For example, a household either redeems all eligible containers or none. Again, a factor of 15 households per MRF sample, an average participation rate of 50 per cent is used along with the same discrete event formulation to calculate the standard error. If the decisions of the households that contributed to the sample were independent, the standard error of the participation rate would be roughly 7 percentage points within the sample.

Focusing on the proportion of eligible containers, the increase in the variability of the proportion of eligible containers due to the variability in the participation rate can be seen as the product of two binomial distributions, both with a probability of 0.5 but with different numbers of trials, 300 and 45. A small numerical simulation shows that the standard deviation of the proportion of eligible beverage containers would increase from about 2.9 per cent to about 4 per cent. This is a large reduction in the effective physical sample size at a MRF, a little less than 50 per cent. This would also increase the variability of the weight of beverage containers.

The reduction in effective sample size is overstated as common demographics within neighbourhoods are likely to lead to a reasonably strong correlation in the participation decisions of households. Further, the decision to opt in or opt out may not be all or nothing. Nevertheless, the variation in participation rates will add substantially to sampling variability and needs to be reflected in the assumptions made in determining the initial sample sizes for aluminium and plastics. This is reflected directly in the assumed relative standard errors for the weight of beverage containers and a reasonably large reduction in the effective physical sample size taken at a MRF.

A Monte Carlo simulation model is used to generate a large number of potential samples that could be taken at a MRF. Subsamples of fixed size were then repeatedly taken from this population of potential samples and the level of precision of the estimate of the number of eligible containers calculated. The size of the subsample was then adjusted until the desired level of precision was achieved, giving the initial guess at the minimum required sample size.

The structure of the Monte Carlo simulation model is provided in Attachment A, along with the parameter assumptions.

[2] The standard error of the mean is the standard deviation of the population divided by the square root of the sample size so we can equate a reduction in effective sample size with an increase in the variability of the population.

4.2 Initial Sampling Plan

4.2.1 Material types included

The initial Sampling Plan is based on estimates of outputs of materials which were provided by all MRFs that have registered to claim using the MRF Protocol. The materials included in the Sampling Plan are:

- aluminium;
- PET;
- HDPE;
- glass (contamination only);
- mixed plastics - excluding segregated PET and HDPE; and
- mixed plastics - including PET and HDPE.

Sampling trials commissioned by the EPA have showed that the number of eligible glass containers cannot be effectively sampled at a MRF due to breakage. During the trials, approximately 80 per cent of glass could not be identified. Estimates of the number of eligible glass containers will instead be determined by undertaking kerbside audits and then determining the level of contamination in processed glass at MRFs.

Liquid paperboard has been excluded from the Sampling Plan as it generally remains in the paper recycling stream, where it is a contaminant, and is therefore not reused or recycled. Steel has been excluded due to the extremely small number of eligible beverage containers in the steel stream. MRF operators who want to claim a processing refund for these material types may use the direct counting method.

4.2.2 Stratification

The Sampling Plan is stratified by:

1. Material type - because sampling will be undertaken on product lines. The stratification is done on the basis of the expected contribution of each material type to the total number of eligible containers claimed.
2. Output by material type - four MRF strata have been defined:
    a. 'Segregators' - segregators of PET and HDPE;
    b. 'Large mixed' - large processors of only mixed plastic;
    c. 'Medium mixed' - medium sized processors of only mixed plastic; and
    d. 'Small mixed' - small processors of only mixed plastic.

There are two levels of MRF stratification: between strata and within strata. Aluminium and glass is allocated jointly across all strata. Plastics are allocated to either those that process PET and HDPE and those that process only mixed plastics.

The assignment to the strata for each of the registered MRFs is shown in Table 2 below.

Table 2. Assigned strata of registered MRF

MRF	Stratum	MRF	Stratum
Solo Gateshead MRF	Large mixed	Cleanaway - Albury MRF	Segregator
PAR RECYCLING	Large mixed	Challenge Narrabri Recycling	Segregator
Polytrade Rydalmere	Large mixed	Northaven MRF	Segregator
Taren Point MRF	Large mixed	Cowra MRF	Segregator
Spring Farm MRF	Medium mixed	Lismore MRF	Small mixed
Cairncross MRF	Medium mixed	Suez Moruya MRF	Small mixed
Polytrade Chinderah	Medium mixed	Shoalhaven Recycling MRF	Small mixed
Tuncurry Materials Recycling Facility	Medium mixed	Challenge Tamworth Recycling	Small mixed
Grafton Resource Recovery Facility	Medium mixed	Armidale Regional Council Waste Facility	Small mixed
Smithfield MRF	Segregator	Mudgee Waste Facility	Small mixed
Handybin MRF	Segregator	Gunnedah Recyclit	Small mixed
Kurrajong Recycling	Segregator	Gilgandra Waste Facility	Small mixed

4.2.3 Preliminary plans

The preliminary plans were made using Monte Carlo simulations at a confidence level of 95 per cent. A minimum sample size of 30 is imposed for glass contamination, which does not impact on the determination of sample size for the remaining materials. The minimum sample size for all other materials is 20. Two basic types of sampling plans were evaluated, with:

- margins of error set by individual material types; and
- margins of error set on the total number of containers to be claimed.

The plans are detailed in Table 3. An overall summary of the sample size, as well as the margin of error and mean absolute error achieved in the Monte Carlo simulations is shown in Table 4.

Table 3. Alternative sampling plans for alternative error margins (at 95 per cent confidence level)

Plan Margin of Error	Aluminium		PET		HDPE		Glass		Mixed Segregated		Mixed Combined	
	Size	Error (%)	Size	Error (%)	Size	Error (%)	Size	Error (%)	Size	Error (%)	Size	Error (%)
All < 5%	38	5.0	45	5.0	148	5.0	-	-	85	5.0	110	5.0
Total < 5%	21	5.6	20	6.4	20	12.8	30	-	20	9.1	35	10.6
Total < 4%	35	5.6	20	6.4	20	12.8	30	-	28	9.1	57	8.7
Total < 3%	59	3.8	27	6.4	20	12.8	30	-	48	6.9	97	5.5

Table 4. Summary of the alternative sampling plans, sample size excludes glass

Expected Plan Error	Sample Size	Error Margin (%)	MAE (%)
All < 5%	426	3.9	1.5
Total < 5%	116	4.9	2.0
Total < 4%	160	3.7	1.6
Total < 3%	251	3.0	1.3

Setting a five per cent margin of error for each material type requires a very large sample size and, in terms of overall accuracy against the total number of eligible containers, is very inefficient.

A sample size of 160 would achieve a higher level of overall accuracy as measured by the margin of error. The absolute mean error is about the same and the reduction in sample size is over 60 percent. There are two reasons for this reduction:

- First, it is unlikely that all errors in each container type will be highly correlated, although there may be some correlation due to the clustering of samples within days. In other words, it is unlikely that all the materials will have either very high or very low numbers of eligible containers at the same time. So, across all material types, the variability will be less extreme. The structure of the Monte Carlo simulation treats each material type as being independent which is the most favourable case and the error margins may be somewhat understated.
- Second, the sampling is stratified by the throughput of eligible containers which is more efficient. Throughput is known in the Monte Carlo simulations, so the gain in efficiency is greater than would really be expected and again the error margins may be somewhat understated.

Nevertheless, the expected margin of error with a sample size of 160 is well below five per cent. It is also clear that targeting lower margins of error results in rapidly escalating sample sizes. Going from a four per cent target to a three per cent target increases the sample size by over 50 per cent (91 sample points).

Note that subsequent sampling plans will improve upon this margin of error since the relationship between sample size and accuracy will be better understood. Further, the data will accumulate over time adding to the precision of the estimates.

Conclusion: An overall minimum sample size of 190 (30 for glass contamination, 160 for other material types) should be adopted in the initial Sampling Plan.

4.2.4 Potential gains from adaptive sampling

The data obtained from the initial MRF sampling is likely to result in substantial revisions and improvements to future designs. The stratification of the samples between MRFs and over time will need to take into account the trade-off between spreading the sample to achieve a greater degree of representativeness and the need to build up a time series component to the data to changes in the composition of eligible materials over time. This trade-off will have a substantial impact on the rotation strategy, particularly with the smaller MRFs.

It is possible to illustrate the sort of efficiency gains that can be obtained through an empirically based sampling design. In the Monte Carlo simulations, it is possible to stratify the sample based on the true variation in the number of eligible containers in each material category. Weighing the strata by their respective standard errors would give the maximum gain in efficiency in the sample design and is shown in stabilise.

Table 5.

Subsequent sample designs can be informed by sample estimates of the variances in eligible containers. As these become more precise, the design would, with substantial assurance, approach the optimal level of efficiency. However, the point here is really to illustrate that substantial reduction in sample numbers may be possible as the CDS continues to operate and participation rates in the scheme stabilise.

Table 5. Potential maximum gain in the efficiency of the sample design

	Stratify on Throughput			Stratify on Contribution to Variance		
	Sample Size	Error (%)	MAE (%)	Sample Size	Error (%)	MAE (%)
All < 5%	426	3.9	1.5	100	4.9	2.0
Total < 5%	116	4.9	2.0	107	3.9	1.7
Total < 4%	160	3.7	1.6	139	2.9	1.2

5 KERBSIDE SAMPLING

5.1 GLASS

The number of eligible containers per tonne of glass claimed against the CDS is to be determined from kerbside sampling to obtain estimates of the average household:

- number of eligible glass containers in a recycling bin and
- the total weight of glass recycled.

5.2 OTHER MATERIAL TYPES

The decision to conduct kerbside audits for glass may be expanded to cover aluminium and plastic for information on, for example, the demographic characteristics affect CDS participation rate or to cross validate surveys conducted at MRFs. While this will involve a consideration of benefits and costs which is outside the scope of the Sampling Strategy, the sample design for a kerbside audit of all eligible materials would not necessarily change.

Kerbside audits of recycling bins in NSW have been conducted in NSW for a number of years, including three recent audits that have identified eligible containers under the CDS. The audits have been and are proposed to be conducted at the level of a Council or local government area (LGA). Hence, the sample design is directed toward obtaining a representative household sample from a subset of LGAs. The design needs to cover sampling within an LGA, as well as the selection of LGAs with a representative spread of demographic characteristics.

The trade-off between the within-LGA sample size and the number of LGAs selected for sampling is quite similar to that considered in relation to the physical sample size and the number of samples taken at a MRF. It is the total number of household samples, as opposed to households per LGA, that will determine the expected level of accuracy of the estimates.

Large sample sizes taken from only a few LGAs may provide adequate numbers but fail to be representative due to clustering that is likely to arise from common demographic characteristics within an LGA, such as income, housing and transportation costs. This would favour spreading the samples over a larger number of LGAs.

There is a second reason for selecting a larger number of LGAs. In previous audits, bin contents have been collected from households by truck, and the total sample was then audited at one location for occupational health and safety as well as efficiency reasons. This process, which is taken to continue, has one major limitation. The household level of variation in recycled materials is not able to be calculated. Therefore, it is not possible to access the expected level accuracy that might be achieved with a particular sample design at any finer level than the levels of variation between LGAs and within LGAs over time. This may not be so important in the initial design but the number of LGAs selected will have a substantial impact on how effectively the sample design can be improved on the basis of the information collected.

There are also two reasons why reasonably accurate measures at an LGA level would be desirable. These are to:

- construct a consistent time series of observations to determine how households' responses to the scheme may have changed over time; and
- match the catchments of one or more MRFS to allow a comparison of MRF inputs and outputs.

The purpose would, in large part, be to see if changes in direct participation rates in the CDS are reflected in the number of eligible containers processed at a MRF. Accurate measures at the level of an LGA may also be useful for cross-validating survey data with regional redemption rates. This information would be most valuable if the household survey in the selected LGAs covered all materials being claimed by MRFs under the scheme.

5.3 Representative sample size

The statistic to be estimated is the average number of eligible glass containers per unit of weight of all glass in a household recycling bin. Given the types of glass that are acceptable for recycling, it is reasonable to assume that glass from sources other than beverage containers is contamination. Given that the average weight of glass container is known, the statistic to be estimated is the proportion of eligible glass containers in a household recycling bin, for which we can derive required sample sizes based on sampling theory.

Starting from the premise that there is a representative household sample, a conservative estimate can be made of the number of household samples required to achieve a given level of accuracy, based on the binomial distribution. Assuming the true proportion of eligible glass containers is 50 per cent and requiring a confidence level of 95 per cent, the sample size required to achieve:

- a ten per cent error margin would be 97; and
- a five per cent error margin would be 385.

These would be the requirement for an aggregate NSW estimate or for an individual LGA. An LGA sample size of 100 would be expected to achieve under a 10 per cent error margin. Drawing these samples from four LGAs would be expected to give an aggregate NSW estimate with and error margin of less than 5 per cent. The latter would be in line with the level of precision expected for the MRF sampling of materials other than glass

A sample drawn from four LGAs is unlikely to be representative. At the same time, a sample drawn from a large number of LGAs may by well in excess of what is required under the CDS. A reasonably sophisticated stratified random sampling technique could be used to try and minimise the number of sample units needed to achieve a representative sample.

Conclusions: In relation to kerbside surveys:

A sample of 100 households should be adopted to ensure an adequate level of precision**.**

In addition to a count and weight of eligible glass containers, and the weight of all recycled glass, sampling for aluminium and plastic containers should also be included in the kerbside surveys.

A stratified random sampling approach should be adopted in order to try and minimise the number of sample units required to achieve a represented sample.

The selection of LGAs should cover at least one MRF catchment to allow for validation of MRF data.

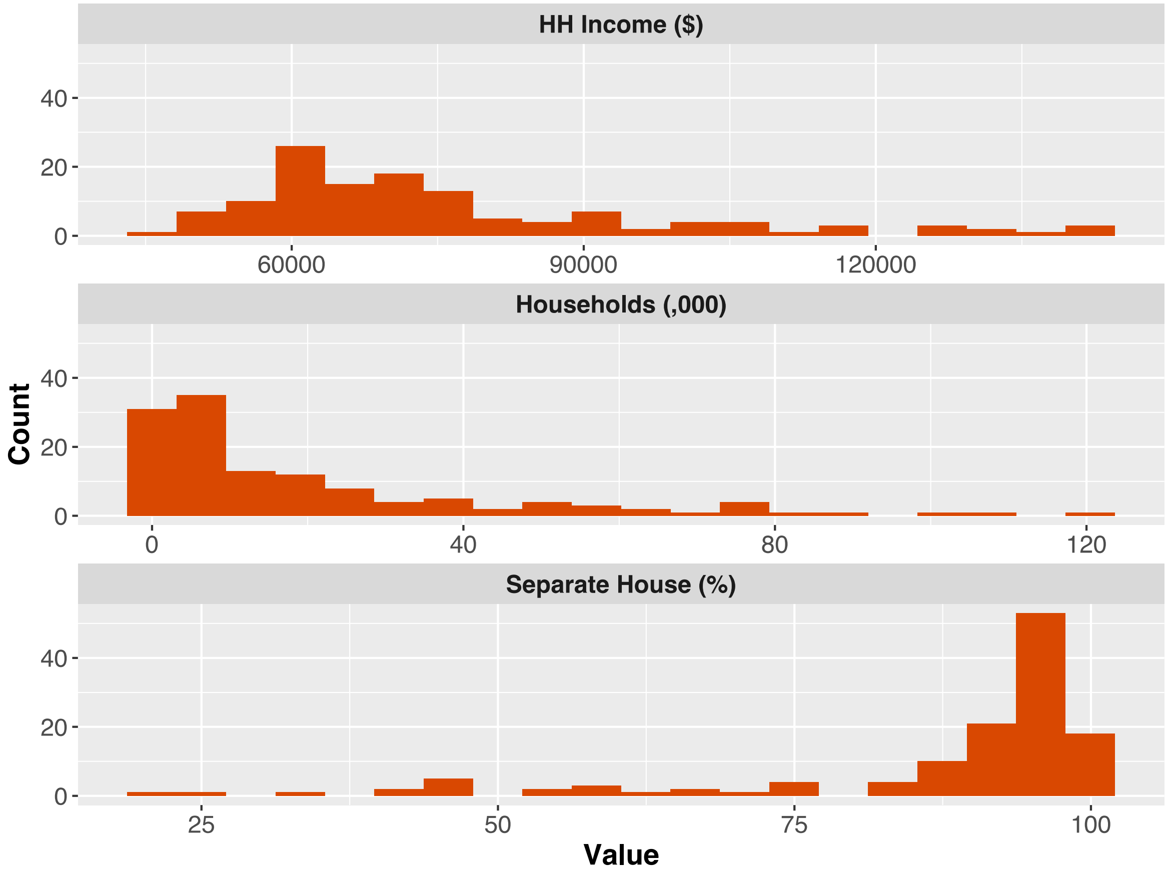
5.4 SELECTING SAMPLE UNITS

In the 2016 census there were 129 LGAs in NSW. Having this number of Councils as the sampling units presents an issue in the context of obtaining a sample representative of the demographic characteristics that are likely to affect direct participation rates in the CDS and the number of eligible containers recycled. These demographic characteristics would include: household numbers; household income; and the percentage of separate dwellings versus units.

The distribution of these characteristics is shown as a histogram in Figure 2. It is clear that there is considerable demographic diversity across Councils. However, the number of strata needed to create a representative sample depends on the correlation between the number households, household income and the percentage of separate housing, which is as follows:

- households and household income – 41.6 per cent;
- households and separate housing – (-)-40.01 per cent;
- household income and separate housing – (-)71.7 per cent.

Figure 2. Distribution of the number of households, household income and the percentage of separate housing across LGAs in NSW in 2016



Source: AnalytEcon

Together these correlations indicate that there is a positive relationship between population density and income that is reflected in these correlations. Population density will in turn be related to a range of other urban and rural characteristics such road congestion or the proximity of cafes and bars.

Perhaps more importantly this correlation will tend to lead to clustering of demographic characteristics. This, in turn, can be used to define subpopulations or strata. The idea is that if the sample is spread across the strata, it will be representative of the full range of demographic characteristics.

The objective of clustering is to assign LGAs to clusters so that the demographic variation within each cluster is small, while the variation between clusters is large. In one extreme case, each LGA could be its own cluster and all the variation would be between clusters. However, sampling each cluster would equate to a census. Suppose instead that the 129 LGAs were divided into 15 clusters such that the demographic characteristics within each one was the same. Then a sample of one from each perfectly homogenous cluster would give the same result as the census. The actual data will fall somewhere in between, and there will be a trade-off between demographic similarity within a cluster and the number of clusters formed. The stronger the correlations between demographic characteristics, the more favourable this trade-off will be.

There are a number of ways to define clusters. A hierarchical or tree based algorithm known as 'Ward's method' was used here. Household numbers and income were first converted to natural logarithms to place them on comparable scale as the percentage of separate dwellings. With five clusters the variation in demographic characteristics between-clusters accounted for over 90 per cent of total demographic variation. The within cluster demographic variation is less than 10 per cent.

The clustering is illustrated with plots of the percentage of separate dwellings against households and household income, shown in Figure 3 and Figure 4, respectively. The clusters, which are colour coded, are reasonably distinct. They are separated along a negative trend between the percentage of separate dwellings and both the number of households and household income. In each case the LGAs are spread more broadly about trend as the proportion of separate dwelling falls.

Figure 3. Demographic clusters defining the strata for the NSW LGAs (number of households versus the share of detached dwellings)

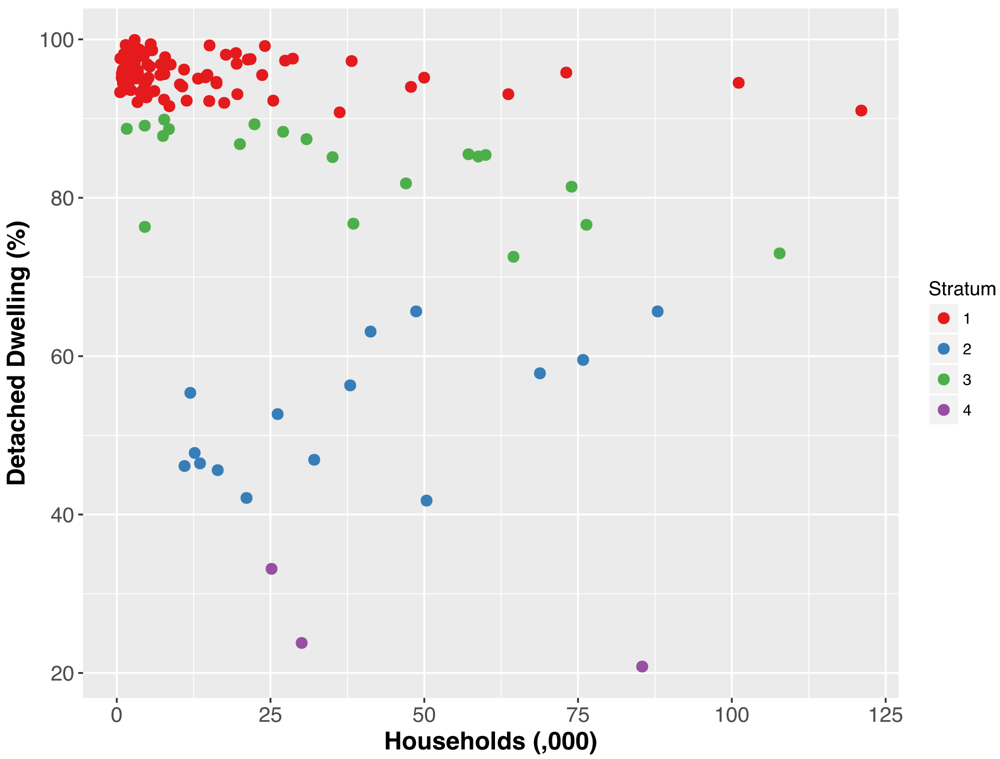
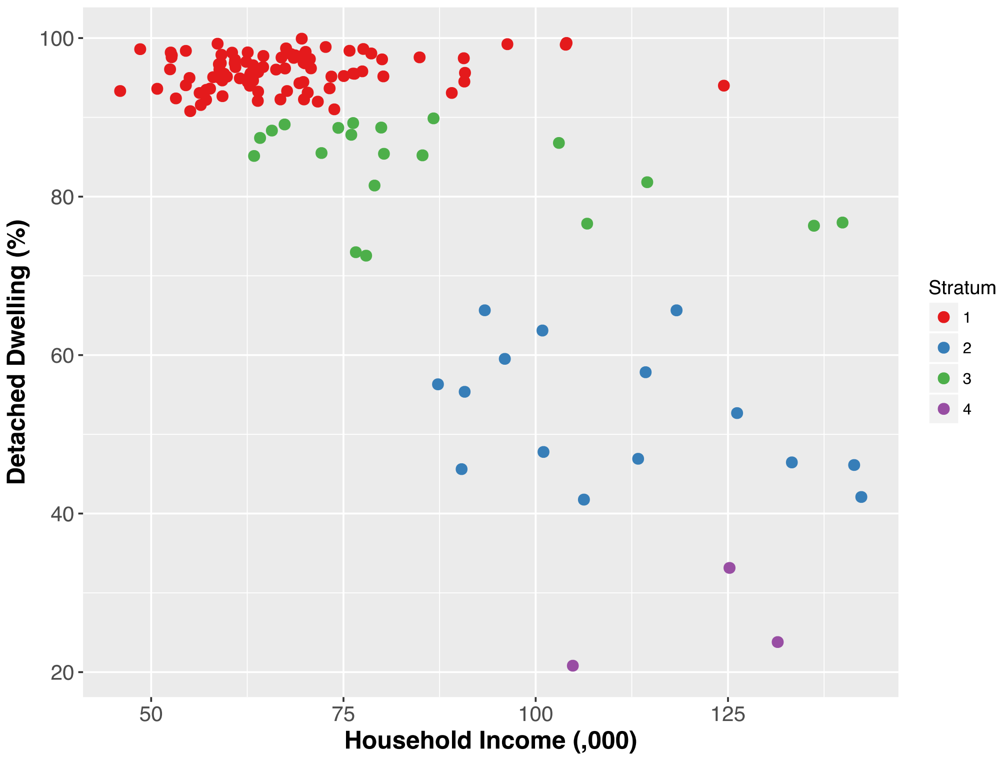


Figure 4. Demographic clusters defining the strata for the NSW LGAs (household income versus the share of detached dwellings)



Using clustering to define strata is a data mining technique, as opposed to a structured spatial analysis. However, there is a strong demographic pattern that reflects increasing population density, levels of urbanisation, and household income. They are characterised by:

- Strata 1: Non-urban with lower household income (91 LGAs);
- Stratum 2: High density urban with high levels of household income (15 LGAs).
- Stratum 3: Outer urban areas and regional centres with medium levels of household income (20 LGAs); and
- Stratum 4: Central Sydney (3 LGAs).

The LGAs in each stratum are shown in Table 6a and 6b.

Starting with the constraints of an overall minimum of 30 sample units and a minimum strata size of four sample units or the number of LGAs within the strata, the allocation of units to each stratum is based on the proportion of households contained in each stratum.

Sampling within the stratum can also be directed to be more representative. The criterion is again based on a measure of demographic variation with each stratum known as a ‘principal component’. Each stratum has a centre defined by the multidimensional mean of its demographic characteristics. The distance of each LGA from the centre of its stratum is a measure of how close it is to this mean.

Ranking each LGA in a stratum by this distance (its first principal component score) stratifies the LGAs from furthest ‘below’ to furthest ‘above’ the centre. Picking the LGAs at equal intervals gives a relatively uniform cross-section of demographic characteristics.

Table 6a Classification of NSW LGAs into five clusters or stratum: first stratum

Cluster 1				
Albury	Campbelltown	Gwydir	Murrumbidgee	Upper Hunter Shire
Armidale Regional	Carrathool	Hawkesbury	Muswellbrook	Upper Lachlan Shire
Ballina	Central Coast	Hay	Nambucca	Uralla
Balranald	Central Darling	Hilltops	Narrabri	Walcha
Bathurst Regional	Cessnock	Inverell	Narrandera	Walgett
Bega Valley	Clarence Valley	Junee	Narromine	Warren
Bellingen	Coolamon	Kempsey	Oberon	Warrumbungle Shire
Berrigan	Coonamble	Kyogle	Orange	Weddin
Blacktown	Cowra	Lachlan	Parkes	Wentworth
Bland	Dungog	Lake Macquarie	Penrith	Western Plains
Blayney	Edward River	Leeton	Port Stephens	Wingecarribee
Blue Mountains	Eurobodalla	Lismore	Richmond Valley	Wollondilly
Bogan	Federation	Lithgow	Shellharbour	Yass Valley
Bourke	Forbes	Liverpool Plains	Shoalhaven	
Brewarrina	Gilgandra	Lockhart	Singleton	
Broken Hill	Glen Innes Severn	Maitland	Snowy Valleys	
Byron	Goulburn Mulwaree	Mid-Coast	Tamworth Regional	
Cabonne	Greater Hume Shire	Mid-Western	Temora	
Camden	Gundagai	Murray River	Tenterfield	

Table 6b. Classification of NSW LGAs into five clusters or stratum: strata two through four

Cluster 2	Cluster 3	Cluster 4
Botany Bay	Canterbury-Bankstown	North Sydney
Burwood	Cobar	Sydney
Canada Bay	Coffs Harbour	Waverley
Georges River	Cumberland	
Inner West	Fairfield	
Lane Cove	Griffith	
Mosman	Hornsby	
Northern Beaches	Hunters Hill	
Parramatta	Kiama	
Randwick	Ku-ring-gai	
Rockdale	Liverpool	
Ryde	Moree Plains	
Strathfield	Newcastle	
Willoughby	Port Macquarie-Hastings	
Woollahra	Snowy Monaro Regional	
	Queanbeyan-Palerang	
	Sutherland Shire	
	Tweed	
	Wagga Wagga	
	Wollongong	

5.4.1 Selected LGAs

The LGAs selected to be included in the kerbside recycling survey are detailed in the Sampling Plan and are selected based on the methodology outlined above.

5.5 Within LGA stratification

As noted previously a sample of 100 household bins in an LGA would be adequate to obtain a reasonable level of precision, in the order of plus or minus 10 per cent. Larger sample sizes would not contribute substantially to the accuracy of the overall estimates for NSW.

Within each LGA, a representative sample of separate dwellings and units is required. This could be achieved in two ways. First, separate samples of separate dwellings and units could be made and estimates weighted by the proportion of each. Second, a stratified sample (with strata defined by separate households and units) could be taken with the proportion in each stratum.

The second option will have the least cost and may be more practical in rural areas due to the limited number of units. In either case a random sample of households will need to be taken in each stratum within each LGA. The logistics of sampling will vary between LGAs and plans will likely need to be adapted to meet on the ground requirements. The coordinator of the auditing process is likely to be best positioned to construct the individual LGA designs.

Conclusion: The random sample of households within each selected LGA should be stratified on the basis of the proportion of separate dwellings and units. The households to be sampled and the timing of this sampling should be developed by the coordinator of the kerbside auditing.

5.5.1 Revision of the kerbside Sampling Plan

Data from the initial and subsequent kerbside sample will be the primary input into a revised sample plan and the longer-term Sampling Strategy. Data from MRFs and regional redemption centres is also likely to be important. Ensuring that the ongoing design allows tracing changes in the composition of eligible containers in household recycling as well as maintaining coverage of the inputs to one or more MRF catchments is a large part of the longer-term Sampling Strategy.

A key aspect of this is the selection of when and which LGAs to rotate out of and into the sample and the households to be sampled. There is a complex trade-off to consider. Less frequent rotations provide a better perspective on changes over time while limiting the geographic spread of the sample. On the other hand, more frequent rotations ensure that the sample units that are selected are efficient and representative. This trade-off needs to be considered in the light of the data being collected.

Conclusion: The initial revision of the kerbside Sampling Plan and the longer-term kerbside Sampling Strategy should be developed over the first year of the CDS under the guidance of a statistical consultant with specific expertise in sampling design.

6 Calculating eligible container factors

Prior to estimation of the parameters of interest, the data need to be checked for the presence of outliers and coding errors. The container counts obtained from the MRF samples should be used to produce direct estimates of the number of eligible containers per tonne of recyclate for the period covered by the sample. For each material type (excluding glass) this will be the total number of eligible containers contained in the sampled units divided by the total tonnage of these sampled units. To determine the eligible container factor for glass, the total number and weight of eligible containers at the kerbside will need to be coupled with the contamination level data from the MRFs.

7 METHOD 2 – DIRECT COUNTING

Direct counting is a census of materials and therefore falls outside the scope of a Sampling Plan. The key issue is that the data from direct counting is cross validated using the data obtained from the MRF and kerbside sampling. This needs to be done in a representative way in terms volume of throughput processed and the location of the kerbside audits.

The scope of the monthly audits specified in Section 6.5 of the MRF Protocol should reflect the risks faced by stakeholders of misreporting. There is no one size fits all as this will largely depend on the volume counted within a month. The auditing of machine counts should take into account the technical specifications of the technology such as counting rates, error tolerances and engineering recommendations regarding the frequency of recalibration. The error tolerances and throughput can be compared to tolerances being set for MRFs that are being sampled to place requirements into perspective. Hand counting is likely to be very low volume and minimal audits are likely to be required.

It should be noted that all MRFs, including those who intend to claim using the direct count method will be required to complete an annual recycling statement and engage a suitably qualified independent auditor to complete an annual assurance audit in accordance with Section 9 of the MRF Protocol.

8 STOCKTAKE TO DETERMINE BASELINE

While the CDS commences on a specified date, recyclable material will be received and processed at MRFs before, during, and after this date. An important component of this Sampling Strategy is recognising and addressing the fact that there will be material stored at MRFs that was received and processed prior to the scheme commencement date. To address this, MRF operators are required to undertake a stock on hand survey on 30 November 2017, and the Scheme Coordinator may also engage an independent surveyor to conduct site visits at MRFs to measure the quantity of material stockpiled prior to scheme commencement to assist in determining a baseline.

ATTACHMENT A: THE MONTE CARLO SIMULATION MODEL

The overall design of the Monte Carlo simulations is based on:

- generating a universe of physical samples for each material type of a fixed weight and a variable number of eligible containers per bin;
    - where random samples are taken from the universe to represent samples taken at a MRF at random points in time;
- starting with a small sample size and taking repeated samples from this population to obtain the percentiles corresponding to the confidence level that has be set.
- If either percentile exceeds the specified error margin, increase the sample size, and repeat until the error margins are within tolerance.

The procedure can be used for individual material type or for all material types. In the latter case, the sample is stratified by the expected level of output of each material from all MRFs. The size of the population and the number of repeated samples were set on the basis of repeated trials to ensure the results were stable.

The main output from the simulations was the variance of the number of eligible containers. This was derived from household bin data which accounts for about 80 per cent of inputs into a MRF, adjusted to reflect an assumed direct participation rate in the CDS at redemption centres. There is a sequence of four steps for each material type on a per tonne basis.

1. Randomly generate the participation rate for each member of the population to calculate the correlation between the weight of beverage containers and weight of eligible containers between samples.
2. Use a copula based on the correlation above to generate a multivariate sample of beverage container weights and the proportion of eligible container proportions:
    a. weight of containers from a normal distribution; and
    b. proportion by weight of eligible containers from a binomial.
3. Randomly generate an average eligible container weight for each sample.
4. Calculate the number of eligible containers in each sample as the product of the weight of beverage containers and the proportion of eligible containers divided by the average container weight.

The means and standard deviations of weights, as well as the number of trials for the binomial distribution are intended to reflect the total number of containers in a sample taken a MRF (one cubic meter). For example, a sample might contain over 2,000 containers. The effective sample size of containers was calculated as an assumed percentage of this sample of containers to account for the potential clustering of samples within a MRF. This inflated the standard errors of the weights and reduced the number of binomial trials.

The common parameter setting for all materials are:

- minimum sample size = 30;

- mean CDS redemption rate = 50%;
- clustering reduction in effective sample size = 20%;
- confidence level = 95%;
- population size = 300,000; and
    - number of Monte Carlo trials 3,000.

The assumptions used to calibrate the kerbside data and those for the post CDS simulations are Shown in Table A.

Table A. Parameter assumptions for the Monte Carlo Simulations (weights are in kg)

Material	Pre CDS W_{BC}	p_{EBC}	Post CDS W_{BC}	p_{EBC}	w_{EBC}	RSE w_{EBC}	RSE W_{BC}	Sample Weight
Aluminium	0.87	0.99	0.77	0.98	0.014	10%	15%	40
PET	0.67	0. 82	0.54	0.69	0.030	25%	15%	28
HDPE	0.45	0.12	0.45	0.06	0.040	25%	20%	26
Mixed Segregated	0.39	0.56	0.32	0.39	0.035	25%	20%	38
Mixed Unsegregated	0.39	0.56	0.32	0.39	0.030	50%	25%	40