Loss recovery protocol: Technical description

Prepared for

Encorp Atlantic/Atlantique

Prepared by

Jacques Allard, Ph.D.

jacques.allard@gmail.com

Atlantic Statistical Analysis Inc.

May 12th, 2023

History and objectives

In 2009, Encorp Atlantic Inc. estimated that, system-wide, redemption centres (RCs) inflated container counts on average by approximately 5%, incurring a meaningful loss for the government. Encorp Atlantic Inc. developed and implemented a quality control (QC) program with two goals: to reduce or eliminate the system-wide loss and to reduce the count errors of the RCs. Following the implementation, the loss was rapidly reduced, achieving the first goal of the program: for example, for the 12 months ending July 31st, 2022, the system-wide container count error is estimated to be less than ±0.1%. The second goal of the program was partially achieved: the count errors of many RCs were reduced.

On April 1st, 2023, the successor of Encorp Atlantic Inc. a new not-for-profit producer responsibility organization in the name of *Encorp Atlantic/Atlantique*, will add to its responsibility the collect of alcoholic beverage containers. This addition is projected to lead to a total number of containers approximately 75% greater and to a significantly different mix of products, increasing the proportion of aluminum cans and of glass containers.

These changes require that the quality control protocol be revised. This revision is the subject of this document.

Errors and corrections

The term “***error***” refers to the signed difference [declared value] – [true value]. The error is positive if the declared value is larger than the true value (over-reporting); it is negative if the declared value is smaller than the true value (under-reporting). The expression “***relative error***” refers to the quotient of the error over the ***true value***, often expressed as a percentage.

The term “***correction***” refers to the change that should be applied to the declared value to make it equal to the true value. The expression “***relative correction***” refers to the quotient of the error over the ***declared value*** often expressed as a percentage.

The mathematical relationship between the relative values is the following, where the relative correction and the relative error are expressed in fractions.

$$C_{relative\ to\ the\ recorded\ value} = -\frac{E_{relative\ to\ the\ true\ value}}{1 + E_{relative\ to\ the\ true\ value}}$$

We apply the terms “error” and “correction” to individual bag counts and to averages or totals over time, products, product classes (see below), for a single RC, a group of RCs or all RCs combined.

Numerical example

Suppose that the declared value is 1000 and the true value is 800.

The error is 1000 – 800 = +200 and the error ***relative to the true value*** is 200/800 = +0.25 = +25%.

The correction is 800 – 1000 = –200 and the correction ***relative to the declared value*** is –200/1000 = -0.20 = –20%.

Applying the equation above the box:

$$C_{relative\ to\ the\ recorded\ value} = -\frac{E_{relative\ to\ the\ true\ value}}{1 + E_{relative\ to\ the\ true\ value}} = -\frac{+0.25}{1 + 0.25} = -0.20$$

The difference between the two percentages is due to the denominators: according to the definitions, the first percentage is relative to the true value and the second percentage is relative to the declared value.

Definitions

Bags, tubs, sorts, sort classes

We use the term "***bag***" to refer to bags or tubs. Tubs are used for glass containers.

Encorp uses the term "***sort***" to describe a product or a group of products. Possible examples of sorts are aluminum cans, clear PET containers of non-alcoholic beverages, etc. The definition of the sorts can change over time due to several factors including changes in recycling technology and accounting requirements.

A container ***sort-class*** or a ***class*** is a set of sorts. A class may contain only a single sort. Containers in a class are typically mingled at certain steps of the recuperation process. Several definitions of classes may be required for various purposes. For example, aluminum containers of alcoholic beverages and of non-alcoholic beverages may form a single class from a quality control point of view but different classes from an accounting point of view.

The "declared value" is the number of containers indicated on a label affixed to each bag.

Errors, bias, variability, accuracy

Measurement ***errors*** are generally broken down into two components:

- ***Bias*** is the average of the errors. If the bias is 0, the average declared count and the total count correspond to the true values. We use the term "trueness" as opposite of bias.
- ***Variability*** is the variation of the errors around their average. Highly variable errors are an indication that an RC does not count bagged containers carefully, even if its bias is low. High variability makes QC difficult and expensive. We use the term "precision" as opposite of variability.

Bias and variability are separate concepts. For example, the counts of an RC that counts containers perfectly but mixes the tags between the bags would have 0 bias but a high variability.

We use the term "***accuracy***" to describe the combination of trueness and precision.

In this document, the bias is defined mathematically as the arithmetic mean of the errors and the variability as the standard deviation of errors.

Quality of statistical estimation

We describe the quality of estimates using the half-length of the 95% confidence interval, i.e., $\pm 2\times$[standard error] for large samples. Estimates obtained will be within $\pm$[half-length of the 95% confidence interval] of the estimated value 19 times out of 20.

The quality of estimates depends on the variability of the RCs' counting errors. We report the quality of estimates for RCs' historical variability at the median ("typical RCs") and at the 75th percentile ("RCs with more variable errors") of standard deviation of the RCs' errors.

Rounding

In this document, numbers have been rounded. Therefore, small discrepancies may appear in sums, etc.

Historical data

Encorp Atlantic Inc.'s historical data and projected volume were used to optimize the sampling protocol. Given changes in sorts and quality control, emphasis was given to data from August 2021 to July 2022.

For the purpose of this analysis, containers were classified in 4 sort classes: Aluminum, PET, Glass and Others. Projected alcohol beverage bag and tub volumes, and container volumes, by sort class, and quality control capacity were supplied by Encorp Atlantic Inc.

Quality control capacity and other constraints

The projected number of containers and bags received, by sort, for the fiscal year 2023-2024 (beginning April 1st, 2023) is shown in Table 1.

Table 1 – Projected container and bag volume for fiscal year 2023-2024.

Sort	Projected number of containers received	Projected average number of containers per bag or tub	Projected number of bags received
Aluminum	193,000,000	1,900	103,000
PET	110,000,000	1,000	115,000
Glass	18,000,000	700	25,000
Other	12,000,000	1,200	10,000
Total	**333,000,000**		**253,000**

Non-glass containers and glass containers will be collected by Contractors A and B, respectively. Each contractor will oversee the quality control for the bags it collects.

The planned quality control capacity is between 4% and 6% of the bags. Using the middle value, 5%, the number of bags to be verified, by sort, for fiscal year 2023-2024 is shown in Table 2.

The following constraints are imposed:

- The proportion of bags to be verified by each Contractor should be similar.
- Verified glass tubs will also be used to estimate the average weight of the glass containers. The minimum annual number of glass tubs required to obtain the desired accuracy of this weight has been estimated at 500 tubs.

The New Brunswick network of RCs includes approximately 70 locations. The number of containers and bags shipped by each centre is anticipated to vary from 100,000 to more than 15,000,000.

Table 2 – Projected number of bags verified for fiscal year 2023-2024, based on a 5% QC ratio, with the corresponding number of containers.

Sort	Projected number of bags verified	Projected number of containers verified
Aluminum	5,150	9,600,000
PET	5,750	5,500,000
Glass	1,250	900,000
Other	500	600,000
Total	**12,650**	**16,600,000**

Protocol

The QC sampling program will be divided into two subprograms:

- Monitoring QC: The objectives of Monitoring QC include the measurement of overall compliance and the identification of compliant and non-compliant RCs. When an RC is under Monitoring QC, a random (or systematic) sample of 50 bags per 6-month period is verified. The sample size was selected to provide an anticipated accuracy of ±1.7% or better for the typical RC.
- Accelerated QC: The main objective of Accelerated QC is to measure the counting error of selected RCs and to assess payment adjustments. When an RC is under Accelerated QC, a random (or systematic) sample of 225 bags per 6-month period is verified. The sample size was selected to provide an anticipated accuracy of ±0.8% or better for the typical RC.

Table 3 shows the anticipated QC activity for each level of intensity based on 70 active RCs.

Table 3 – Anticipated QC activity by level of intensity.

	Monitoring QC	Accelerated QC	Total
Number of RCs at a given time by monitoring status	55	15	
Number of standard 6-month QC runs annually	110	30	
Number of bags per 6-month period	50	225	
Total number of bags verified monthly	458	563	1,021
Total number of bags verified annually	5,500	6,750	12,250

Monitoring QC

The main objectives of the Monitoring QC program are the measurement of overall compliance and the identification of ***possibly*** non-compliant RCs.

For each RC in the Monitoring QC program, a random or systematic sample of 50 bags will be verified for each 6-month period, i.e., approximately 8.3 bags per month. The sampling will be stratified by sort class and the 50 bags will be allocated to minimize the standard error of the estimate while respecting the constraints described above. The sample allocation will be as follows:

Table 4 – Sample allocation for Monitoring QC (see below for very small RCs).

Aluminum	PET	Glass	Others	**Total**
20 bags	20 bags	5 bags	5 bags	**50 bags**

The anticipated accuracy of the Monitoring QC for 50 bags over 6 months (or over any continuous time period) is ±1.7% and ±2.5% for typical RCs and for RCs with more variable errors, respectively. The accuracy for 100 bags over any continuous time period is ±1.2% and ±1.8% for typical RCs and for RCs with more variable errors, respectively.

The results of the Monitoring QC sampling can be compiled on a continuous basis and/or after set periods of time.

Note 1: The default sampling rate of 8.3 bags per month and the 6-month time period may be changed to respond to special circumstances. For example, rapid changes in the beverage market in a specific region may incite a momentary increase in the monitoring intensity of RCs in that region.

Note 2: For small volume RCs, a lower sample size or a lengthening of the assessment period may be selected for practical reasons and from a cost/benefit point of view. If the assessment period is kept at 6 months, the following sample allocation is suggested for RCs delivering less than 500 bags and tubs annually.

Table 5 – Sample allocation for the Monitoring QC for RCs delivering ***less than 500 bags and tubs*** annually.

Aluminum	PET	Glass	Others	**Total**
10 bags	10 bags	3 bags	2 bags	**25 bags**

Accelerated QC

The main objectives of the Accelerated QC program are the assessment of the counting errors of selected RCs and the estimation of payment corrections with a sufficient accuracy.

For each RC in the Accelerated QC, a random or systematic sampling of 225 bags will be verified over each 6-month period, i.e., approximately 37.5 bags per month. The sampling will be stratified by sort class and the 225 bags will be allocated to minimize the standard error of the estimate while respecting the constraints described above. The sample allocation will be as follows:

Table 6 – Sample allocation for Accelerated QC (see below for very small RCs).

Aluminum	PET	Glass	Others	**Total**
100 bags	95 bags	20 bags	10 bags	**225 bags**

The anticipated accuracy of the Accelerated QC for 225 bags over 6 months (or over any continuous time period) is ±0.8% and ±1.2% for typical RCs and for RCs with more variable errors, respectively. The accuracy for 100 bags over any continuous time period is ±0.6% and ±0.8% for typical RCs and for RCs with more variable errors, respectively.

Note 1: The default sampling rate of 37.5 bags per month and the 6-month time period may be changed to respond to special circumstances.

Note 2: For small volume RCs, a lower sampling rate or a lengthening of the assessment period may be selected for practical reasons. If the assessment period is kept at 6 months, the following sample allocation is suggested for RCs delivering less than 1,000 bags and tubs annually.

Table 7 – Sample allocation for the Accelerated QC for RCs delivering ***less than 1000 bags and tubs*** annually.

Aluminum	PET	Glass	Others	**Total**
50 bags	50 bags	15 bags	5 bags	**125 bags**

QC resources allocation

Table 8 shows the allocation of the QC capacity by sort-class and QC subprogram. The proportion of bags or tubs verified is projected to be 4.9% for non-glass containers (Contractor A) and 4.6% for glass containers (Contractor B).

Table 8 – Allocation of the QC capacity by sort-class and QC subprogram.

	Monitoring		Accelerated			
Product class	Single run (bags)	Annual subtotal (bags)	Single run (bags)	Annual subtotal (bags)	Annual total (bags)	Proportion of bags received (%)
Aluminum	20	2,200	100	3,000	5,200	5.0%
PET	20	2,200	95	2,850	5,050	4.4%
Glass	5	550	20	600	1,150	4.6%
Others	5	550	10	300	850	8.5%
Total	50	5,500	225	6,750	12,250	4.8%
Non-Glass	45	4,950	205	6,150	11,100	4.9%

Stratification

In sampling, stratification refers to the partition of the statistical population into sub-populations called strata. Stratification may be required for accounting reasons (e.g., containers with different redemption values or following different recuperation streams).

For fixed quality control resources, stratification increases the accuracy of the estimates if the strata are more homogeneous than the whole statistical population (e.g., if, say, errors on counts of aluminum containers are, on average, larger or smaller than those on PET containers).

In some cases, stratification is necessary to obtain correct estimates. For example, using a simple annual average error may lead to incorrect estimates if winter and summer errors are very different.

If the sampling is stratified, mathematical equations must be adapted accordingly.

Sampling procedure

For the present quality control protocol, container sorts handled differently should be in different container classes, except for sorts with very small numbers. The sampling procedure should be applied to each container class separately.

The selection of the bags ***MUST*** be done according to one of the following two procedures.

Random selection: Using a pseudo-random number generator, a sequence of numbers (e.g., 15, 29, 39, 71, 86, 294, 328, …) is created in a way that gives the required sampling rate. These numbers are taken as the order of arrival of bags to be verified (e.g. the 15th, 29th, 39th, 71st, 86th, 294th, 328th … bags, in order of arrival, will be verified).

Systematic selection: Once the sampling ratio is established and expressed as a fraction 1/A (e.g. 1/40), a number i_0 is chosen randomly between 1 and the denominator A (say i0 = 28). Then, in sequential order of arrival, the i_0^{th} is selected for QC and every following Ath bag (e.g. the 28th, 68th, 108th … bags). Systematic sampling is easier to implement.

"Representativity": The selected bags must be IN ALL ASPECTS like the non-selected bags. That means, among other things, that **NO** factor other than the process just described is involved in the selection and that the selected bags must **NOT** be subjected to any particular intervention after the selection process.

Data Collection: For each selected bag, the RC Count and the Quality Control Count of containers in the correct sort are recorded, as well as other information, including the RC, pick-up date, QC date, etc.

Bag weight: Systematically collecting bag weights is recommended. The relationship between the weight and the number of containers in bags can improve the monitoring process. Glass tub weights will be required for the estimation of the average glass container weight.

Transition between Monitoring and Accelerated QC

It is essential that RCs NOT be aware of their QC status. It is recommended that service suppliers and their staff be contractually required to preserve the confidentiality of the QC activities.

The decision to transition an RC from the Monitoring QC program to the Accelerated QC program can be based on several factors including but not limited to the results of the Monitoring QC.

We recommend that one or more RCs be transitioned to Accelerated QC on a random basis.

Accuracy award

We use the following terms:

- Trueness refers to the closeness of the mean of the container counts to the true mean of the container counts. An RC's count will be true if the average of its count errors is close to 0. The opposite of trueness is bias.
- Precision refers to how similar the errors are. An RC's counts will be precise if all bag-level errors are close to each other, say, for example, between -10% and -7%. An RC's counts will be imprecise if errors vary a lot, say, for example, from -50% to +100%. The opposite of precision is variability.

Biased counts lead to over- or under-reimbursements. Imprecise counts mean that an RC poorly counts containers in each bag and makes quality control difficult and expensive.

The accuracy award will be given to RCs whose counts for all sort classes confounded, ***over a period of 12 months***, satisfy the following two conditions:

Condition A – Trueness: The estimated average error is ***between –1% and +1%***.

Condition B – Precision: The estimated relative standard deviation of the errors is ***less than 5%***.

The computation of the estimated average error and of the estimated relative standard deviation of the errors is described in the section “Mathematical equations”. We recommend that the computation be stratified according to container classes and, if feasible, quarters.

The definition of the “variability of the errors” is based on the definition of the coefficient of variation.

The time period and the two thresholds may be adjusted to respond to market composition, technological changes, overall QC intensity and other factors.

Mathematical equations

Definitions

We use the following symbols.

Definitions and conventions:

- $\bar{A}$, read as “A bar”, stands for the mean of the elements of A.
- $\hat{B}$, read as “B hat”, stands for an estimate of a quantity B.
- We use lower case letters for indices and the corresponding upper-case letter for the set of indices.

The following definitions refer to data from a single RC.

h = A container class, $h \in H$.

p = A time period (e.g., a month of a year, a quarter), $p \in P$.

N_{hp} = Number of bags of class h delivered during the period p.

n_{hp} = Number of bags of class h verified during period p.

j = A bag number, $j \in J$ for all bags, $j \in QC$ for verified bags.

r_{hpj} = The RC count of containers for bag j of a class h, period p.

q_{hpj} = The QC count of containers for bag j of a class h, period p.

$e_{hpj} = r_{hpj} - q_{hpj}$ = Count error in the bag.

$R_{hp} = \sum_{j \in J} r_{hj}$ = Total number of containers ***declared*** for all bags of class h reported during period p. In the context of single period, p is omitted.

T_{hp} = True total number of containers ***delivered*** for all bags of class h reported during period p. In the context of single period, p is omitted.

$\hat{T}_{hp}$ = Estimate of the true total number of containers ***delivered*** for all bags of class h reported during period p. In the context of single period, p is omitted.

Single class, single period computations

The following equations apply to containers of a single class h during a single period p.

Average number of containers in bags verified:

$$\overline{q_{hp}} = \frac{1}{n_{hp}} \sum_{j \in QC} q_{hpj}$$

Average error (positive for overcount, negative for undercount) for bags verified:

$$\overline{e_{hp}} = \frac{1}{n_{hp}} \sum_{j \in QC} e_{hpj}$$

Estimated total error (positive for overcount, negative for undercount):

$$\widehat{E_{hp}} = N_{hp} \overline{e_{hp}}$$

Estimated total correction (negative for overcount, positive for undercount):

$$\widehat{C_{hp}} = -\widehat{E_{hp}} = -N_{hp} \overline{e_{hp}}$$

Estimated total true number of containers delivered:

$$\widehat{T_{hp}} = R_{hp} - \widehat{E_{hp}} = R_{hp} + \widehat{C_{hp}}$$

Estimated relative error for the total number of containers:

$$\widehat{E.rel}_{hp} = \widehat{E_{hp}} / \widehat{T_{hp}}$$

Estimated standard deviation of the errors:

$$\widehat{s_{hp}} = \sqrt{\frac{1}{n_{hp} - 1} \sum_{j \in J} \left(e_{hpj} - \overline{e_{hp}}\right)^2}$$

Estimated ***relative*** standard deviation of the errors:

$$\widehat{s.rel}_{hp} = s_{hp} / \overline{e_{hp}}$$

Estimated standard error of the estimate of the average error, average correction and true total number of containers delivered:

$$\widehat{SE(\widehat{E_{hp}})} = \widehat{SE(\widehat{C_{hp}})} = \widehat{SE(\widehat{T_{hp}})} = N_{hp} \sqrt{\left(1 - \frac{n_{hp}}{N_{hp}}\right) \frac{s_{hp}^2}{n_{hp}}}$$

Estimated relative standard error of the estimate of true total number of containers delivered:

$$\widehat{SE(\widehat{T_{hp}})} / \widehat{T_{hp}}$$

Aggregating results over container classes and/or periods

We present the formulae to aggregate the results over container classes. For the aggregation over periods, the role of the class ("h") and the period ("p") are inverted. For aggregation over both classes and periods, summations are taken over all values of h and p.

The following equations apply to a single period p.

Estimated total error (positive for overcount, negative for undercount):

$$\widehat{E_p} = \sum_h \widehat{E_{hp}}$$

Estimated total correction (negative for overcount, positive for undercount):

$$\widehat{C_p} = -\widehat{E_p}$$

Estimated total true number of containers delivered:

$$\widehat{T_p} = \sum_h \widehat{T_{hp}}$$

Estimated relative error for the total number of containers delivered:

$$\widehat{E.rel_p} = \widehat{E_p}/\widehat{T_p}$$

Estimated standard error of the estimate of the average error and average correction:

$$\widehat{SE(\widehat{E_p})} = \sqrt{\sum_h \widehat{SE(\widehat{E_{hp}})}^2}$$

Accuracy Award conditions

Estimated relative average error:

$$\hat{\mathrm{E}} = \left(\sum_{h,p} \widehat{E_{hp}}\right) \Big/ \left(\sum_{h,p} \widehat{T_{hp}}\right)$$

Estimated relative variability of the errors:

$$\hat{s} = \sqrt{\left(\sum_{h,p} N_{hp} \widehat{s_{hp}}^2\right) \Big/ \left(\sum_{h,p} N_{hp}\right)} \Big/ \left(\sum_{h,p} \widehat{T_{hp}}\right)$$

Note: This is a weighted average of standard deviations applied only for the Accuracy Award. The estimated variability of the errors ***cannot*** be used to compute the standard error of the estimated total error.

Interpolation

For a given sort class, when no QC data is available for a period, linear interpolation is used. For periods before the beginning or after the end of the data availability period, the first or last data will be repeated, respectively.

Theoretical basis

We are interested in estimating the true population total since it corresponds directly to the accounting requirements.

For a single sort-class and time period, the estimate of the mean error is the usual estimate of a mean for simple random sampling. The estimate of the true total number of containers is a “difference” estimator. This choice reflects the facts that, according to historical data, the dependency of the errors on the values is weak and that the dependency was inconsistent between sort-classes.

For several sort-classes and/or time periods, the estimated errors are aggregated, making the overall estimator a stratified estimator. For simplicity, we use the same equations whether the stratification is part of the sampling design or the data is post-stratified. We consider that the impact of this simplification is negligible.

For the monitoring and the accelerated programs, the sampling design is optimized within the operational constraints.

Sample size vs sampling ratio: Typically, the quality of an estimate obtained from a random sample depends only on the sample size. Consequently, in our situation, the quality of an estimate obtained from a random sample of 50 or 225 bags will have the same quality for a small or a large RC. Contrary to most people's intuition, the sampling ratio impacts the quality only if it is very high e.g., > 25%.

The following textbook is our general reference for this work:

Sharon L. Lohr, Sampling: Design and Analysis, 2nd edition, Duxbury Press (Dec 2009), 600 pp. ISBN-10: 0495105279.

Terminology partly follows ISO 5725-2:2019:

https://www.iso.org/obp/ui/#iso:std:iso:5725:-2:ed-2:v1:en (viewed on 2023-05-06)

https://www.iso.org/obp/ui/#iso:std:iso:5725:-2:ed-2:v1:fr (viewed on 2023-05-06)