**Booster Guideline for GCS**

**Introduction:**

A Booster sample is often included in GCS across markets as it is the most cost effective way to achieve a threshold level of sample for small brands – both BAT & competition. The alternative approach to this would be to increase the random sample to get the required number of smokers for brand(s) of interest. While the alternative approach is statistically optimum and sound it lacks effectiveness due to much higher requirement of resources i.e., time & budget. Whilst the boosters are more efficient way of achieving the data for smaller sub-groups – there is always a trade-off between efficiency/costs and quality and usability. In order to be statistically sound the boosters come with a list of strings attached regarding their application and the way the data may be analyzed.

The selection of booster sample should be done in a way that it does not affect and hinder the robustness & representative-ness of the GCS sample. Due to this it is recommended only to boost a demo/geo-demographic group or a brand group which has already a decent base representation in the random sample (their representation in random sample may not be sufficient to carry out analysis but sufficient to check the profile for weighting purpose). This is very important as the inclusion of baseless boosters will seriously complicate the weighting process as well as impacts the overall robustness and representative-ness of the data. In the worst case the booster that was intended to provide us with additional information and improve our data reach and quality may end up distorting the base data that we are looking at.

Depending on the type and nature of the boosters there are also limitations to their usage. Especially if we enter the area of low base boosters – they do not provide us a representative sample and therefore cannot be used for the full scale of analysis that the proper base boosters and random sample provide.

During the quality deep dive into our GCS data and processes it has been observed that there are a lot of issues with the way boosters have been set up in the GCS, weighted and added into the Batprobe Database as well as used in analysis. Common issues include:

* Booster samples often drawn from brands that do not have any representation or a too small representation in the random sample
* Booster samples drawn based on additional behavioral elements like Occasional Usage, Recent Trial etc.
* Brand Booster data without sufficient random base is used for a wide range of analysis like profiling, trends, switching – leading us to do analysis which is not statistically robust and in the worst cases completely misleading.
* Boosters without sufficient base have been weighted into the GCS data – causing the main data to suffer and having a negative impact on the robustness of our basic KPI’s.

The purpose of this document is to outline the new Global Guidelines for the Booster sample selection and the usage of booster sample in GCS analysis. The new guidelines are applicable with immediate effect and will be applied as of the Wave 2 2010 onwards.

Global Guidelines **for Booster Sample:**

Whilst there are rules and limitations to the use of boosters they do still remain an integral part of the GCS framework and if correctly used they provide a powerful and efficient way of improving the random sample and widening the analysis that might be otherwise obtained. On broad terms Boosters are classified into three groups:

1. Demographic over-sampling
2. Brand Boosters
3. Tag on sample

The following sections introduce these different types of boosters and the guidelines for their application and usage.

1. **Demographic over-sampling :**

Boosters for demographic sub-groups are called Demographic over-sampling. Examples of this type are boosters such as Female, ASU30, and Capital City etc. This type of over-sampling should be done **randomly with pre-defined quotas**. Integration of such demographic over-sampling to the main random sample is relatively straightforward with a simple cell weighting. Weighting matrix for the over-sampled group should be taken from the weighted Random sample or 1+ manufactured cigarette smokers from Incidence database. The weighting parameters should be only demographics of the smokers and no brand weights should be used.

Demographic over-sampling is a good way to improve the overall sample where it is known to be difficult to reach sufficient amounts of consumers from the random base.

In order to boost a demographic sub-group there needs to be minimum of 25 such smokers in the random base to enable us to see the proportion of these people in the random sample and to do the weighting properly.

1. **Brand boosters :**

The brand boosters should always be done at the family level e.g., Pall Mall Red KS (including any Pall Mall Red KS variants in whatever pack type, tipping, pack size). This is the optimum level for analysis purposes and gives the best balance for the booster to be representative of a meaningful entity.

The basic rule for brand boosters is that the brand family for which the boosting will be done has to have minimum 25 samples in Random smoker base. This is required to correct the demographic profile of the brand family at the weighting stage. Demographic adjustment of the boosted brand is a must otherwise the booster would negatively impact the overall quality of the data.

Brand boosters can also be done at a combined brand family X demographic group e.g., ASU30 Pall Mall Red KS – however the minimum sample size criterion needs to be met i.e., at least 25 sample size of ASU30 Pall Mall Red KS in the Random smoker base.

In case the boosted brand family does not have required representation in the Random smoker base i.e., has less than 25 sample in the Random smoker base – then the booster sample will not be weighted into the main results or integrated into the Batprobe database and it does not qualify as a booster but rather a Tag-On sample (explained later in the document). Thus it is still possible to collect such extra Tag-On samples but the analysis for them will have to be done unweighted and separated from the main data. The coordinating agency responsible for the GCS can provide such additional service where necessary.

The size of the extra booster sample for a given brand family can be at the most the same size as the random sample of the same brand family. As an example – if Pall Mall Red KS has 27 regular smokers in the random base the maximum additional booster that can be applied is 27 booster smokers giving in total 54 smokers for the analysis. The reason for this is that the efficiency of the weighting and the overall representative-ness of the GCS data goes down if booster sample size is higher than random sample of the brand family. This is a strict rule from statistical perspective – if there would be a need to have a higher booster collected it needs to be reviewed from the perspective of a Tag-On sample.

In order to ensure high quality and sound weighting – Brand Boosters should be collected from the Regions/Cities that mirror the profile found in the random sample. It is essential to do proper planning for the Booster sampling and to aim as close as possible to the geographical dispersion of the same family in Random sample.

1. **Tag on Sample :**

In addition to the Demographic Over-Sampling and the Brand Family boosters with sufficient base – it is possible to use the GCS fieldwork to collect additional Tag-On samples with the purpose of answering customized marketing questions for some small/niche brands in a cost efficient way.

This section provides the guidelines for Tag-On samples which have also strict limitations for the type of analysis that may be conducted and the way the data should be treated.

There are 5 types of recommended Tag-On samples:

1. **Type 1 :** Brand booster with less than 25 base for the brand family in the Random smoker base
2. **Type 2:** Brand X demographic booster with less than 25 samples for the combination in the Random smoker base.
3. **Type 3:** Recent switchers to a particular brand family. For example, switched during last 3 months or 6 months into Pall Mall Red KS.
4. **Type 4 :** Tried the brand family in last three months
5. **Type 5 :** Occasionally smoke the brand family

Note that it is recommended to recruit the tag-on sample also at the level of a brand family.

Tag on samples will not be weighted nor integrated with the GCS main sample and will not be available in the Batprobe database. Reason for this is that integration of such data with GCS database with weights will impact negatively the overall robustness/ representative-ness of the GCS data and that is not an acceptable trade-off for the Group.

Note that Brand Values Segments can be created in the Tag on sample using the same Segident equation. So the analysis of Tag on sample can be done within brand values segments however unweighted and restricted to the list given in the next section i.e., Guideline for GCS Analysis with Booster.

As Tag on sample will not be included in the GCS Batprobe database – one may decide to administer only selected sections from GCS questionnaire for the Tag-on sample. Clear objectives and rationale need to be put behind the recruitment of such sample in GCS framework. **Tag on sample needs to be approved by the Global GCS Team and Global Oracle Manager – similarly as with the GCS sampling and Questionnaire.** The management and analysis of a tag-on sample will be the responsibility of the Coordinating Agency for each particular market or in case the initiative is for an Area or Region – respective Coordinating Agency will do the management.

**Guideline for GCS Analysis with Booster:**

As mentioned in the beginning of this document – Boosting does not come without strings attached and some limitations of use compared to proper random sampling. Whilst cost efficient boosting is still often expensive as well as aimed to get us information for critical strategic brand families thus is it of utmost importance that the data is statistically sound in relation to the intended usage.

This section provides guidelines how the different types of Booster samples and Tag-On sample may be analyzed. It is the responsibility of the Coordinating Agency to follow these guidelines in all the booster analysis done and to inform BAT of any issues/doubt of statistical robustness.

Key principles for the Booster Analysis:

1. Demographic over-sampling should always be integrated with the Random sample with demographic weight adjustments. This type of boosters, after appropriate weighting, can be used for all different types of GCS analysis thus there are no limitations to the use.
2. Brand boosters (with minimum 25 regular smokers of the brand family in Random base) will also be integrated with the Random sample after both demographic & brand share adjustment through weighting (Rim Weighting). This type of boosters, after appropriate weighting has been done can be used for all types of GCS analysis thus there are no limitations to the use.
3. Tag-on sample should always be left unweighted as this type of sample is purely purposive and no weight adjustment will be done due to lack of robust information from Random smoker base. Demographic profile analysis, switching analysis & trend analysis should **strictly not** be done with a tag-on sample. Due to the lack of robust base and weighting the information is very likely to be misleading at the best.

The analysis that can be done with a Tag-On sample is concentrated around Regular Brand Disposition and Imagery or some behavioral questions like purchasing behavior. It can give us indications of Loyalty and brand health but cannot be extended to measure for instance the CDF or to review trials etc... The reason being that these measures would heavily depend on the random data to be able to track development and movement across the waves. With the nature of the Tag-On samples this is not possible.

The following tables provide the list of analysis that can be done with GCS sample with valid boosters versus the limitations of Tag-on sample:

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Type of Analysis** | **GCS sample with valid Boosters** | **Tag on Sample (Unweighted)** |
| 1. | Regular Brand Share | **✓** | **×** |
| 2. | Demographic profile of Brand Franchise e.g., % Female, % ASU30 etc. within the brand franchise | **✓** | **×** |
| 3. | Demographic profile of Occasional smokers, Tried Smokers | **✓** | **×** |
| 4. | Consumer Disposition Funnel | **✓** | **×** |
| 5. | Brand Switching i.e., source of gain/ destination of loss | **✓** | **×** |
| 6. | Occasional Brand % | **✓** | **×** |
| 7. | Substitute brand % | **✓** | **✓** |
| 8. | Trial brand % | **✓** | **×** |
| 9. | Regular Brand Disposition | **✓** | **✓** |
| 10. | Regular Brand Imagery | **✓** | **✓** |
| 11. | Occasional Brand Imagery | **✓** | **✓** |
| 12. | Tried Brand Imagery | **✓** | **✓** |
| 13. | Regular brand quality problem (If included) | **✓** | **✓** |
| 14. | Purchase Behavior (Outlet type, Mode of Purchase i.e., SKU size) | **✓** | **✓** |
| 15. | Venues visited in last one month | **✓** | **✓** |
| 16. | Cig. Brand communication recall (if included) | **✓** | **✓** |
| 17. | Recall of Brand messages by Touchpoint | **✓** | **✓** |
| 18. | Trend Analysis on any measure | **✓** | **×** |

------0------