



## KNOWLEDGEPANEL SAMPLING AND WEIGHTING METHODOLOGY

Significant resources and infrastructure are devoted to the recruitment process for KnowledgePanel (KP) so that the resulting panel can properly represent the adult population of the US. This representation is not only achieved with respect to a broad set of geodemographic distributions, but also hard-to-reach adults – such as those without landline telephone or Spanish language dominant individuals – are recruited in proper proportions as well. Consequently, the raw distribution of KP mirrors that of the US adults fairly closely, barring occasional disparities that may emerge for certain subgroups due to differential attrition rates among recruited panel members.

For selection of general population samples from KP, a patented methodology has been developed that ensures the resulting samples behave as EPSEM (equal probability of selection method). Briefly, this methodology starts by weighting the entire KP to the benchmarks secured from the latest March supplement of the Current Population Survey (CPS) along several dimensions. This way, the weighted distribution of KP perfectly matches that of the US adults – even with respect to the above mentioned few dimensions where minor misalignments may result from differential attrition rates. Typically, the geodemographic dimensions used for weighting the entire KP include the following dimensions, with additional nesting of dimensions as well:

- Gender (Male/Female)
- Age (18–29, 30–44, 45–59, and 60+)
- Race/Ethnicity (Hispanic and non-Hispanic White, Black/African American, Other, and 2+ Races)
- Education (Less than High School, High School, Some College, Bachelor and beyond)
- Census Region (Northeast, Midwest, South, West)
- Household Income (<\$10k, \$10K to <\$25k, \$25K to <\$50k, \$50K to <\$75k, \$75K to <\$100k, \$100K to <\$150k, \$150k+)
- Homeownership status (Own, Rent/Other)
- Metropolitan Area (Yes, No)
- Hispanic Origin (Mexican, Puerto Rican, Cuban, Other Hispanic, and non-Hispanic)
- Language Dominance (non-Hispanic and English Dominant, Bilingual, and Spanish Dominant Hispanic)

Using the above weights as the measure of size (MOS) for each panel member, in the next step a PPS (probability proportional to size) procedure is used to select study specific samples. It is the application of this PPS methodology with the above MOS values that produces fully self-weighting samples from KP, for which each sample member can carry a design weight of unity. Moreover, in instances where the study design has required any form of oversampling of specific subgroups, such departures from an EPSEM design are corrected by adjusting the corresponding design weights accordingly with the CPS benchmarks serving as reference points.

### Study-Specific Final Weights

Once the study sample has been selected and fielded, and all the survey data are edited and made final, design weights are adjusted for any survey nonresponse as well as any under- or over-coverage imposed by the study-specific sample design. Depending on the specific target population for a given study, geodemographic distributions for the corresponding population are obtained from the CPS, the American Community Survey (ACS), or in certain instances from the weighted KP profile data. For weighting adjustments, an iterative proportional fitting (raking) procedure is used to produce final weights that will be aligned with respect to all study benchmark distributions simultaneously. In the final step, calculated weights are examined to identify and, if necessary, trim outliers at the extreme upper and lower tails of the weight distribution. The resulting weights are then scaled to the sum of the total sample size of all eligible respondents.