CHAPTER VII

SAMPLE SELECTION AND MAINTENANCE

INTRODUCTION

The purpose of this Chapter is to address a number of subjects related to the selection and maintenance of the HPMS sample and sample panels. The data reported for sampled roadway sections are the source of the condition, use, and operational information provided by the HPMS. Expanded sample data are used for apportionment of funds, for monitoring trends and impacts in performance data over time, and for analyses in support of national budgeting for highway improvements through the *Condition and Performance* reports to Congress. Selection and maintenance of an adequate, up-to-date HPMS sample should be a high priority, continuing activity.

Although developed for the HPMS standard sample, most of the following discussion applies equally for donut area supplementary sample panels; Appendix G contains additional discussion of the donut area sample panel.

SAMPLE PANEL CONSIDERATIONS

The HPMS sample includes the arterial and collector functional systems, excluding the rural minor collector system. While it is assumed that there is a "technically best" way to collect sample data, a sample design also must be simple and cost efficient to meet manpower and cost considerations; this involves tradeoffs. The required number of samples for HPMS are derived by formula from the normal dispersion characteristics of AADT values within a framework of preselected AADT groups (strata). The sample size requirements relate to the critical data element, AADT, whose values can be conveniently stratified. Information obtained from the existing sample or universe data in each State is used to optimize and maintain the sample panel.

Procedures for determining necessary sample size based on the analysis of existing data are described in this chapter. In order to obtain cost-effective, valid comparisons of system performance over time and to reduce technical effort, the sample was designed as a fixed sample panel. With a fixed sample panel, the same sections that are inventoried are then updated in future years on a cyclical basis. This means of obtaining data is efficient because:

1. The need for the periodic drawing of a complete new sample is eliminated.
2. The need to update or reinventory all data elements for every cycle is eliminated.
3. Only those data elements that change over time need be updated on a cyclical basis.

The length of the cycle is determined by the known statistical characteristics of individual elements, the intended use and accuracy needed, and the time and cost required to collect and report the data.

However, the use of fixed panel sections is not without disadvantages. These include:

- the possible loss of the sample's representativeness as the highway networks and traffic patterns change, and
- the inability to assess the correctness of the estimates by comparing them with those of a different sample.

Procedures have been developed to ensure the representativeness of the sample and FHWA has determined that the practicability of the fixed panel approach outweighs the disadvantages. HPMS sampling procedures are both simple and efficient and, if applied properly, the selected sample will achieve the predetermined levels of desired precision and yield an adequate sample for performance monitoring.

SAMPLE STRATIFICATION AND PRECISION LEVELS

Data needs vary for rural, small urban, and urbanized areas; this variation is reflected in the sample design. The design is capable of producing valid estimates of highway condition, operational, and performance characteristics on a State-by-State basis. Rural and small urban functional systems are sampled on a statewide basis; all Federal-aid urbanized areas must be sampled individually.

The sampling plan consists of the random selection of a panel of road sections within predetermined AADT volume groups (strata) for each functional system in the rural, small urban, and urbanized areas of the State. The stratification of sections (sampling units) into relatively homogeneous AADT volume groups produces estimates of greater accuracy with respect to VMT for a smaller number of samples at the functional system (summation) levels. Although stratification for sample selection is based on the critical data element AADT, tests have shown that AADT stratification is compatible with the sampling of nonvolume related data elements.

The required sample size is a function of the variability of AADT within a volume group, the functional system volume group precision level, and the number of sections available for sampling in the volume group (the universe). The term "precision level" is defined as the degree of confidence that the sampling error of a produced estimate will fall within a desired fixed range. For a precision level of 80-percent confidence with 10-percent allowable error (80-10), there is the probability that 80 times out of 100 the error of a data element estimate will be no greater than $\pm$10 percent of its true value. The precision levels specified in Appendix C represent minimum FHWA requirements for rural, small urban, and urbanized area functional system volume groups. The precision levels determined for the sample design apply specifically to the individual volume strata. The sample size estimating procedures shown in Appendix D are used to determine the required number of samples to meet the target precision levels at the volume group level. The sample adequacy software described in Appendix K produces a table that contains a standard sample size estimation based on Appendix D criteria.

Sample size requirements by functional system will vary by State according to the total number of road sections, the number of predetermined volume groups, the validity of the State's AADT data, and the design precision level. The HPMS sample size requirements are more stringent for the arterial systems, where a higher level of precision is needed because of higher Federal interest. In rural and small urban areas, the sample sizes are based on a 90-5 precision level for the volume groups of the principal arterial system, 90-10 for the minor arterial system, and 80-10 for the collector (excluding minor collector) systems. The sample for individually sampled urbanized areas is broken into two major categories of precision levels:

1. For individual urbanized areas with a population of 200,000 or more and those that are in an NAAQS nonattainment or maintenance area, the design precision levels are 90-10 for the arterial systems and 80-10 for the collector system.
2. For other individual urbanized areas with a population of less than 200,000, the design precision levels for individual volume strata are 80-10.

The statewide summation of individual urbanized functional system data element estimates will result in an overall precision level of at least 80-10.

The higher precision levels at the State level are necessary to obtain comparable urban and rural precision levels and to obtain precision levels that can adequately accommodate desired levels of accuracy for estimates of proportionate values as well as average and aggregate values. That is, although the HPMS sample is designed to measure AADT, the same samples are used to estimate the proportionate values of data such as pavement condition. Since the level of accuracy for estimated proportions is closely related to sample size, precision levels have been set sufficiently high to produce reasonable proportionate estimates at the functional system level.

PREPARATION FOR SAMPLE SELECTION

Before a sample can be drawn, the universe from which it is selected must be defined. This is very important since expansion factors, and the reliability of the expanded sample data, relate directly to the universe definition. Initial steps include:

- First, delimit the boundaries between rural, small urban, and urbanized areas using FHWA-approved, adjusted urban boundaries.
- Next, identify the functional system of all arterial and collector routes within each of these areas.
- Then, break the arterial and collector routes into logical roadway sections.
- Finally, assign all road sections in these functional systems to the predetermined AADT groups shown in Appendix C.

An AADT volume group assignment is required for all roadway sections on functional systems subject to sampling (all but rural minor collector and rural and urban local). Assigning sections to proper volume groups and maintaining proper volume group assignments is an important step. Because of economic growth and development, AADT growth may require periodic adjustments to volume group assignments over time. The HPMS software will assign a volume group for each section where an AADT is provided. Maintaining accurate volume groups requires States and other data providers to maintain comprehensive, high quality, traffic count programs (see Appendix F).

If volume groups other than the predetermined volume groups shown in Appendix C are selected, the AADT limits of these volume groups must be reported to FHWA, and the State will need to assign each section to the appropriate volume group.

Each HPMS section should be relatively homogeneous as to geometrics, traffic volume, cross section, and condition, and should be long enough to constitute a logical section for various analyses such as determining highway investment requirements. In general:

- Rural section lengths should range from 0.5 to 16.1 kilometers (0.3 to 10.0 miles)
- Urban access controlled facility section lengths should usually not exceed 8.0 kilometers (5.0 miles)
- All other urban section lengths should range from 0.2 to 4.8 kilometers (0.1 to 3.0 miles).

These suggested lengths are intended to keep the sample normalized on a national basis. Shorter lengths may be warranted where there are nonhomogeneous roadway elements; longer sections reduce the number of universe sections and result in a somewhat smaller number of initial samples. However, longer

sections may have to be split in later years in order to maintain sample homogeneity; this will increase the number of universe sections and may result in an increase in the required number of samples.

Finally, it is important to precisely document the exact location of each sampled section to assure that yearly and cyclical updates, field reviews, traffic counts, etc., are performed on the appropriate roadway sections.

CALCULATION OF EXPANSION FACTORS

The purpose of the HPMS sample panel is to provide an expandable base for rural, small urban and urbanized area data in each State, stratified by functional system and traffic volume group. An expansion factor is calculated for each volume stratum within each functional system. This is accomplished by dividing the total kilometers (miles) in the stratum by the kilometers (miles) included in that stratum's sample. As noted above, the total universe length in each stratum is a known value based on the AADT volume group identifier. Expansion factors are calculated by the HPMS software for each sample section. The expansion factor allows sample data to be expanded to represent entire functional systems for rural, small urban, and urbanized areas.

SAMPLING RURAL AND SMALL URBAN AREAS

Both rural and small urban area data are sampled on a statewide basis, stratified only by functional system and volume group. The volume group for each road section in the sampling universe must be identified using the tables in Appendix C before sample selection can begin. The number of sections to be included in the sample is determined using the calculation procedure in Appendix D. A minimum of three unique sample sections is required for each volume group; if less than three universe sections exist in a volume group, they must all be sampled and the expansion factor will be 1.000. Sections should be selected from the universe of each functional system and volume group using a random number table or random number generation computer software, until the required number of samples is reached.

SAMPLING INDIVIDUAL URBANIZED AREAS

Urbanized area data are sampled on an individual area basis stratified by functional system and volume group. Each State must individually sample urbanized areas regardless of population size. The volume group for each road section in the sampling universe must be identified using the tables in Appendix C before sample selection can begin.

The number of sections to be included in the sample is determined using the calculation procedure in Appendix D. As with rural and small urban areas, sections should be selected from the universe of each functional system and volume group using a random number table or random number generation computer software, until the required number of samples is reached; a minimum of three sections per stratum is required.

Each State must sample its part of a multi-State urbanized area individually. The sample in each State should not be less than its pro rata share for the entire urbanized area by functional system and volume group nor in any case less than one section per applicable volume group. If length does not exist in a particular volume group in one (or more) portion(s) of a multi-State urbanized area, all of the sampling

should take place in the State(s) where the length does exist. In such areas, expansion factors must be calculated separately for each State's portion. To ensure a consistent sampling approach, States having multi-State urbanized areas are urged to coordinate with the appropriate neighboring State(s) so that all portions of an urbanized area are sampled in the same manner.

An individual sampling approach must be applied to all parts of multi-State urbanized areas if expanded estimates are desired for the complete urbanized area.

SAMPLE MAINTENANCE

An HPMS sample adequacy review should be performed annually as part of a State's sample maintenance activities. The review should be completed shortly after the annual submittal of the HPMS data set; this permits the data provider to assess the adequacy of the sample in time to make changes for the next reporting cycle. Timing of the State's HPMS data processes is an important issue since sampling is dependent to a certain degree upon up-to-date traffic and functional classification data. A number of elements should be considered when making a review of HPMS sample adequacy. These should include not only the assessment of number of samples by volume group, but also checks for potential sample biases.

When conducting a sampling review, the State also should check for biases that may have been inadvertently introduced into the sample. Although the HPMS sample was designed as a fixed panel sample, additions, deletions, and other changes may have been made to the sample to account for system and other changes that occurred over time. And, although changes to the HPMS sample are to be made on a random basis, this may or may not have been a closely observed practice.

As a result, sample bias may have been introduced in areas such as samples on State versus non-State owned roads, subarea biases by highway district or county, or nonrandom selection of adjacent roadway sections as new samples. Some of these biases may be disclosed by comparing the number of miles sampled; for instance, the percent of State owned miles sampled compared with the percent of non-State owned miles sampled. Others may require a more detailed examination of the sample and its distribution; are samples clustered in groups on the same facility, for instance. A periodic review of the sample provides an opportunity to identify if any of these problems exist; as further changes are made to the HPMS sample, any biased sample selection procedures should be eliminated to improve sample randomness.

Need for Sample Panel Adjustments

There are any number of occurrences that may result in a need to reconsider the suitability of the existing sample panel. Some of the more common reasons for considering sample panel adjustments include:

- The decennial census of population is likely to require changes in HPMS sample panels. The sampling basis may need to change because the numbers of small urban areas and individual urbanized areas may change, and/or the FHWA approved, adjusted Census urban boundaries of existing urban areas may be altered.
- The addition of new areas and the expansion of current urban boundaries are likely to require the functional reclassification of roadways within the new boundaries. This will in turn likely require transferring universe and sample sections from one area's panel to another and randomly drawing additional samples to satisfy urban area requirements. Also, the loss of samples caused by movement from rural to small urban or from rural or small urban to urbanized areas may cause a deficiency in the rural or small urban area panels.

- The EPA's designation or redesignation of NAAQS nonattainment or maintenance areas. This may result in new or revised donut areas depending on the geographic boundaries of the new nonattainment or maintenance areas, requiring adjustment to the donut area sample panel. Please see Appendix G regarding the donut area.
- Changes in the existing functional system length and HPMS sample panels are likely to result from functional reclassification, non-Census related changes in urban boundaries, or new road construction.
- Migration of sections among and between volume groups may also result in a need to change HPMS sample panels. Each volume group contained in a functional system is a separate sampling universe; normally, over the short term (less than 3 years), there should be only minor changes in sample section and universe length assignments to specific volume groups as a result of traffic increases or decreases. Universe volume group information for each roadway section must be kept up to date so that correct volume group reassignments can be made (see Appendix C).

A thorough sample adequacy review, conducted on a 3-year cycle basis, provides an opportunity to update the HPMS sample panels when necessary to meet the changed conditions reflected above.

Making Sample Panel Adjustments

Sample panel adjustments should be made as necessary upon completion of a sample adequacy review; use of a 3-year cycle will minimize the burden of completing this task. The following general procedures should be considered when adjusting sample panels:

- Functionally reclassify roadway sections that have moved from rural areas into new or expanded urban/urbanized areas or out of contracting urban areas into rural areas; use appropriate classification criteria and good engineering judgment to determine the extent of change warranted.
- Stratify the reclassified roadway sections within these same areas into traffic volume groups consistent with the groups established for the latest HPMS sample.
- Transfer rural, urban, or urbanized sample sections that have moved from one area type to another into the appropriate functional systems and volume groups in the new panel.
- Calculate the required number of standard samples required for the revised rural, small urban, and urbanized area panels in accordance with Appendix D procedures and select additional samples where necessary.

In using these general procedures, the user should keep in mind that:

- When small urban or urbanized areas contract in size, changes to small urban or rural sample panels will occur; universe and sample sections affected by such changes should be assigned to the correct functional system and volume group in the new panel.
- Make adjustments to standard and donut area sample panels independently. It is better to update the standard sample panels prior to updating the donut area sample panels, since existing standard samples in the donut areas become donut area samples; in general, the same procedures apply (see Appendix G).
- If a new urbanized nonattainment area is designated by EPA, a new donut area sample must be drawn for that nonattainment area in accordance with the procedures in Appendix G.

Selecting Additional Samples

The selection of additional sample sections for a given volume group is straightforward for most sample panel updates. Basically, the number of existing sample sections is compared to the required number as determined from the Appendix D procedures, and additional sample sections are randomly drawn from the nonsampled universe sections in the same volume group to cover any shortfalls. Again, maintaining accurate volume groups requires States and other data providers to maintain comprehensive, high quality, traffic count programs (see Appendix F). This procedure is to be used for the standard sample panels in rural, small urban and individually sampled urbanized areas, and for the donut areas of nonattainment areas; it also is to be used when newly designated urbanized areas are sampled as individual areas.

Sample Permanence

Once a roadway section has been selected for a sample panel, it must be maintained as a sample regardless of changes in volume group assignment, functional system, or geographic area. Sample sections transferred to other geographic areas become part of the sample base for those areas. Samples may be dropped in cases where a roadway is truly abandoned and not relocated, where sample sections are reclassified to the rural or urban local or rural minor collector functional systems, or where sample sections are dropped from use as a result of a sample reduction plan. When samples are deleted, the State must submit a list of the sample numbers, the reasons for the deletions, and where the deletions will occur.

Deleting Samples

Since the standard sample panel has been in existence for some time, the addition of samples and the movement of universe and sample sections from one volume group to another are likely to have caused over sampling in some volume groups. Significant oversampling is not encouraged because of cost and efficiency impacts; sample reductions should be considered a normal part of sample maintenance. Before proceeding with a sample reduction exercise, the State should prepare a sample reduction plan and provide it to the FHWA for evaluation. A sample reduction plan should take account of the following:

1. All sampling criteria must be met; sample size requirements in Appendix D must be maintained for each standard sample functional system.
2. Sample AADT and universe section volume group data must be up to date and accurate.
3. Individual volume group reductions of less than three sample sections in any volume group should not be considered.
4. Random deletion of the samples within each over sampled volume group is required.
5. The three samples per volume group minimum must be maintained.
6. Trends of sample/universe section migration among volume groups should be examined; volume groups that continually gain samples may warrant keeping a few excess samples.
7. An expansion factor maximum of 100.000 should be observed.
8. A State using the HPMS analytical package or the HPMS data base for other purposes may want to keep an oversampled sample panel intact or consider using higher precision levels.
9. The sample reduction process may require more effort than the apparent resulting benefit of maintaining fewer samples; however, a periodic review and adjustment of the sample is needed to maintain the overall viability of the HPMS sample program.
10. A sample reduction should be considered as part of the annual sample review.

Updating Expansion Factors

When updating sample panels, any change in sample length and/or the length of the sampling frame (the universe) requires updating the expansion factors related to affected volume groups. Expansion factors are recalculated before every HPMS submittal to ensure that all changes to volume groups (both universe and sample), whatever the cause, have been properly accommodated. Expansion factor recalculation is one of the final data preparation steps when the HPMS submittal software is used.

A Tabular Summary

The following table provides an overview of conditions which generally require making changes to HPMS sample panels, both standard and donut area. It is divided into those changes triggered either directly or indirectly by Bureau of Census actions, or by changes unrelated to Census actions. The table outlines the "causes" for potential sample panel change and the "Recommendations" to deal with the change.

CAUSE	RECOMMENDATION
CENSUS RELATED	
New Small Urban Areas (Rural to Small Urban)	Adjust all rural sample section records within the new area to urban requirements. Verify statewide rural and small urban area sample and universe bases and select additional samples as necessary.
New Urbanized Areas (Small Urban and/or Rural to Urbanized)	Adjust all rural and small urban area sample section records within the new area to urbanized area requirements. Procedures for drawing new standard samples for individual panels are discussed above. Procedures for drawing donut area samples are discussed in Appendix G. Verify all sample and universe bases and select additional samples as necessary.
Expansion of the Adjusted Boundaries of Small Urban or Urbanized Areas (Rural to Small Urban and Rural and/or Small Urban to Urbanized)	Adjust all affected rural sample section records to urban requirements. Verify all affected sample and universe bases and select additional samples as necessary.
Functional System Reclassification--Any Area	Reassign reclassified sections (universe and sample) to appropriate areas and volume groups. Sample new sections as necessary to maintain required volume group precision levels.
Losses in Urban Population	No action until Census area designation changes.
Major Revision of Boundaries Based on New Census	Redraw sample panel and include old samples, if possible.
Changes or Additions to Nonattainment/Maintenance Area(s).	Updates to the donut area samples are made based on the procedures in Appendix G.
NON-CENSUS RELATED	
New Length by Functional System	Verify sample and universe base; sample new sections, if necessary.
Functional System Reclassification in Any Area	In addition to the movement of sections because of reclassification, there may be a need for possible volume group changes for universe and/or sample sections, precision level changes, and additional samples.
Volume Group Reassignment of Sections	Reassign sample sections but no further action is needed if changes are minor; if changes are major, verify volume group sample and universe bases for all affected volume groups and add samples, if necessary.
Expansion Factor	Adjust expansion factor values for sample section records in the affected group.

Splitting Samples

Sample sections should be split only when significant HPMS related changes in an existing sample occur. The most critical reasons for splitting samples involve changes in traffic (AADT) on the section, county

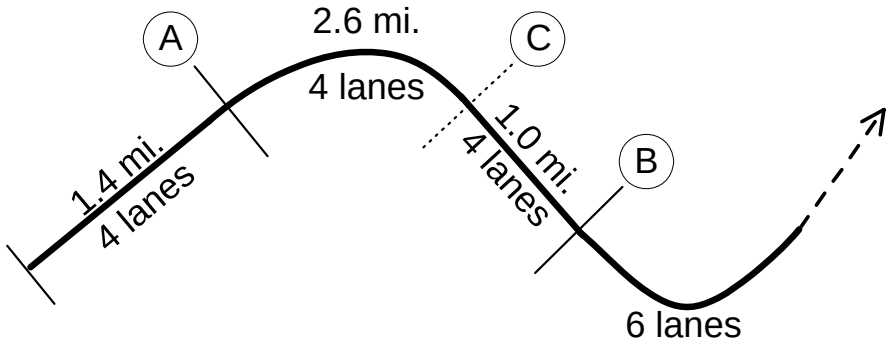
code, functional system, urban/rural/urbanized status, and number of lanes. Many times these changes are the product of adjacent land development, which may result in the improvement of a portion of an existing sample, an increase in traffic on the improved portion, and/or the inclusion of the improved portion in an urbanized area boundary. In general, a change in any of the following key HPMS data items on an existing sample section should result in splitting an HPMS sample:

Item Number	Data Item
4	County Code
13	Rural/Urban Designation
15	Urbanized Area Code
16	NAAQS Nonattainment Area Code
17	Functional System
19	National Highway System
25	Governmental Ownership
26	Special Systems
27	Type of Facility
33	AADT
34	Number of Through Lanes

It may be necessary to split HPMS samples when there are changes in other HPMS data items; however, the State needs to make a reasoned judgment of the particular case beyond these minimum specifications. In most cases, for changes in other HPMS data items on existing sample sections, it is more than adequate to code the predominant or typical condition on the existing sample section and retain the entire sample. If a shoulder type, for instance, is changed on a portion of an existing sample section, it is acceptable to code the resulting predominant type of shoulder on the entire section in lieu of splitting the sample. If part of a sample is improved and the remaining portion is to be improved in the next construction season, it is likewise acceptable to code the predominant condition and retain the entire sample. Samples should not be split for changes in non-HPMS related State inventory items, such as guardrail changes or highway district boundary, etc.

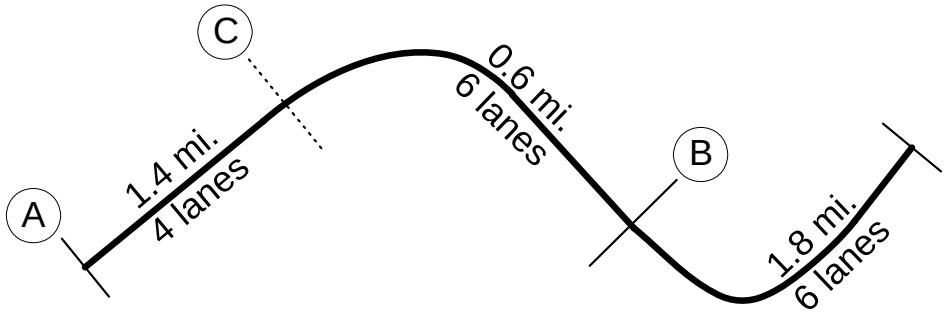
When splitting an existing sample section, the State should select the portion to remain as the sample based either on length - keep the longest portion - or by random pick; the selected portion becomes the sample and retains the existing sample identifier. Although FHWA suggests that the longest section be retained as the sample, either is acceptable. The remaining portion should be converted to a universe section or merged with an adjoining universe section. It is important that all count information is adjusted to reflect actual conditions on the retained sample section. Two examples follow.

Splitting Samples



Existing sample section (A-B) with 4 lanes = 2.6 mi.
Urban/rural change = 1.0 mi.
New universe section (C-B)
Remaining sample now = 1.6 mi. (A-C)

Splitting Samples



Existing sample section (A-B) = 2.0 mi.
Major improvement (add 2 lanes) = 0.6 mi.
Merge (C-B) with adjacent universe section (1.8 mi.) for total = 2.4 mi.
Sample section now = 1.4 mi. (A-C)

ELIMINATING SHORT SAMPLE SECTIONS

As part of sample maintenance activities, existing samples should be reviewed to see if they can be deleted or recombined with adjacent sample sections. In the past, excessive splitting of HPMS sample sections has resulted in the accumulation of many short adjacent sample sections. Adjacent short samples not meeting minimum length requirements should be recombined into longer sample sections if they have similar roadway characteristics and the key data items listed under the "Splitting Samples" discussion are the same. Excessive short samples resulting from previous sample splits can also be considered for deletion if HPMS-critical roadway characteristics are the same for a number of adjacent samples. In this case, the State may wish to retain the longest sample and recombine the remaining samples with an adjacent universe section or merge them into a new universe section. The sample section to be retained can also be selected randomly. Either way, a reduction in an excessive number of short samples may save the State financial and personnel resources and will improve sample representation.