

**Information Collection Request for the 8th Drinking
Water Infrastructure Needs Survey and Assessment (DWINSA)**

PART B OF THE SUPPORTING STATEMENT (FOR STATISTICAL SURVEYS)

Table of Contents

8 th DWINSA.....	1
B.1 SURVEY OBJECTIVES, KEY VARIABLES AND OTHER PRELIMINARIES.....	1
B.1.a Survey Objectives.....	1
B.1.b Key Variables.....	2
B.1.c Statistical Approach.....	3
B.1.d Feasibility.....	5
B.2 SURVEY DESIGN.....	6
B.2.a Target Population and Coverage.....	7
B.2.b Sample Design.....	8
B.2.c Precision Requirements.....	19
B.2.d Data Collection Instrument Design.....	22
B.3 PRE-TESTS AND PILOT TEST.....	23
B.3.a Pre-tests.....	23
B.3.b Pilot Test.....	23
B.4 COLLECTION METHODS AND FOLLOW-UP.....	24
B.4.a Collection Method.....	24
B.4.b Survey Response and Follow-up.....	24
B.5 ANALYZING AND REPORTING SURVEY RESULTS.....	25
B.5.a Data Preparation.....	25
B.5.b Analysis.....	25
B.5.c Reporting Results.....	26

INTRODUCTION TO PART B

The Environmental Protection Agency (EPA) will conduct the following type of statistical survey for the 8th Drinking Water Infrastructure Needs Survey and Assessment (DWINSA). The EPA will survey large drinking water Community Water Systems (CWSs) (serving over 100,000 persons) via a census approach, medium CWSs (serving between 3,301 and 100,000 persons) through a statistically based sample, and small CWSs (serving 3,300 or fewer persons) through a re-sample of those that were surveyed in the 7th DWINSA. Consistent with the principles of cooperative federalism, the information collection and quality review for medium and large CWS will be facilitated by the CWSs' respective state SRF programs. For non-profit noncommunity water systems (NPNCWSs) and American Indian and Alaska Native Village water systems, the EPA will not collect new data on their capital investment projects. Instead, the EPA will adjust the estimated capital improvement needs from the 7th DWINSA by adjusting to 2026 dollars and updating for changes in the inventory of systems. A sample of American Indian and Alaska Native Village utilities will receive a set of supplemental questions for the 8th DWINSA related to operations and maintenance (O&M) and the attributes of technical, managerial, and financial (TMF) capability. The EPA will also engage with key stakeholders to better understand and incorporate the differential in infrastructure costs faced by small community water systems into the 8th DWINSA needs estimate (see Supporting Statement Part A, Section 8a and 8b).

8th DWINSA

B.1 SURVEY OBJECTIVES, KEY VARIABLES AND OTHER PRELIMINARIES

B.1.a Survey Objectives

The primary objective of the 8th DWINSA is to collect information from public drinking water systems on the capital improvements they need to make to continue to provide safe drinking water to consumers. These data are used to produce a national estimate as well as state-specific estimates of water systems' 20-year infrastructure investment need. In addition, the EPA will collect information through supplemental questions in the data collection instrument. As previously mentioned in Part A, to address the mandate for the survey to collect lead service line replacement information by sections 1452(h) and 1452(i)(4) of the Safe Drinking Water Act (SDWA), as amended by America's Water Infrastructure Act (AWIA) of 2018, and to avoid duplication of efforts and to obtain the best quality information, the EPA will use service line inventory data submitted by all DWSRF-eligible systems for to meet Lead and Copper Rule requirements once these data are available. The EPA will ask state small, medium and large CWSs supplemental questions to identify potential funding barriers for accessing DWSRF funding. The EPA will ask American Indian (AI) and Alaska Native Village (ANV) utilities supplemental questions about operations and maintenance (O&M) and their technical, managerial, and financial (TMF) capability. The EPA has established approaches to ensure that the overarching goals of the survey are met:

- Estimate the total national 20-year need.
- Update small system needs and apply the differential in infrastructure costs for small systems.
- Estimate the total 20-year need for medium and large community water systems (CWSs) for each full participation state.
- Collect lead service line replacement project cost data.
- Collect data on DWSRF funding barriers.
- Collect information about the O&M and TMF capability of AI and ANV utilities.
- Provide complete and accurate data to Congress.
- Inform the distribution of DWSRF capitalization funds to states.
- Maintain the credibility of the DWINSA findings.

The EPA proposes to collect information on the cost of water systems' infrastructure needs. If cost data are not available from systems, the EPA proposes to collect information that will enable the Agency to model costs. In the data collection instrument, the respondent will identify needs on a project-by-project basis and list the "type(s) of need" that the project will meet. The types of need include raw water source, transmission, source water treatment, storage, distribution, pumping stations, and other needs.

The EPA will use the information from the DWINSA to estimate capital investment needs of drinking water systems. The information will be used to allot DWSRF monies among states.

For the 8th DWINSA, the EPA is proposing to continue to use the modified panel approach to select medium and large system survey respondents used for the 6th and 7th State DWINSAs. The EPA is proposing to use the 7th DWINSA sample for small systems. The sampling design is discussed in detail below.

As previously discussed in Part A of this ICR, the 8th DWINSA includes new supplemental questions for states and territories. EPA did not establish separate precision targets for the responses to these questions, so the supplemental questions will be submitted to all state CWSs sampled for the 8th DWINSA, including those in partial participation states. Medium CWSs in partial participation states will not be sent a survey asking for information about capital improvement projects, but a sample of medium systems will be asked to respond to the supplemental questions. As previously discussed in Part A, this ICR assumes that large, medium and small systems in all states will be asked the supplemental questions. AI and ANV utilities will be sent supplemental questions but will not be asked about their infrastructure investment needs.

B.1.b Key Variables

Several key variables are available from the Safe Drinking Water Information System (SDWIS). To ensure accuracy, the 8th DWINSA will verify these data by asking respondents to confirm existing information (pre-populated on the data collection instrument) or correct it. These variables include population served, number of service connections, primary source of supply, ownership type (private or public), and whether the system purchases water from another water system. Respondents will also be asked to provide information on total length of pipe in the system and the total treatment design capacity. Total length of pipe and total design capacity are not available in SDWIS. The pre-populated questionnaires will include this information from the 7th DWINSA for systems that were in the previous survey.

Information on capital needs will be collected from respondents on a project-by-project basis. For each project, respondents will be asked to provide the following types of information: type of need; reason for need; documentation of need; if the project is a new project or is to replace, rehabilitate or expand existing infrastructure; if the project is needed now to protect public health or if it is needed over the next 20 years to continue to provide safe drinking water; the federal regulation or state requirement (if the project is to meet a current regulation or state requirement); design capacity of source, storage, and treatment projects; length and diameter of pipe projects; diameter and number needed for projects such as water meters; cost of the project and date of the cost estimate (if documentation of the cost is available). For most of these variables, respondents will choose the appropriate “documentation,” “type of need,” “reason for need,” or “regulation or requirement” from the EPA’s Lists of Codes (See Appendix B).

The principal variable of interest is total projected capital needed for each water system in the 8th DWINSA for the time period of January 1, 2026, through December 31, 2045. The total capital need for all systems in each state (to be derived from the statistical sample of systems) is the key variable that decision-makers at the EPA use to allot funds to states based on need.

The method of data collection has been designed to minimize burden on respondents while ensuring that information is collected in a consistent manner. Collecting information on a project-by-project basis, for example, will be particularly helpful in reducing burden since most respondents develop Capital Improvement Plans (CIPs) on a project-by-project basis or report infrastructure-specific projects based on their asset inventory.

Information on type of need will be used to disaggregate total capital needs for the EPA’s Report to Congress. Information on the reason for need will be used to verify the public health benefit of the need. Information on the date of the cost estimate will be used to provide a consistent basis for cost

estimates across systems. Information on a regulation or requirement will be used to determine the reported project costs related to federal regulations or state requirements.

If a system cannot provide cost estimates, additional data are necessary so that the Agency can impute costs. Each of these variables is described in greater detail later in this document.

Key variables among the supplemental questions for AI and ANV utilities are annual revenue and expenses. Questions about their workforce, managerial capacity, and financial management support the analysis of the O&M and TMF issues facing AI and ANV utilities.

B.1.c Statistical Approach

The proposed sampling plan for the 8th DWINSA is designed to achieve a desired level of precision for state-level estimates of total capital needs for medium and large CWSs. The EPA proposes to use a census of large CWSs and a statistical sample of medium-sized CWSs to estimate total capital needs. The sample also will be used to collect information on the supplemental questions. EPA proposes using the small systems sample that was selected for the 7th DWINSA as described below. Average national small system needs are applied to each small system in each state's inventory to estimate their small system need. The proposed approach also includes a statistical sample of AI and ANV utilities designed to achieve a desired level of precision within categories of these utilities. This statistical approach minimizes burden while achieving the desired level of precision (see Section B.2.c.i., *Precision Targets*).

Medium and Large CWSs

For purposes of the sampling plan, the 8th DWINSA design divides CWSs serving populations of more than 3,300 persons into two groups: CWSs serving populations of more than 100,000 (large systems) and systems serving populations of 3,301 to 100,000 (medium systems). The EPA proposes to sample with certainty systems serving more than 100,000 persons. These systems have the largest capital needs and they have the staff to respond efficiently to the 8th DWINSA. The EPA proposes a stratified random sample within each state of systems serving 3,301 to 100,000 persons. This methodology reduces burden and while meeting the DWINSA data quality objectives.

To further reduce burden, the EPA proposes using a modified panel approach for medium systems in the 8th DWINSA. Rather than select a completely new sample of systems for the 8th DWINSA, the EPA will reassess the needs of most of the systems that participated in the 7th DWINSA. The EPA determines the number of systems needed in each state to meet the survey's precision target. It starts with the 7th DWINSA sample, adding or dropping medium systems as needed to meet the precision target. EPA will then randomly select 25 percent of the sample to drop within each state and will then randomly select replacement systems from the sampling frame. By starting with information from the 7th DWINSA, this approach reduces the amount of time needed for systems to prepare and states to review the responses from systems resurveyed for the 8th DWINSA. This approach will maintain the EPA's sampling sizes need for each stratum and ensure that the EPA meets its precision targets for each state. Additionally, by using information from the 7th DWINSA, the modified panel approach would reduce the amount of time necessary for the EPA to review each system's response. By replacing 25 percent of the medium system sample, the EPA will reduce a potential source of bias introduced by the panel. (When a completely new sample was selected for each assessment, the sampling error was a random component that changed from survey to survey. With the panel approach, this error becomes systematic.) By refreshing 25 percent of the sample, the approach alleviates this potential source of bias and helps to ensure that the

8th DWINSA sample represents the need as it exists in 2026. This approach was used for the 6th and 7th DWINSAs.

To meet the state-level precision targets, the EPA will use the same strata as in the 7th DWINSA. As previously mentioned, the EPA will adjust the sample size to accommodate changes in the sample frame. These changes may address new systems, systems that are no longer active and systems that have “migrated” between strata (become smaller or larger, or changed source). As in the 7th DWINSA, the EPA will first determine the total sample size for each state to meet the target level of precision. The EPA will then allocate the sample to strata in order to maximize the efficiency of the design.

Medium CWSs in Partial Participation States

The 8th DWINSA survey of 20-year capital investment needs will not be administered to medium CWSs in partial participation states, but a sample of these systems will receive the supplemental questions. Unlike medium systems in states that fully participate in the survey, the EPA will develop national-level estimates for these systems. The EPA will sample 311 medium systems in partial participation states to receive the supplemental questions.

Small CWSs

The object of the 8th DWINSA is to update state-level estimates of total capital needs for CWSs, including systems serving 3,300 or fewer persons (small systems). The 7th DWINSA collected data from a sample of approximately 600 small CWSs through site visits. The 7th DWINSA used a national sample of small CWSs rather than the larger state-level samples because of the high cost of the site visits and the large number of small systems that would need to be sampled for state-level estimates. In the 7th DWINSA, EPA used the national sample to estimate total capital needs of small CWSs and used that information to infer the total capital needs for small CWSs in each state. The 8th DWINSA will build on this approach by conducting phone interviews with the small CWSs surveyed as part of the 7th DWINSA to update their capital improvement needs reported in the 7th DWINSA and to capture new needs.

The small system sample is a stratified random sample, stratified by primary water source and the size of the population served. EPA will retain the full sample of the small systems that participated in the 7th DWINSA rather than replace 25 percent of the 7th DWINSA small system sample as it does with medium systems. EPA will drop small CWSs from the sample that migrated to medium or large strata between the 7th DWINSA and the date of frame validation for the 8th DWINSA or ceased operations since the 7th DWINSA. By using information from the 7th DWINSA and collecting capital needs information through phone calls with small CWS personnel, EPA will reduce the burden on survey respondents.

American Indian and Alaska Native Village Utilities

The 8th DWINSA will not collect new data about the total capital needs for AI and ANV systems. The survey will ask questions about the O&M and TMF concerns of water utilities, including drinking water utilities and combined drinking water and wastewater utilities. These utilities will receive supplemental questions that have been tailored to AI and ANV utilities.

B.1.d Feasibility

To fully assess feasibility, the EPA convened a workgroup (see Part A Section 5.b) to comment on the proposed data collection approach and its feasibility. The data collection instrument to be used for the

8th DWINSA, which was designed with the capabilities of the typical respondent in mind, is generally the same form as used for the past four DWINSAs. For the 4th DWINSA, the EPA conducted a pre-test in which the EPA's contractor met with individual CWS operators and discussed the proposed survey. System operators were asked to comment on all proposed data elements and the feasibility of collecting information by mailed survey.

The Agency recognizes that most systems may not have cost data or documentation of costs for some projects needed within the 20-year period captured by the DWINSA. In those cases and consistent with the previous seven DWINSAs, the 8th DWINSA data collection instrument requests other readily available information that the EPA can use to model costs. The EPA will emphasize to respondents that they are not to develop cost estimates for the purposes of the 8th DWINSA. In addition, the EPA (or states) will provide systems with technical assistance for completing the data collection instrument.

The EPA has developed cost models for most of the infrastructure needs included in the 8th DWINSA based on the size and capacity of a project. These cost models were originally developed during the 1st DWINSA; have been updated during subsequent assessments, including the 7th DWINSA; and will be used again for the 8th DWINSA. New cost models may be developed as part of the 8th DWINSA when deemed appropriate.

The time frame for the 8th DWINSA is acceptable to the users of data within the Office of Ground Water and Drinking Water (OGWDW) and sufficient to complete a report to Congress by its anticipated due date in 2027. The schedule also is acceptable to other users of the data.

B.2 SURVEY DESIGN

This section contains a detailed description of the statistical survey design including a description of the sampling frame, sample identification, precision requirements, modified panel approach, and data collection instrument.

Medium CWSs in Fully Participating States and Large CWSs in All States

The EPA's precision target for the 8th DWINSA is to be 95 percent confident that an interval of plus or minus 15 percent of the estimated need includes the true need for each state. This precision target has been modified slightly from the target for the 7th DWINSA, which was 95 percent confidence plus or minus 10 percent of the estimated need. The sample sizes will be adjusted to account for changes in the inventory of systems, if necessary, to ensure the 8th DWINSA sample meets the precision targets.

The sample plan for the 8th DWINSA is stratified random sampling within each state. Each state sample is used to estimate the total need of medium and large systems in each state.

Stratification increases the precision of estimates compared with a simple random sample of the target population of systems. In stratified samples, the target population is divided into non-overlapping groups, known as strata, from which separate samples are drawn. The goal of stratified sampling is to choose sample sizes within each stratum in a manner designed to obtain maximum precision in the overall estimate for the population. Stratification variables for this study include population size (populations of: 3,301 to 10,000; 10,001 to 25,000; 25,001 to 50,000; 50,001 to 100,000 and more than 100,000) and primary sources of supply (surface and ground). Systems serving more than 100,000 persons are selected with certainty. For the 8th DWINSA, the survey will rely on a modified panel

approach in which 75 percent of the 7th DWINSA respondents serving populations of 3,301 to 100,000 were resampled, and 25 percent were put back into the frame and a new 25 percent were drawn.

Medium CWSs in Partial Participation States

In cases where the state is not participating in the data collection of capital improvement needs for systems serving 3,301 to 100,000 persons, the EPA will not provide state-specific results for capital improvements, as the data collection for these states does not meet the DWINSA data quality objectives. The EPA will include an overall national result for these systems using the average need by stratum of the systems in states that are participating in the full 8th DWINSA. For systems serving 3,301-100,000 persons in states that are partially participating and will therefore receive supplemental questions, the 8th DWINSA will use a modified panel approach for sampling these systems within each state. This approach is described in more detail in Section B.2.b.

For systems in partial participation states, a stratified random sample of CWSs will be selected from each partial participation state to determine which systems will receive the supplemental questions. The sample design for these systems will be identical to that used for the medium CWSs in fully participating states.

Small CWSs

For the 8th DWINSA, EPA will use the same sample of small systems that participated in the 7th DWINSA, excluding those systems that have migrated to a medium or large system strata or ceased operations since the 7th DWINSA. The sample design for the 7th DWINSA was national-level stratified random sampling. Stratification variables for the 7th DWINSA included population size (populations of fewer than 101, 101 – 500, 501 – 1,000, and 1,001 to 3,300) and primary sources of supply (surface and ground). For the 7th DWINSA, EPA estimated that the sample size for small systems would allow EPA to estimate the national capital need of these systems with a 95 percent confidence interval equal to plus or minus 15 percent. This design for small CWSs was driven by budgetary constraints and the substantial small system sample that would be needed to increase the confidence interval. EPA's objective in sampling was to achieve the maximum level of precision on a national basis without exceeding that budgetary constraint. The national estimate is divided amongst the states based on each state's small system inventory.

EPA's experience in previous DWINSAs demonstrated that small systems could not independently respond to the survey with accurate information. For the 7th DWINSA EPA sent site visitors to each small system in the sample to collect information about system need. To reduce the cost of the site visits, the small system sample was selected in clusters so site visitors could visit several systems in each trip. Therefore, the sample was selected in two stages, as described in section B.2.b. For the 8th DWINSA, EPA will use the 7th DWINSA small system sample and will not include a new set of site visits. Through phone conversations with system representatives, EPA will update the capital improvement needs for the approximately 600 small systems for which capital improvement needs were documented through site visits as part of the 7th DWINSA. The EPA is also collecting information from stakeholders that could inform a cost differential adjustment for small systems.

American Indian and Alaska Native Village Utilities

The 8th DWINSA will select a stratified random sample of AI and ANV utilities. The 8th DWINSA will use the list of eligible utilities used by the Indian Health Service (IHS) in its 2020 operations and maintenance

cost evaluation of Tribal utilities updated with 2025 IHS data and information provided by the EPA Regions as its sampling frame. The list comprises 282 AI utilities in the contiguous 48 states and Alaska and 142 ANV utilities.

The AI sample will be stratified by the type of utility (water utilities and combined water and sewer utilities), water source (systems that use surface water and systems that use ground and purchased sources) and the size of the population served (populations of less than 501; 501 to 1,000; 1,001 to 3,300; 3,301 to 10,000; and greater than 10,000). ANV utilities, which are mostly small (serving fewer than 3,301 persons) are stratified by the type of drinking water distribution and wastewater collection systems used by the utilities, distinguishing among pressure, circulating, individual, and community haul drinking water distribution systems and gravity, vacuum, low-pressure, honey bucket, and community haul wastewater systems. The strata are shown in section B.2.b.iii.

The EPA's precision target is to be 90 percent confident that the interval of plus or minus 30 percent of total expenses for AI utilities and ANV utilities includes true total expenses.

B.2.a Target Population and Coverage

The target population for the 8th DWINSA state survey is the number of CWSs in the nation that are eligible for DWSRF funding. A CWS is a public water system (PWS) that serves at least 15 service connections used by year-round residents or regularly serves at least 25 year-round residents (40 CFR 141.2). Federally owned systems are excluded. The 8th DWINSA is designed to produce estimates of the capital need of CWSs for each full participation state. In partial participation states, the EPA will be able to provide state-specific results for systems serving 100,000 or more persons. The EPA will include an overall national result for the systems serving 3,301 to 100,000 persons in partial participation states using the average need by stratum of the systems in full participation states and the total number of systems by stratum in the partial participation states. The 8th DWINSA will produce a national-level estimate of capital needs of small CWSs. EPA will calculate the average need per small system. EPA will multiply this average by the number of small systems in each state to estimate each state's small systems' need. EPA will use regional construction cost data to adjust the need for differences in prices among the states.

For Tribal utilities, the target population for the 8th DWINSA is AI and ANV utilities serving primarily Tribal residential communities. The utilities include PWSs that are eligible for DWSRF funding. The 8th DWINSA is designed to produce estimates of the total revenue and expenses of AI and ANV utilities, as well as to estimate the proportion that meet several criteria related to O&M and TMF.

B.2.b Sample Design

This section describes the sample design. It includes a description of the sampling frame, target sample size, stratification variables and sampling method. The sampling design employed is a stratified random sample of CWSs and AI and ANV utilities. The strata employed in the design are discussed in Section B.2.b.iii.

B.2.b.i Sampling Frame

The sampling frame for the state systems is developed from SDWIS. SDWIS is a centralized database for information on PWSs, including their compliance with monitoring requirements, maximum contaminant

levels (MCLs) and other requirements of the SDWA. The following information was extracted from SDWIS for the statistical survey and verified by participating states:

- Name of system.
- Type of system (CWS).
- Retail population served.
- Consecutive population served.
- Total population served.
- Primary source (surface water or ground water).
- PWS identification number (PWSID).
- Ownership type.

From these data, the EPA will develop the frame from which the EPA will calculate summary statistics (e.g., number of systems per state in pre-defined strata) for use in calculating sample size. SDWIS includes approximately 150,000 PWSs, of which approximately 50,000 are CWSs.

Justification for the Use of SDWIS

The following criteria are used to assess a proposed sampling frame:

- It fully covers the target population.
- It contains no duplication.
- It contains no foreign elements (i.e., elements that are not members of the population).
- It contains information for identifying and contacting the units selected in the sample.
- It contains other information that will improve the efficiency of the sample design.

The units of observation for this survey of state systems are CWSs, a subset of PWSs. SDWIS is the ideal choice for a sample frame because of its inclusive coverage of all units of observation for the 8th DWINSA. In addition, SDWIS has two other advantages: it contains information that will facilitate contacting the respondents and it contains other information that is useful in stratifying the sample, thereby improving the efficiency of the sample design.

In previous surveys where SDWIS was used as a sample frame, there have been criticisms of its utility. Since 1989, the EPA has conducted audits of the quality of SDWIS data. As a result, the EPA is aware of the limitations of SDWIS. The audits, however, show that errors in classification of systems by strata proposed for the 8th DWINSA are rare. The audits show that systems are misclassified by population or source in less than one percent of all cases.

To mitigate any potential problems with the sample frame, the 8th DWINSA design anticipates substantial state involvement in the DWINSA process. For example, states will check the sample frame of systems that will be used to determine the final sample. In the EPA's experience, states often have in-house data systems with very accurate data. Even if these data are not transmitted to SDWIS, they are available and can be used by states to check the sample frame. Examples of state information that ensures an accurate frame is developed include the population served by consecutive systems that purchase sufficient water to affect the size of the infrastructure of the wholesale system, and knowledge of recent or pending source water changes to ground water or to surface water.

The sampling frames for the AI and ANV utilities are based on the list of eligible utilities developed for the IHS operations and maintenance evaluation completed in 2020. The following information will be extracted and verified by the EPA Regions:

- Name of utility.
- Type of utility (water only or combined water and wastewater).
- Population served.
- Primary source (surface water or ground water).
- Utility identification number.
- List of PWSs included in the utility.
- IHS area.
- Drinking water distribution system type.
- Wastewater collection system type.

From these data, the EPA will develop the frames from which the EPA will calculate summary statistics (e.g., number of utilities per stratum) for use in calculating sample size requirements and selecting the samples.

Justification for the Use of the IHS data

As with SDWIS, the IHS data meets the following criteria:

- It fully covers the target population.
- It contains no duplication.
- It contains no foreign elements (i.e., elements that are not members of the population).
- It contains information for identifying and contacting the units selected in the sample.
- It contains other information that will improve the efficiency of the sample design.

The units of observation for this survey are utilities, which consist of PWSs and, for combined utilities, wastewater systems. The IHS list is the ideal choice for the sample frames because of its inclusive coverage of all units of observation for the 8th DWINSA. It also contains information that will facilitate contacting the respondents and it contains other information that is useful in stratifying the sample, thereby improving the efficiency of the sample design.

B.2.b.ii Sample Size

Medium and Large CWSs

Exhibit B-2-1 at the end of this subsection shows the preliminary sample sizes for the 8th DWINSA. For the modified panel approach, sample sizes for the 8th State DWINSA will be similar to the 7th DWINSA, with changes to accommodate: (1) the addition of large systems serving more than 100,000 persons since the 7th survey; (2) partial participation states from the 7th survey that fully participate in the 8th survey; (3) the change to the precision target of 95 percent confidence plus or minus fifteen percent, and (4) changes in the inventory of medium systems, including systems changing size or source categories or the creation of new systems. As shown on Exhibit B-2-1, the sampling design will be implemented to achieve state-level precision targets for CWSs serving more than 3,300 persons. The total sample includes CWSs serving 3,301 to 100,000 persons in 15 partially participating states that will receive supplemental questions but will not receive questions about their capital needs. Precision targets are discussed in Section B.2.c. Exhibit B-2-2 shows the sample size needed for the AI and ANV surveys.

For state systems, the task of determining the sample size for each stratum requires two steps. The first step determines the sample size for each state that achieves the precision target for that state. The second step allocates the sample among the relevant strata in the state. For AI and ANV utilities, the precision target is set for each stratum, so the second step is not necessary. The strata are described in section B.2.b.iii.

The first step calculates the total number of systems required at the state level to meet the precision requirements. The sample size is given by:

$$1. \quad n_{0g} = \frac{\left(\sum_{h=1}^H N_{gh} S_{gh} \right)^2}{V_g}$$

Where: n_{0g} = the sample size for state g (prior to the finite population correction)

N_{gh} = the total number of systems in the g^{th} state in the h^{th} stratum (taken from SDWIS)

s_{gh} = the standard deviation of the variable of interest in the g^{th} state for the h^{th} stratum (estimated using data from the data from previous assessments)

H = the number of strata defined in the sample design for the g^{th} state

V_g = the desired sampling variance for the total system (those serving more than 3,300 persons) capital needs estimate for state g .

The desired error in the sample is expressed as a relative error. In the above equation, $V_g = (d/Z_{\alpha} * \hat{Y}_g)^2$. $\hat{Y}_g$ is an estimate of the total capital needs for a given state. $\hat{Y}_g$ is computed for each state by calculating the mean total capital needs for stratum h (from the prior DWINSAs) and multiplying this mean by the actual number of systems in each stratum for that state (N_{gh}). Summing across strata

provides an estimate of $\hat{Y}_g$. d is the half-width of the desired confidence interval (0.15 for DWINSAs). Z_α is the value of a standard normal distribution for a confidence level of $1 - \alpha$, (1.96 for DWINSAs).

Because the number of water systems is known and finite, the following population correction is applied:

$$2. \quad n_g = \frac{n_{0g}}{1 + \frac{1}{V_g} \sum_{h=1}^H N_{gh} s_{gh}^2}$$

The second step allocates the total sample to each sampling stratum. The EPA will randomly draw this number of samples from each of these strata. The Neyman allocation is used to determine the sample size for each stratum:¹

$$3. \quad n_{gh} = n_g \left(\frac{N_{gh} s_{gh}}{\sum_{h=1}^H N_{gh} s_{gh}} \right)$$

(Because systems serving populations of more than 100,000 are to be sampled with certainty, H is the number of strata of systems serving 100,000 or fewer persons.)

In order to implement these sample size and sample allocation equations, the EPA needs estimates for V_g , N_{gh} , s_{gh} and mean total capital needs by stratum. Information on mean total capital needs by stratum and s_{gh} were estimated using data from the prior DWINSAs.

¹ J. Neyman, "On the Two Different Aspects of the Representative Method: The Method of Stratified Sampling and the Method of Purposive Selection," *Journal of the Royal Statistical Society*, Vol. 97 (1934), pp. 558-606; as cited in William G. Cochran, *Sampling Techniques* (New York: John Wiley & Sons), 1977.

Medium CWSs in Partial Participation States

The total sample for medium CWSs in partial participation states is 311. These systems are included in the counts in Exhibit B-2-1 represented in the column titled “Number of Systems Receiving Supplemental Questions Only.”

Small CWSs in States

For the 8th DWINSA, EPA is using the sample of small systems that participated in the 7th DWINSA. In the 7th DWINSA, the total small system sample was set at approximately 600 by available resources and the data quality objective. EPA allocated the sample among eight strata to produce the most efficient estimate of small sample need, given this sample size. The sample for systems serving 3,300 or fewer persons was allocated among source water and population-served strata using a Neyman allocation. The sample is selected in two stages. In the first stage, a probability proportionate to size sample of counties is selected, in which the number of small systems in the county is the measure of size; if a county contains few small systems, it is combined with an adjacent county. In the second stage, a sample of six CWSs is selected from each county (or group of counties). The total sample for small CWSs is 600. These systems are included in the counts in Error: Reference source not found represented in the column titled “Total Number of Small Systems in Sample Receiving Both Capital Needs and Supplemental Questions.”

American Indian and Alaska Native Village Utilities

The approach for AI and ANV utilities is similar. Equation 4 is used to determine the sample size for each stratum and computed using average annual O&M expenses:

$$4. \quad n_{0gh} = \frac{CV_{gh}^2}{V_{gh}}$$

Where: n_{0gh} = the sample size for utility in location g (AI or ANV) in stratum h, prior to the finite population correction

cv_{gh} = the coefficient of variation (standard deviation divided by the mean) of the annual expenses of utilities in g in the hth stratum (estimated using data from the data from IHS O&M study)

V_{gh} = the desired sampling variance for utilities in g in stratum h. It is equal to the half width of the desired confidence interval (.30) divided by the value of a standard normal distribution for a confidence level of 90 percent (1.64), or approximately 0.18.

The finite population correction is applied to each stratum.

5.

$$n_{gh} = \frac{n_{0gh}}{1 + \frac{n_{0gh}}{N_{gh}}}$$

The Neyman allocation used for the state sample is not needed because the sample size is determined by stratum to achieve the stratum-specific precision target. The total sample for AI utilities and ANV utilities is the sum of each stratum's sample. The number of eligible AI and ANV utilities in the IHS inventory and the size of the samples are shown in Exhibit B-2-2.

Exhibit B-2- 1: State Sample Sizes

State	Number of CWSs Serving More Than 3,300 Persons				Number of CWSs Serving 3,300 or Fewer Persons
	Inventory of Systems	Estimated Sample Size			
		Number of Systems Receiving Both Capital Needs and Supplemental Questions*	Number of Systems Receiving Supplemental Questions Only**	Total Number of Systems in Sample	Total Number of Small Systems in Sample Receiving Both Capital Needs and Supplemental Questions***
Alaska *	19	2	11	13	6
Alabama	335	57	0	57	6
Arkansas	196	47	0	47	6
American Samoa	1	1	0	1	2
Arizona	142	37	0	37	6
California	697	166	0	166	36
Colorado	187	48	0	48	12
Connecticut *	57	7	15	22	12
District of Columbia	1	1	0	1	0
Delaware *	33	3	13	16	6
Florida	384	76	0	76	24
Georgia	256	46	0	46	12
Guam	3	3	0	3	0
Hawaii *	34	2	18	20	6
Iowa	155	53	0	53	12

State	Number of CWSs Serving More Than 3,300 Persons				Number of CWSs Serving 3,300 or Fewer Persons
	Inventory of Systems	Estimated Sample Size			
		Number of Systems Receiving Both Capital Needs and Supplemental Questions*	Number of Systems Receiving Supplemental Questions Only**	Total Number of Systems in Sample	Total Number of Small Systems in Sample Receiving Both Capital Needs and Supplemental Questions***
Idaho *	57	3	25	28	18
Illinois	500	58	0	58	12
Indiana	217	42	0	42	6
Kansas	115	33	0	33	6
Kentuckv	253	93	0	93	6
Louisiana	264	49	0	49	12
Massachusetts	258	42	0	42	6
Marvland	67	24	0	24	12
Maine *	40	1	19	20	6
Michigan	311	47	0	47	18
Minnesota	188	35	0	35	12
Missouri	239	43	0	43	18
Northern Mariana Is.	2	2	0	2	2
Mississippi	206	44	0	44	6
Montana *	36	1	18	19	12
North Carolina	290	49	0	49	24
North Dakota	48	18	0	18	12
Nebraska *	48	2	17	19	12
New Hampshire *	44	2	23	25	6
New Jersev	246	53	0	53	6
New Mexico *	66	4	24	28	6

	Number of CWSs Serving More Than 3,300 Persons				Number of CWSs Serving 3,300 or Fewer Persons
Nevada State	35	16	0	16	6
New York	336	50	0	50	18
Ohio	327	52	0	52	6
Oklahoma	173	42	0	42	18
Oregon	123	35	0	35	12
Pennsylvania	349	48	0	48	30
Puerto Rico	98	26	0	26	6
Rhode Island *	28	4	9	13	0
South Carolina	170	40	0	40	6
South Dakota *	55	2	26	28	6
Tennessee	285	46	0	46	12
Texas	1,259	106	0	106	36
Utah	133	34	0	34	6
Virginia	159	39	0	39	12
Virgin Islands	2	2	0	2	2
Vermont *	36	0	34	34	6
Washington	225	46	0	46	30
Wisconsin	185	40	0	40	6
West Virginia *	119	2	40	42	12
Wyoming *	33	0	19	19	6
Total	10,125	1,824	311	2,135	600

* Fifteen states are estimated to partially participate in the assessment. CWSs serving 3,301 to 100,000 people in the partial participation states will receive the supplemental questions but will not receive the questions about their capital needs. CWSs serving more than 100,000 people and 3,300 or fewer people in the partial participation states will receive questions about their capital needs and the supplemental questions.

** This column shows the number of CWSs serving 3,301 to 100,000 people in partially participating states that will receive only the supplemental questions.

*** The small system sample is not used to estimate each state's small system need, instead national averages are applied to the inventory in each state. See Section B.1.c.

Exhibit B-2-2 American Indian and Alaska Native Village Utility Sample Sizes

	Number of AI and ANV Utilities	
	Inventory of Utilities	Sample Size
American Indian Utilities		
Smalls (3,300 & fewer)	220	71
Medium (3,301 to 10k)	47	19
Large (>10k)	15	13
Total	282	103
Alaska Native Village Utilities		
Smalls (3,300 & fewer)	139	41
Medium (3,301 to 10k)	2	2
Large (>10k)	1	1
Total	142	44

B.2.b.iii Stratification Variables

The objective of stratification is to increase the efficiency of the sampling design (thereby reducing the number of systems to be sampled for a given level of precision). Stratified sampling may produce a gain in precision in the estimates of the characteristics of the target population as compared to simple random sampling. In stratified sampling, the target population (i.e., CWSs and AI and ANV utilities) is divided into non-overlapping strata that are internally homogeneous, in that the measurements vary little from one unit to another (i.e., the within-stratum variance is minimized). If the within-stratum variance is relatively small, then a precise estimate of the variable of interest can be obtained with relatively small samples. Each of the strata estimates can be combined to obtain a precise estimate for the overall target population. If the strata are constructed correctly, the target population estimate can be achieved with greater precision and with fewer samples than the estimate obtained from simple random sampling.

The EPA's drinking water programs have historically evaluated CWSs based on (1) the number of persons served and (2) the primary water source (ground water or surface water).² Using total capital need information obtained from prior DWINSAs, the EPA evaluated several classification schemes. This analysis showed that the stratification scheme used in prior assessments (10 strata based on size and source) would be appropriate for the 8th DWINSA. For some states, the EPA may combine the 10,001 to 25,000 and 25,001 to 50,000 size categories within each source category, resulting in 8 rather than 10 strata. The EPA will combine these two size categories only if the sample using 8 strata is more efficient than the sample using 10 strata.

The proposed strata and sampling method for systems serving more than 3,300 persons are as follows:

² For the purposes of the 8th DWINSA, purchased surface water systems are included with ground water systems. This design yields lower within-stratum variance and has been used since the 2nd DWINSA.

Size of Population Served	Water Source	Sample Methodologies
3,301 – 10,000	Ground	Panel approach with 25 percent refresh using a random sample.
3,301 – 10,000	Surface	
10,001 – 25,000	Ground	Panel approach with 25 percent refresh using a random sample. In some states the number of strata will be reduced based on analysis of optimal stratum boundaries. Specifically, in some states systems serving between 10,001 and 50,000 will be in one size group rather than two.
10,001 – 25,000	Surface	
25,001 – 50,000	Ground	
25,001 – 50,000	Surface	
50,001 – 100,000	Ground	Panel approach with 25 percent refresh using a random sample
50,001 – 100,000	Surface	
More than 100,000	Ground	Sampled with certainty
More than 100,000	Surface	

EPA is using the same systems serving 3,300 or fewer persons as were selected for the 7th DWINSA. The 7th DWINSA utilized a national two-stage stratified random sample. The strata for these systems in the 7th DWINSA were as follows:

Water Source	Population Served
Surface Water Systems	Fewer than 101
	101 – 500
	501 – 1,000
	1,001 – 3,300
Ground Water Systems	Fewer than 101
	101 – 500
	501 – 1,000
	1,001 – 3,300

The EPA uses a similar approach for AI utilities, stratifying them based on (1) the number of persons served and (2) the primary water source (ground water and surface water). It also distinguishes between combined water and wastewater utilities and water-only utilities. The EPA uses expenditure data from the 2020 operation and maintenance cost evaluation to estimate total expenditures and the standard deviation in total expenditures, which are used to determine the required sample size for each stratum.

The 8th DWINSA uses the following 20 sampling strata and sampling method for AI utilities:

Utility Type	Water Source	Population Served	Sample Methodologies
Combined Water and Wastewater	Ground Water	LT 501	Simple Random Sample within Each Stratum
		501-1,000	
		1,001-3,300	
		3,300-10,000	
		GT 10,000	
	Surface Water	LT 501	
		501-1,000	
		1,001-3,300	
		3,300-10,000	
		GT 10,000	
Water Only	Ground Water	LT 501	
		501-1,000	
		1,001-3,300	
		3,300-10,000	
		GT 10,000	
	Surface Water	LT 501	
		501-1,000	
		1,001-3,300	
		3,300-10,000	
		GT 10,000	

Three strata are empty. The inventory does not contain combined water and wastewater surface water utilities serving more than 10,000 people. It also does not include water only surface water utilities serving 3,300 to 10,000 people or more than 10,000 people.

The EPA uses a similar approach for ANV utilities, but stratifies utilities based on the type of drinking water distribution and wastewater collection systems. As with the AI utilities, the EPA uses expenditure data from the 2020 operation and maintenance cost evaluation to estimate total expenditures and the standard deviation in total expenditures, which are used to determine the sample sizes.

The 8th DWINSA uses six sampling strata for ANV utilities:

Drinking Water Distribution System	Wastewater Collection System	Sample Methodologies
Pressure	Gravity	Simple Random Sample within Each Stratum
Circulating	Gravity	
Circulating	Vacuum	
Circulating	Low Pressure System (LPS)	
Individual Haul -Washeteria	Honey Bucket -Washeteria	
Community Haul or Water Utility-Only	Community Haul or Wastewater-Only	

B.2.b.iv Sampling Method

As indicated above, for the state survey all CWSs serving populations of more than 100,000 will be sampled with certainty.

For CWSs serving 3,301 to 100,000 persons, all CWSs will be allocated to each stratum within each state, based on the population served and primary water source. The sample size for each stratum in each state will be determined by the sampling strategy outlined above. As previously described, the modified panel approach that will be used for the 8th DWINSA sample will begin with the 7th DWINSA sample and make targeted modifications to account for changes in the sample between the 7th and 8th DWINSAs. The EPA will then “refresh” the sample by randomly replacing 25 percent of the 2020 sample of systems serving 3,301 to 100,000 persons. To refresh the sample, 25 percent of systems serving 3,301 to 100,000 persons will be dropped from the survey and returned to the pool of systems that were not selected in the 7th DWINSA. The “refresh” will then randomly select systems from among the pool of systems that were not in the 2020 sample, the systems that were randomly dropped, and new systems so that the number of systems in the sample reaches the target sample size.

The sampling method for the 25 percent refresh will be similar to the approach used in the 7th DWINSA. An equal probability random sample will be drawn from each stratum. Anticipating a level of non-response, the EPA will over-sample the refresh systems to achieve the desired number of completed data collection instruments. Since the expected response rate for systems serving 3,301 to 100,000 persons is approximately 90 percent, the EPA has increased the sample by approximately 10 percent. However, as discussed below, the DWINSA has consistently achieved a higher response rate than estimated. Therefore, the EPA has included the full sample size estimate in the burden estimate of this ICR.

For systems in partial participation states, a stratified random sample of CWSs will be selected from each partial participation state to determine which systems will receive the supplemental questions. The sample design for these systems will be identical to that used for the medium CWSs in fully participating states.

The same CWSs serving 3,300 and fewer persons that were surveyed for the 7th DWINSA will be surveyed for the 8th DWINSA, with the exception of CWSs that have migrated to a medium or large system strata or ceased operations between the 7th DWINSA and the date of frame validation for the 8th DWINSA. There will be no refresh for these small systems.

AI and ANV utilities will be randomly selected from within each stratum. All systems will be selected in a stratum if the sample would require all but 1 or 2 utilities.

B.2.c Precision Requirements

B.2.c.i Precision Targets

The sampling design for the 8th DWINSA will be implemented at the state level for state medium and large CWSs. The EPA's goal is to be 95 percent confident that the margin of error, when estimating the total capital needs facing these systems in each state, will be plus or minus 15 percent of the total need for full participating states. For example, if the total need for these systems in a state is estimated to be \$2 billion, the EPA will be 95 percent confident that the actual total need is between \$1.7 billion and \$2.3 billion. For partial participation states, the EPA will develop national-level capital need estimates for these systems, although the precision target was applied to establish the state-level size for supplemental questions.

The sampling design for small systems for the 8th DWINSA is the same as was applied for the 7th DWINSA, which was implemented at the national level. EPA will use the same CWSs serving 3,300 and fewer persons that were surveyed for the 7th DWINSA, excluding CWSs which have migrated to a medium or large system strata or ceased operations between the 7th DWINSA and the date of frame validation for the 8th DWINSA. For the 8th DWINSA, EPA's goal is to be 95 percent confident that the margin of error for small system national capital need will be equal to plus or minus 15 percent of the total need for these systems. This precision level is at a national level whereas the level for estimates developed for medium and large CWS is at the state level, but it will not materially reduce the overall precision for total cost estimates at the state level.

The small system random sample was designed to provide an unbiased estimator of total small system 20-year need for systems in 2021. Applying these estimates—adjusted for the new information collected in the 2026 effort—may introduce bias because the systems are a random sample of systems selected in 2021, not 2026. The bias is likely small because the inventory of small systems has not changed dramatically since the 7th DWINSA. Furthermore, EPA will adjust the total need for changes in the inventory. The typical need per system may change, but the 8th DWINSA is collecting information about per system need which EPA will use to modify the estimate of the total need.

For AI and ANV utilities, the EPA's goal is to be 90 percent confident that the margin of error when estimating mean expenditures will be plus or minus 30 percent of mean expenditures within each stratum. For example, if average expenditures are \$200,000, the 90 percent confidence interval will be between \$140,000 and \$260,000.

B.2.c.ii Nonsampling Error

The EPA has developed an assessment approach that will employ several quality assurance techniques to maximize response rates, response accuracy, and processing accuracy to minimize non-sampling error.

Particular emphasis will be placed on maximizing response rates. Standard methods that have proved effective in other surveys involving states and water systems are proposed, including the following:

- The EPA and the states will coordinate in the production of a cover letter for the 8th DWINSA. The EPA's opinion (shared by state drinking water administrators and trade associations) is that surveys on state letterhead will be better received than surveys on the EPA letterhead. Therefore, states can use state-level cover letters signed by a senior state official instead of the EPA letter. The EPA Regions will coordinate the letters to the AI and ANV utilities.
- The states and EPA Regions will place a telephone call to each participating medium or large system for the state survey or utility for the Tribal survey to ensure that they understand the survey process and their role.
- The data collection instrument design, content and format were reviewed by states that participated in the 1st, 2nd, 3rd, 4th, 5th, 6th, and 7th DWINSAs.
- Questions being asked are those that owners or operators of systems and utilities should know. The EPA does not ask questions that require monitoring, research, or calculations on the part of the respondent.
- The data collection instrument design for fully participating states is limited to a cover page of system information and characteristics and one project table, with three tables the system may use to record general information about the system's infrastructure inventory and one set of supplemental questions. By limiting the information requested, the EPA believes that the average CWS respondent in fully participating states can complete the data collection instrument in approximately 6.22 hours. Exhibit A-14 shows the breakdown of the total burden hours for CWSs by system size in states fully participating in the 8th DWINSA.
- Medium CWSs in partial participation states and AI and ANV utilities will receive a survey instrument that consists only of a cover page and their respective supplemental questions. The burden on these systems and utilities will be minimal because they will not receive the full survey instrument. The EPA estimates that the average medium CWS in partial participation states would spend approximately 1.35 hours to respond to the supplemental questions. The EPA estimates that the average AI and ANV utility will spend 2.53 hours to respond to the supplemental questions. The EPA Regions will work with the utilities to help them respond to the survey. Exhibit A-15 shows the breakdown of burden hours for medium CWSs in partial participation states.
- Small CWSs will be contacted by the EPA. EPA will complete the survey instrument through a conversation with small CWS personnel. The EPA will make repeated attempts to contact small system personnel using contact information from the 7th DWINSA. The EPA estimates that the average small CWS, can complete the data collection instrument in approximately 3.35 hours.

- Respondents will be encouraged to call state and regional personnel who will be trained to answer questions. In addition, the EPA will provide technical assistance to states and water systems and the Regions, Navajo Nation, and the AI and ANV utilities.
- The electronic format of the survey will make returning the data collection instrument convenient. Small CWSs will not have to return the data collection instrument, as the EPA will be responsible for completing and submitting the instrument based on input from the phone interviews.

Standard methods to reduce other sources of non-sampling error also will be used:

- The EPA expects complete coverage of the target population using SDWIS, supplemented by state review of all systems.
- The EPA expects complete coverage of the target population of AI and ANV utilities in the frame, supplemented by the review of the EPA Regions and Navajo Nation of all utilities.
- The data collection instrument is pre-coded to improve accuracy by eliminating unnecessary processing steps.

Supplementing these standard methods, the EPA proposes several unique steps to eliminate non-sampling errors that have been developed in concert with organizations representing the states and water systems. These organizations believe that the 8th DWINSA is important and that a high level of participation by all water systems is essential to its success. Because of the substantial commitment being made by states and water systems to the 8th DWINSA, the EPA believes that response rates will be higher than most surveys of similar respondents. To ensure success, states and organizations representing water systems are taking the following steps.

- ***Participation of the states and the EPA Regions.*** Because the 8th DWINSA will be used to allocate DWSRF funds to states, each entity has a strong interest in achieving a high response rate. The EPA believes that their participation will be a key factor in guaranteeing high response rates and low item non-response. Personnel who work with water systems every day are in a strong position to encourage systems to complete the 8th DWINSA form. States have committed to assisting the EPA in achieving a high response rate by participating in follow-up activities. The EPA will provide technical assistance to any system that has questions about the 8th DWINSA. The EPA Regions and IHS will support the AI and ANV utilities to help achieve a high response rate.
- ***Participation of Organizations Representing Water Systems.*** The EPA anticipates public support of organizations representing water systems. The prior assessments were supported by groups such as the American Water Works Association (AWWA), the National Association of Water Companies (NAWC), and the Association of Metropolitan Water Agencies (AMWA).

B.2.d Data Collection Instrument Design

Questions about system characteristics (name, population served, number of connections, and other customary business information) will be pre-populated on all data collection instruments. The respondent needs only to enter accurate information if any pre-populated information is not correct.

The 8th DWINSA is based on a matrix project table that requests a list of capital water system infrastructure projects that the system plans for the period 2026 through 2045. For each project listed, the water system is asked to provide:

- Type of need.
- Reason for need.
- Documentation of need.
- If the project is for new infrastructure or to replace, rehabilitate, or expand existing infrastructure.
- If the project is needed now to protect public health or if it is needed over the next 20 years to continue to provide safe drinking water.
- The federal regulation or state requirement if the project is needed to meet a current federal regulation or state requirement.
- Design capacity of source, storage, and treatment projects.
- Length and diameter of pipe projects.
- Diameter for projects such as meters.
- Cost of the project (if available).
- Date of the cost estimate (if necessary).
- Documentation of cost (if necessary).

For most of these variables, respondents will choose the appropriate “documentation,” “type of need,” “reason for need,” or “regulation or requirement” from the EPA’s “Lists of Codes” (Appendix B). The data collection instrument has been designed to be concise, to avoid jargon, and to avoid ambiguous words or instructions. Terms and formats have been standardized to the extent possible. There is no intentional bias in the ordering of the items.

The data collection instrument will also include the supplemental questions. These questions have been designed to be simple, clear, and concise, to avoid jargon, and to collect information that survey respondents are likely to provide. For the states, supplemental questions will identify potential funding barriers for accessing DWSRF funding. For the AI and ANV utilities, the supplemental questions will collect information about O&M and TMF concerns, annual expenditures, and annual revenue.

B.3 PRE-TESTS AND PILOT TEST

B.3.a Pre-tests

For the 4th DWINSA the data collection instrument and some approaches were modified substantially. The EPA conducted two pre-tests of the data collection instrument for the 4th DWINSA. These pre-tests were conducted by the EPA's contractor, The Cadmus Group LLC. The pre-tests gathered feedback on the effectiveness of the data collection instrument; highlighted imprecise, ambiguous, or redundant questions; and indicated where further inquiry was needed. A pre-test was held in both Maine (four participants) and Montana (three participants). These states were chosen because they were both partial participation states and therefore most of their systems did not participate in the 4th DWINSA. Also, the contractor conducting the pre-tests maintained offices in both these states and by conducting the pre-test in these states was able to reduce costs. The names of the seven systems were provided to the EPA by the 4th DWINSA state contacts. Based on the comments received, the EPA made modifications to the data collection instrument. Since the EPA's pre-tests of the 4th DWINSA data collection instrument were so extensive, and because few changes have been made to the data collection instrument since the 4th DWINSA, the EPA believes that a pre-test is not needed for the 8th DWINSA.

The data collection instrument was modified for the 7th DWINSA by the addition of supplemental questions to gather information on lead service lines (LSLs), pipe and storage iron and steel (I&S) materials, and operator workforce (OpW) issues.

The data collection instrument will be modified for the 8th DWINSA by the removal of the 7th DWINSA questions on lead service lines and OpW, and the addition of other supplemental questions (e.g., O&M and TMF). The I&S questions from the 7th DWINSA remain in the 8th DWINSA but will be prepopulated with 7th DWINSA responses and are clearly identified as optional. The 8th DWINSA supplemental questions for state CWSs were developed based on feedback from the DWINSA Coordinator Workgroup and intra-agency review and were included in the first Public Comment opportunity. Public comments and the EPA's responses are summarized in Appendix C. The supplemental questions for AI and ANV utilities went through an agency review and received stakeholder feedback. The questions do not change any procedural aspects of the DWINSA, so no pre-tests were needed.

B.3.b Pilot Test

To eliminate unnecessary burden on states and water systems, it has been decided that no pilot test for the 8th DWINSA will be conducted. A pilot test was conducted for the 1st DWINSA and consisted of 60 CWSs from New York and Texas.

Starting with the 6th DWINSA, the EPA began delivering all data collection instruments electronically rather than mailing hard copies, as was done for the 1st, 2nd, 3rd, 4th, and 5th DWINSAs. States used the electronic data collection instruments in the 6th and 7th DWINSAs and expressed no concerns about using the electronic format. The EPA believes this approach has been well tested and has proven to be successful; therefore, it is not necessary to repeat this testing step.

B.4 COLLECTION METHODS AND FOLLOW-UP

B.4.a Collection Method

The proposed collection method is an electronic survey. The data collection instrument including the Lists of Codes will be sent to the states via e-mail. State drinking water agencies will provide the data collection instrument (with the project table prepopulated for systems that participated in 7th DWINSA and blank for those that did not) and other necessary documents to the systems in the sample. They will assist systems with questions on survey completion and provide follow-up if the data collection instrument has not been returned in 30 days. For a complete description of the follow-up procedures proposed to increase the response rate, see section B.2.c.ii.

The EPA also will conduct phone conversations with a sample of CWSs serving 3,300 or fewer persons included in the 7th DWINSA to update their capital investment needs and collect their responses to the supplemental questions.

The proposed collection method for the AI and ANV utility survey is for the EPA Region to contact each utility and provide technical support to the respondent. By having EPA Regions and IHS assisting the survey in this manner, the information collection burden on these water system respondents will be minimized.

B.4.b Survey Response and Follow-up

The target response rate (defined as the ratio of responses to eligible respondents) for the 8th DWINSA is 90 percent. The EPA realizes that this is an ambitious target, but the EPA believes that there are special circumstances that warrant such a target. Also, overall response rates of 94, 97, 96, 93, 97, 99.7 and 97 percent were achieved in the 1st, 2nd, 3rd, 4th, 5th, 6th and 7th surveys, respectively. In the first seven surveys, the EPA conducted the following activities to achieve that high response rate.

- ***Seek Support from the Respondent Population.*** This is a national survey of infrastructure needs for drinking water systems. The EPA will work to bring to the attention of water systems, as well as all national organizations representing these systems, the importance of the DWINSA results. As with the previous seven surveys, national organizations will be contacted by the EPA to seek their endorsement of the DWINSA and to communicate to their members the importance of a high response rate to their members. As discussed in Section B.2.c, in past surveys, organizations have provided access to their newsletters and magazines to publicize and endorse participation in the DWINSA. For the 8th Survey, the EPA will seek similar efforts by these organizations.
- ***Follow-up by States and Respondent Peer Groups.*** Since a majority of states have indicated their willingness to participate in follow-up activities, the EPA has requested that state personnel, most of whom are personally familiar with the respondents, conduct follow-up procedures including the use of reminder letters and telephone calls to systems that have not responded with the needed information or documentation. If the follow-up fails after three attempts (one reminder letter plus two telephone follow-up calls), the EPA plans to shift to a second approach of peer-group follow-up by members of a trade association, such as AWWA.

B.5 ANALYZING AND REPORTING SURVEY RESULTS

B.5.a Data Preparation

State personnel will check all cost data and documentation for CWSs serving more than 3,300 persons to ensure that it is consistent with state and national standards. States will then send the completed and reviewed data collection instruments to the EPA for a second round of review by the EPA. As mentioned in Section B.5.a., small system responses will be collected via phone calls by the EPA.

Once data have been checked, the EPA will verify the data. Missing cost data will be modeled, using other information provided by the respondents on the data collection instrument or applicable cost modeling information. When modeling is insufficient, missing data will be imputed using standard methods such as cell means and regression. The sample of water systems will be weighted so that stratum estimates can be summed to prepare state-level estimates for medium and large CWSs in the 8th DWINSA.

For the AI and ANV DWINSAs, EPA Regions will check supplemental questions responses for completeness and submit the survey forms to the data system. No data preparation is needed other than ensuring data upload.

B.5.b Analysis

The EPA will prepare a report that tabulates the results of the 8th DWINSA and explains the precision of the estimates of total capital needs. Examples of statistics that will be produced include:

- Total capital improvement needs by state and by types of need for medium and large CWSs.
- Total capital improvement needs for the nation by types of need for small CWSs, Tribal water systems, and NPNCWSs.
- Total capital improvement needs by domains within the total population, e.g., systems serving populations greater than 100,000.
- Total capital improvement needs by system type.
- Service line materials by state for small, medium and large CWSs based on compliance data submitted for the LCRR or LCRI.
- Potential funding barriers for accessing DWSRF funding by system size and state.
- Average annual revenue and expenditures of AI and ANV utilities.
- Summary of O&M and TMF issues and concerns facing AI and ANV utilities.
- Standard errors calculated for key statistics.

The analysis will be similar to that of previous DWINSAs.

B.5.c Reporting Results

The 8th DWINSA results will be made available to the EPA and the public through:

- A report that is submitted to Congress on drinking water infrastructure needs. This report will be made available to the public through the EPA's website.
- The EPA will also make information available on their website.