

**Information Collection Request for the 2007 Drinking Water
Infrastructure Needs Survey and Assessment (DWINSA)
(Supporting Statement for ROCIS)**

August 2006

Prepared by:

**The Cadmus Group, Inc.
57 Water Street
Watertown, MA 02472**

Prepared for:

**U.S. Environmental Protection Agency
Office of Ground Water and Drinking Water
Drinking Water Protection Division**

This page intentionally left blank.

Table of Contents

PART A OF THE SUPPORTING STATEMENT.....	5
A.1 IDENTIFICATION OF THE INFORMATION COLLECTION.....	5
A.1.a Title of the Information Collection Request.....	5
A.1.b Short Characterization.....	5
A.2 NEED FOR AND USE OF THE COLLECTION.....	7
A.2.a Authority and Need for the Collection.....	7
A.2.b Use and Users of the Information.....	7
A.3 NONDUPLICATION, CONSULTATIONS, AND OTHER COLLECTION CRITERIA.....	9
A.3.a Nonduplication.....	9
A.3.b Public Notice Required Prior to ICR Submission to OMB.....	10
A.3.c Consultations.....	10
A.3.d Effects of Less Frequent Collection.....	11
A.3.e General Guidelines.....	11
A.3.f Confidentiality Questions.....	12
A.3.g Sensitive Questions.....	12
A.4 THE RESPONDENTS AND THE INFORMATION REQUESTED.....	13
A.4.a Respondents/NAICS Codes.....	13
A.4.b Information Requested.....	14
A.4.b.i Data Items.....	14
A.4.b.ii Respondent Activities.....	16
A.5 INFORMATION COLLECTED: AGENCY ACTIVITIES, COLLECTION METHODOLOGY, AND INFORMATION MANAGEMENT.....	19
A.5.a Agency Activities.....	19
A.5.b Collection Methodology and Management.....	20
A.5.c Small Entity Flexibility.....	24
A.5.d Collection Schedule.....	24
A.6 ESTIMATING THE BURDEN AND COST OF THE COLLECTION.....	25
A.6.a Respondent Burden.....	25
A.6.a.i Burden to Public Water Systems.....	25
A.6.a.ii Burden to Primacy Agencies.....	29
A.6.b Respondent Costs.....	33
A.6.b.i Costs to Public Water Systems.....	33
A.6.b.ii Cost to States.....	34
A.6.c Agency Burden and Cost.....	35
A.6.d Estimating Respondent Universe and Total Burden and Costs.....	38
A.6.e Bottom Line Burden Hours and Costs.....	38
A.6.f Reasons for Change in Burden.....	41
A.6.g Burden Statement.....	41
 PART B OF THE SUPPORTING STATEMENT (FOR STATISTICAL SURVEYS).....	 43
B.1 Survey Objectives, Key Variables and Other Preliminaries.....	45
B.1.a Survey Objectives.....	45
B.1.b Key Variables.....	45
B.1.c Statistical Approach.....	46
B.1.d Feasibility.....	47
B.2 Survey Design.....	49
B.2.a Target Population and Coverage.....	49
B.2.b Sample Design.....	50
B.2.b.i Sampling Frame.....	50

B.2.b.ii	Sample Size.....	51
B.2.b.iii	Stratification Variables.....	54
B.2.b.iv	Sampling Method.....	55
B.2.b.v	Multi-Stage Sampling.....	56
B.2.c	Precision Requirements.....	56
B.2.c.i	Precision Targets.....	56
B.2.c.ii	Nonsampling Error.....	57
B.2.d	Data Collection Instrument Design.....	59
B.3	PRE-TESTS AND PILOT TEST.....	61
B.3.a	Pre-tests.....	61
B.3.b	Pilot Test.....	61
B.4	COLLECTION METHODS AND FOLLOW-UP.....	63
B.4.a	Collection Method.....	63
B.4.b	Survey Response and Follow-up.....	63
B.5	ANALYZING AND REPORTING SURVEY RESULTS.....	65
B.5.a	Data Preparation.....	65
B.5.b	Analysis.....	65
B.5.c	Reporting Results.....	65
Appendix A Public Notice Required Prior to ICR Submission to OMB.....		67
Appendix B Data Collection Instrument and Lists of Codes.....		71
Appendix C Comments and Response to Comments Received on the 1 st <i>Federal Register</i> Notice.....		91

LIST OF TABLES

Exhibit A-5-1 Collection Schedule.....	24
Exhibit A-6-1 Estimated Unit Burden for Systems Serving More Than 50,000 People.....	26
Exhibit A-6-2 Estimated Unit Burden for Systems Serving 3,301 – 50,000 People.....	28
Exhibit A-6-3 Estimated Unit Burden for Small CWSs.....	28
Exhibit A-6-4 Overall State Burden Summary.....	29
Exhibit A-6-5 State Unit Burden for Up-Front Activities.....	30
Exhibit A-6-6 State Unit Burden for Systems Serving More Than 50,000 People.....	31
Exhibit A-6-7 State Unit Burden for Systems Serving 3,301 – 50,000 People.....	33
Exhibit A-6-8 State Unit Burden for Small Systems.....	33
Exhibit A-6-9 Total Burden and Cost to Water Systems.....	34
Exhibit A-6-10 Total Burden and Cost to States.....	35
Exhibit A-6-11 Burden/Cost to EPA (Excluding Contractor Activities).....	36
Exhibit A-6-12 Burden/Cost of Contractor Activities.....	37
Exhibit A-6-13 Bottom Line Respondent Burden.....	38
Exhibit A-6-14 Burden Hours and Costs for Respondents per Year.....	38
Exhibit A-6-15 Bottom Line Burden Hours and Costs for EPA (including its contractor).....	39
Exhibit A-6-16 Disaggregated Burden by Affected Information Collection (IC) Entities.....	40
Exhibit B-2-1 State Sample Sizes.....	52

This page intentionally left blank.

PART A OF THE SUPPORTING STATEMENT

A.1 IDENTIFICATION OF THE INFORMATION COLLECTION

A.1.a Title of the Information Collection Request

The title of this information collection request (ICR) is *Information Collection Request for the 2007 Drinking Water Infrastructure Needs Survey and Assessment (DWINSA)*¹. The Office of Management and Budget (OMB) control number for this ICR is 2040-new; EPA ICR No. 2234.01.

A.1.b Short Characterization

The Environmental Protection Agency (EPA) will conduct an assessment to estimate the capital investment needs for drinking water systems eligible to receive Drinking Water State Revolving Fund (DWSRF) monies. The nationwide assessment will be conducted by the Drinking Water Protection Division (DWPD) of EPA's Office of Ground Water and Drinking Water (OGWDW). The data collection is authorized by Sections 1452(h) and 1452(i)(4) of the Safe Drinking Water Act (SDWA) and will be used to estimate the cost of providing safe drinking water to consumers over a 20-year period. The data from the report will also be used to allot DWSRF monies among states. The focus of the 2007 DWINSA is collecting information on systems' needs and on the projected costs associated with those needs. All states² have committed to help EPA administer the 2007 DWINSA with at least the minimum of activities. Thirteen states have chosen not to participate in the statistical portion of the survey (i.e., collecting data from systems serving 3,301 – 100,000 people). For the states that opt out of the statistical portion of the survey, the needs of the participating states will be used to determine the needs for systems serving 3,301 – 100,000 people. All states with systems serving more than 100,000 people will participate in census portion of the survey.

For the 2007 DWINSA, EPA will conduct a census of all community water systems (CWSs) serving populations more than 100,000 and select a random sample of CWSs that serve populations of 3,301 – 100,000. EPA will conduct a site visit for 600 randomly selected CWSs serving 3,300 and fewer people and will estimate infrastructure needs for not-for-profit noncommunity water systems (NPNCWSs)³ in the states and all Alaska Native and American Indian systems based on the 1999 DWINSA.

EPA will send the data collection instrument to all CWSs serving more than 100,000 people and a random sample of CWSs serving between 3,301 and 100,000 people (i.e., medium⁴ and large⁵ system sample). The data collection instrument consists of project tables in which the water systems list all their capital improvement projects through the year 2026. The same data collection instrument will be used for CWSs serving a population of 3,300 and fewer (i.e., small system sample); however, the data collection instrument will be completed by an EPA site visit contractor.

¹ EPA's previous assessments of infrastructure need in 1995 and 1999 were called "needs surveys" because the assessment relied primarily on survey methods. In 2003, EPA relied in part on surveys but also on analysis of previous survey data. Accordingly, the term "assessment" is more appropriate. Hereinafter, these studies will be referred to as "assessments."

² The term states refer to all 50 states, as well as the U.S. territories (Guam, U.S. Virgin Islands, Northern Mariana Islands, and American Samoa), Puerto Rico, and the District of Columbia.

³ NPNCWSs are also eligible for DWSRF funding.

⁴ Medium systems refer to systems serving between 3,301 and 50,000 people.

⁵ Large systems refer to systems serving more than 50,000 people.

The effort will involve 3,193 respondents (3,137 CWSs and 56 states⁶), requiring 46,030 hours at a total cost to the respondents of \$1,541,361. Section A.6, *Estimating the Burden and Cost of the Collection*, provides a detailed description of the unit burden and costs for this collection. The average water system burden per response is 7.25 hours. The average burden per response to complete the data collection instrument is 8.20 hours per response.⁷ (This burden is higher because the small systems do not complete the data collection instrument.)

⁶ Fifty-six states include the 50 states plus the District of Columbia, Puerto Rico, Northern Mariana Islands, American Samoa, Guam, and the U.S. Virgin Islands. For several of these entities, primacy activities are actually implemented by EPA Regional offices. However, as a simplifying assumption, they are included with the states for respondent calculations under this ICR.

⁷ See Exhibit A-6-9.

A.2 NEED FOR AND USE OF THE COLLECTION

A.2.a Authority and Need for the Collection

EPA (the Agency) is conducting this DWINSA pursuant to its authority under Sections 1452(h) and 1452(i)(4) of the SDWA. Section 1452(h) requires that “the Administrator shall conduct an assessment of water system capital improvements needs of all eligible public water systems in the United States and submit a report to the Congress containing the results of such assessment within 180 days after the date of the enactment of the SDWA Amendments of 1996 and every 4 years thereafter.”

A.2.b Use and Users of the Information

The results of the 2007 DWINSA will be used as a basis for allocation of DWSRF funds among states. In addition, many water systems have empirical data on the cost of compliance with SDWA regulations. A national assessment will improve the Agency’s ability to gauge the real capital cost of SDWA regulations.

EPA will collect two types of system-specific information: (1) system inventory and characteristics data (i.e., name and address of the system, contact person, population served, total design capacity, number of connections, primary source, whether the system is privately or publicly owned, and whether the system purchases/sells water from/to another public water system (PWS)); and (2) information on capital improvement projects. The specific uses of each data type vary. EPA will use system inventory and characteristics data to characterize CWSs nationwide, and, in some cases, to model individual systems’ capital improvement projects. EPA will use all data collected to estimate state and national needs.

On 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. EPA will collect information on the proposed infrastructure to be installed, replaced, rehabilitated, upgraded, or expanded. EPA will use the information to assess project reasonableness and develop the cost models.

The respondent will also identify either a documented cost estimate for the project or will provide adequate information so that EPA can model the cost of the project. The information needed to model the cost will depend on the type of need. For example, EPA may collect information on the type and number of valves or the diameter and length of transmission or distribution lines. EPA expects that modeling will be required to project the capital needs for small systems, many medium systems, and some large systems. For the 2003 DWINSA, approximately 18 percent of the projects reported had documented costs. The costs for the remaining 82 percent of projects were modeled.

The data collected by the 2007 DWINSA will likely have several secondary uses, both inside and outside of EPA. For example, EPA will use the information to support various program activities, such as the development of general enforcement strategies and new regulations. Congress may use occurrence and cost information in considering new drinking water legislation. States have indicated to EPA that they plan to use the data collected to help identify projects that should be included on the state’s DWSRF priority list and to implement capacity development strategies. The public may use information on costs associated with SDWA compliance.

This page intentionally left blank.

A.3 NONDUPLICATION, CONSULTATIONS, AND OTHER COLLECTION CRITERIA

The following sections verify that this information collection satisfies the OMB's nonduplication and consultation guidelines, and does not duplicate another collection.

A.3.a Nonduplication

To the best of EPA's knowledge, up-to-date state-by-state information on water system's capital needs is not available from any other source. Some of the data collection efforts EPA considered include the following:

- **Safe Drinking Water Information System (SDWIS).** Inventory data and information on system characteristics have been collected by states and regions and entered into the SDWIS. For the statistical sample, EPA will pre-print the SDWIS system characteristics data (i.e., name and address of the system, contact person, address, population served, total design capacity, number of connections, primary source, whether the system is privately or publicly owned, and whether the system purchases/sells water from/to another PWS) on the 2007 DWINSA form and ask the respondents to provide information only if the SDWIS data are inaccurate or missing. SDWIS does not contain information on water systems' capital needs.
- **Community Water System Survey (CWSS).** EPA completed a statistical survey in 2000 that focused on the operating and financial characteristics of CWSs. The CWSS is addressed in the ICR for National Survey of the Financial and Operating Characteristics of Community Water Suppliers. The CWSS had a different objective than the DWINSA. The CWSS was designed to characterize the technical and financial aspects of CWSs. In contrast, the DWINSA will be used to develop national estimates of capital needs. In addition, the CWSS's targeted precision was on a national basis; whereas the DWINSA will provide state-by-state estimates. It is anticipated that EPA will conduct a new CWSS in 2007. Where possible, EPA will combine any efforts for the CWSS and DWINSA, such as the small system site visits. Even though the data collected are different, the burden may be reduced if only one site visit is conducted at a system to collect data for both the CWSS and DWINSA.
- **Economic Analyses (EAs) for National Primary Drinking Water Regulations.** The Agency has developed EAs (formerly referred to as Regulatory Impact Analyses) for its National Primary Drinking Water Regulations. These documents estimate the costs of complying with proposed regulations. The scope of the EAs is limited to the cost associated with the implementation of a given proposed regulation. EAs do not include an estimate for on-going capital projects to maintain compliance with existing regulations. Therefore, the EAs are not an adequate substitute for the DWINSA. In addition, the EAs provide nationwide estimates. As discussed above, EPA is conducting the DWINSA because the Agency needs a state-by-state estimate to develop the allocation formula for the DWSRF. Also, many EAs are several years out of date. They do not consider currently available contaminant occurrence data, or current or emerging treatment technology costs.
- **State Needs Surveys.** Several states have conducted needs surveys of their own drinking water systems. The state results cannot be extrapolated to the nation as a whole because the state surveys do not use consistent methodologies and do not account for national variations in system characteristics and needs.

- American Water Works Association (AWWA) Water Industry Data Base: Utility Profiles. Like the CWSS, this AWWA survey and associated database focuses on financial and operating characteristics of water systems. Moreover, the AWWA database is not statistically representative.
- Drinking Water Treatment Screener and Detailed Questionnaire. EPA plans to conduct a survey of drinking water treatment plants as part of its effort to evaluate the need for national effluent guidelines regulations for the drinking water treatment point source category. This survey will request information on drinking water treatment plants operating during the 2005 calendar year. This survey will be conducted under the authority of Section 308 of the Clean Water Act (Federal Water Pollution Control Act, 33 U.S.C. Section 1318). This survey will only collect data from systems serving more than 10,000 people and information specific to drinking water treatment residuals. As discussed above, EPA is conducting the DWINSA because it requires a state-by-state estimate to develop the allocation formula for the DWSRF. The data in this survey will not support that effort.
- 1995, 1999, and 2003 DWINSAs. Under the SDWA, EPA must conduct the DWINSA every 4 years. The approach for the 2007 DWINSA will incorporate some data collected during the previous assessments, as well as “lessons learned” from the earlier assessments. In addition, the approach for the 2007 DWINSA ensures that up-to-date data on infrastructure needs are collected for all CWSs. Small, medium, and large CWSs regulated by the states reported approximately 94 percent of the national need for the 2003 DWINSA. The remaining 6 percent were for needs associated with American Indian and Alaska Native water systems, NPNCWSs, and recently promulgated regulations (including the new Arsenic Rule).

A.3.b Public Notice Required Prior to ICR Submission to OMB

To comply with the 1995 Amendments to the Paperwork Reduction Act (PRA), EPA solicited public comment on this ICR for a 60-day period before it was submitted to OMB. Specifically, EPA published a notice in the *Federal Register* ([FR](#)) requesting comment on the estimated respondent burden and other aspects of this ICR ([71FR32344](#)). This notice is included in Appendix A. Before submission to OMB, EPA considered any comments received and determined if any adjustments were needed to the burden and cost calculations or to the supporting statement for this ICR. Comments received and EPA’s responses are included in Appendix C. An additional *Federal Register* notice will be published when this ICR is submitted to OMB. The public comment period for this additional notice is 30 days.

A.3.c Consultations

For the 2007 DWINSA, EPA assembled a workgroup that consisted of EPA headquarter, EPA regional, and state representatives. In October 2005 and April 2006, EPA held meetings with workgroup members. The purpose of the meetings was to gather information on state and/or regional concerns, to discuss lessons learned during the 2003 DWINSA, and to discuss new policies for the 2007 DWINSA. The information gathered during the meetings was used to develop the methodology for the 2007 DWINSA. EPA also convened an informal peer input group to involve stakeholders such as trade associations. In addition, EPA conducted a pre-test of the data collection instrument (see B.3 for more information on the pre-test) and a formal peer review of the 2007 DWINSA statistical methodology and policies. Based on comments received from the peer input group, the peer review, and the pre-test, EPA made modifications to the data collection instrument, statistical procedures, and survey polices.

Specific changes EPA will make to the statistical methodology based on the results of the peer review include:

1. The survey methodology required that at least one system be selected from each stratum that contains at least one system. Based on the suggestions of reviewers the sampling plan will be adjusted to also ensure that at least two systems are selected from each stratum that contains more than one system to allow for variance calculations by strata and maintain the precision of the estimates.
2. Upon completion of the data collection, EPA will consider recommendations for alternative approaches to determining the system weights.

EPA will also make changes to the survey policies based on the results of the peer review. Two of the changes that EPA will make are:

1. EPA will clarify and provide more guidance to state reviewers on what type of documentation will be acceptable for the different types of water system projects.
2. In the survey instructions, EPA will add the definition of “need” as defined for the 2007 DWINSA.

A.3.d Effects of Less Frequent Collection

The 2007 DWINSA is a single collection and does not involve periodic reporting or recordkeeping.

A.3.e General Guidelines

The 2007 DWINSA does not violate any guidelines for information collection activities specified by OMB. Specifically, the 2007 DWINSA respondents are not required to:

- Report information to EPA more often than quarterly.
- Retain records for more than 3 years.
- Complete the data collection instrument in fewer than 30 days.
- Maintain or provide information in a format other than that in which it is customarily maintained.
- Submit proprietary, trade secret, or other confidential information.
- Submit more than one original and two copies of any document.

The information collection:

- Is a statistical assessment designed to produce data that can be generalized to the universe of the study (see Section B.2).

- Does not provide remuneration to participants.
- Will transcribe information collected into an automated format.
- Is designed with small entities particularly in mind (see Part A.5.c).
- Does not concern grants or grantees.
- Is voluntary.

A.3.f Confidentiality Questions

This information collection does not require the respondent to disclose any confidential information. Respondents are not obliged to respond to this strictly voluntary information collection. Further, respondents could eliminate any confidential business information from their reply.

A.3.g Sensitive Questions

The 2007 DWINSA does not ask sensitive questions, such as those pertaining to sexual attitudes or behavior or religious beliefs.

A.4 THE RESPONDENTS AND THE INFORMATION REQUESTED

A.4.a Respondents/NAICS Codes

NAICS Codes

The respondents for the 2007 DWINSA are CWSs and states. According to 40 CFR Part 141.2, a CWS is a “public water system which serves at least 15 service connections used by year-round residents or regularly serve at least 25 year-round residents.” The North American Industry Classification System (NAICS) code for PWSs is 221310. State agencies that include drinking water programs are classified as NAICS code 924110 (Administration of Air and Water Resources and Solid Waste Management Programs) or 926130 (Regulation and Administration of Communications, Electric, Gas, and Other Utilities). Ancillary systems (i.e., those that supplement the function of other establishments like factories, power plants, mobile home parks, etc.) cannot be categorized in a single NAICS code. For ancillary systems, the NAICS code is that of the primary establishment or industry.

Respondents

EPA will gather information from large and medium CWSs, as well as a limited number of small water systems. Because of their variability and significant contribution to the overall drinking water capital investment need, EPA will survey 999 systems that serve more than 50,000 people. Systems serving more than 100,000 will be sampled with certainty (589 systems). EPA will select a statistically representative sample of systems serving 50,001 – 100,000 (410 systems).

There are 8,109 systems that serve populations of 3,301 – 50,000. Surveying all of these systems would impose a large burden on respondents, EPA, and states.⁸ Therefore, EPA will select a statistically representative sample of systems serving 3,301 – 50,000. This will result in 1,538 systems receiving the mailed data collection instrument. Part B of the supporting statement describes the sampling methodology.

There are 41,398 small systems.⁹ Surveying all of these systems would impose a large burden on respondents, EPA, and states. Therefore, EPA will select a statistically representative sample of 600 systems serving 3,300 and fewer people. The 2007 DWINSAs will be completed by EPA site visitors. Part B of the supporting statement describes the sampling methodology.

EPA will not collect new needs data from NPNCWSs, Alaska Native water systems, or American Indian water systems.

In addition, 56 states (50 states plus the District of Columbia, Puerto Rico, Northern Mariana Islands, American Samoa, Guam, and the U.S. Virgin Islands) will provide support and information for the 2007 DWINSA.

⁸ States include Regional offices that act as the primacy agency for drinking water programs in Wyoming, District of Columbia, U.S. Virgin Islands, Guam, American Samoa, and the Northern Mariana Islands.

⁹ The DWSRF defines small systems as those with populations of 10,000 or fewer, a different threshold than that listed in the typical drinking water regulation (i.e., systems serving 3,300 or fewer persons). In the final Report to Congress resulting from this survey, the results will be presented in different formats to accommodate various definitions of small system.

A.4.b Information Requested

As previously discussed, EPA will collect two types of information from systems: (1) system inventory and characteristics; and (2) information on capital improvement projects. EPA anticipates that respondents will provide varying levels of information by system size category. Based on experience from the previous three DWINSA, EPA expects larger systems to have a good understanding of their capital needs and the costs for meeting them. Almost all of these systems will have detailed Capital Improvement Plans (CIPs). Most large systems will be capable of providing accurate information on cost. Most medium systems can provide reliable data on their needs and some can provide cost estimates for meeting their needs. The information that respondents will be asked to provide is generally maintained and reported as a function of the management and operation of the water system.

A.4.b.i Data Items

Medium and Large CWSs (Systems serving more than 3,300 people)

The data collection instrument asks respondents to verify or correct system characteristic information (i.e., name and address of the system, contact person, address, population served, total design capacity, number of connections, primary source, whether the system is privately or publicly owned, and whether the system purchases/sells water from/to another PWS). It is Customary Business Practice (CBP) for the system to maintain this information. The respondent will either indicate that the information is correct as printed or enter the correct information in the space provided. States will verify this information in advance of the data collection instrument being sent to the systems. Based on previous assessments, EPA anticipates that very few systems will need to correct the information provided.

In addition, the respondent is asked to provide information on tables associated with specific types of projects:

- Source.
- Treatment.
- Finished or Treated Water Storage, Pumping, and Other.
- Transmission and Distribution.
- Backflow Prevention Devices/Assemblies, Flushing Hydrants, Service Lines, Valves, and Water Meters.

For each of these projects, the respondent is asked to:

- Briefly describe the needed capital projects (e.g., “routine distribution system replacement,” “filtration plant upgrade,” “high service pump replacement,” “corrosion control treatment,” or “storage tank rehabilitation”). Information is collected on a project-by-project basis because it is most commonly available to respondents in that form, and because documentation, when available, is usually developed on a project-by-project basis.

- Provide the code that best describes the project from List 1 of the Lists of Codes and that best describes the reason for the need from List 2. EPA will use this information to:
 - Develop separate cost estimates for source water treatment, transmission, storage, distribution, and other needs. (EPA will disaggregate the costs when projects meet multiple needs, if necessary.)
 - Help verify that adequate documentation of the need has been submitted.
 - Help determine if the project is an allowable need.
 - Help gauge cost-reasonableness.
- Indicate if the project is to install new infrastructure to meet current population demands, replace old infrastructure, expand or upgrade existing infrastructure (such as treatment plants to meet current population demands, or rehabilitate existing infrastructure).
- Indicate if the project is needed now to protect public health or not needed now, but will be necessary to continue providing safe drinking water over the next 20 years.
- Indicate if the project is associated with a regulation, the codes for which are listed on List 3 in the Lists of Codes. EPA will use this information to determine which needs are required because of an SDWA regulation or state requirement.
- Provide design capacity when applicable—millions of gallons per day (MGD) for treatment and pumping; or millions of gallons (MG) for storage; the diameter and number of feet of distribution or transmission lines that will be replaced or added; or the size and number of backflow prevention devices/assemblies, flushing hydrants, service lines, valves, and water meters. EPA will use these parameters to model project costs.
- If available, provide the capital cost estimate and year and month (if known) of the estimate. EPA will use this information to identify the cost of the project. The year and month are important because they will allow EPA to account for differences in the value of money over different years and to convert all costs to a common year.
- Provide inventory data on the total length of pipe in the water system. This information will only be required for water systems that submit pipe projects, but do not have independent documentation (i.e., planning document, sanitary survey, or leak and break records). It is expected that not all systems will need to provide this inventory information. This information is necessary to allow EPA to determine that the need reported is what is reasonably expected to be replaced or rehabilitated in a 20-year timeframe.
- Indicate the type of documentation from List 4 of the Lists of Codes that explains why the project is needed and, if a cost estimate is available, indicate the documentation that explains the breakdown of the cost. This will verify the cost for the project. NOTE: EPA does not expect systems to develop cost estimates for the purposes of the 2007 DWINSAs.

The respondent is also asked to provide his or her name, title, address, phone number, and e-mail address. This information is requested in case EPA or the state must contact the respondent for clarification or explanation of any response.

The respondent is asked to attach documentation for all needs and costs reported in the 2007 DWINSA. Systems are encouraged to provide inventory data on their systems. Only where noted above will the inventory data be required.

The data collection instrument is attached as Appendix B.

Small CWSs (Systems serving 3,300 and fewer people)

As discussed above, EPA will collect data from small systems through site visits. EPA will use the large and medium system data collection instrument for the site visits. Contractor personnel will complete the instrument during the site visit. EPA anticipates that most system operators will make themselves available to accompany contractor personnel. The operators may be asked very basic questions about the physical design of the plant and configuration of their system.

American Indian and Alaska Native Water Systems and NPNCWSs

EPA will base the needs of Alaska Native and American Indian water systems and NPNCWSs on the 1999 DWINSA results. EPA conducted site visits in 1999 to estimate the needs of these systems and EPA will update these estimates for the 2007 DWINSA.

A.4.b.ii Respondent Activities

Medium and Large CWSs (Systems serving more than 3,300 people)

To complete the data collection instrument, the following activities are anticipated for medium and large CWSs:

- Participate in an informational telephone call from the state. Respondents will receive a call from the state describing the purpose of the DWINSA, the information that will be requested, and the timetable for completing and returning the data collection instrument.
- Read the cover letter and data collection instructions. Respondents will review the cover letter and instructions accompanying the data collection instrument.
- Collect and copy supporting documentation. Respondents will locate the necessary supporting documentation in system files and copy it.
- Complete the data collection instrument and inventory data. Respondents will fill out the data collection instrument and attach supporting documentation. Inventory data must be filled out for any pipe project that is submitted without independent documentation of need (e.g., a planning document).

In addition, some systems may contact states (or an EPA-established helpline) to obtain clarifying information on the data collection instrument.

Small CWSs (Systems serving 3,300 and fewer people)

The 2007 DWINSA methodology has been designed to minimize the burden on small CWSs. These systems' role will be limited to answering basic questions during a phone call and accompanying the 2007 DWINSA team during the site visit. They will:

- Participate in an informational/scheduling telephone call from EPA. Respondents will receive a call from EPA to describe the purpose of the DWINSA and to schedule the site visit.
- Answer simple questions posed by the 2007 DWINSA team. Respondents will be expected to answer very basic questions about the physical design of the plant, system configuration, and capital needs.

State Activities

All states have committed to help EPA administer the 2007 DWINSA with at least the minimum of activities. Region 3 (Philadelphia) will act as the state for the District of Columbia. Region 8 (Denver) will act as the state for Wyoming. Region 9 (San Francisco) will perform state activities for the three Pacific Islands (Northern Mariana Islands, American Samoa, and Guam). However, 13 states have chosen not to participate in the statistical portion of the survey (i.e., collecting data from systems serving 3,301 – 100,000 people). For non-participating states, the needs of the participating states will be used to determine an average need per strata. This will be applied to the inventory of systems in non-participating states to estimate the needs for systems serving 3,301 – 100,000 people. All states that have systems serving more than 100,000 people will participate in the census portion of the survey.

The activities described in this section represent a level of participation that will ensure nationally consistent results. Some states will participate at a higher level.

State Up-Front Activities

This first activity category includes the states' "fixed burden" for helping EPA prepare for the 2007 DWINSA.

- Participate in training and other pre-mailout efforts. This activity includes participating in training sessions offered to the states and becoming familiar with the survey design and policies. In addition, it includes activities such as reviewing the draft data collection instrument.
- Help EPA verify SDWIS data. There are several important variables for which SDWIS data must be verified. Critical inventory data for the statistical sample will need to be reviewed by states. Such data include PWS identification number (PWSID), system name, address, telephone numbers (if any), primary source, population served, number of service connections, whether the facility is publicly or privately owned, and whether the system is a consecutive system. In addition, states will need to review address information to ensure the street address for each system selected in the sample is accurate. To help states with this verification activity, EPA will provide the information that must be reviewed in electronic form.
- Perform miscellaneous administrative activities. States will perform various administrative duties prior to the 2007 DWINSA (e.g., establishing system files). In

addition, state management will explain the 2007 DWINSA to staff and allocate resources.

State Data Collection Activities for Medium and Large CWSs (Systems serving more than 3,300 people)

States will conduct the following activities for medium and large CWSs during the data collection phase of the 2007 DWINSA:

- Telephone systems to ensure participation. To improve response rates, states that participate in the 2007 DWINSA will telephone the medium and large systems early in the process to ensure that they have received the survey package and understand how to complete the data collection instrument.
- Provide technical assistance. Participating states will provide technical assistance to systems by answering their questions about the data collection instrument and how needs should be represented.
- Call systems that do not return the data collection instrument by a certain date. To improve response rates, participating states will telephone systems that have not returned their assessment by a specific date to encourage participation.
- Review completed data collection instruments and documentation. The data collection instrument will be returned directly to the state. State personnel will have the opportunity to review the information on the data collection instrument, as well as any accompanying documentation.
- Discuss results with EPA. After the state reviews the submission and documentation, the state forwards the data to EPA for review and data entry. EPA performs a second quality control/quality assurance check to ensure all data are documented and reasonable. Any differences of opinion regarding the documentation of the data will be resolved by EPA and the state.

State Data Collection Activities for Small CWSs (Systems serving 3,300 and fewer people)

States will conduct the following activities for small CWSs during the data collection phase of the 2007 DWINSA:

- Brief contractor conducting site visits. Participating states will brief site teams on water systems that they will visit. Some states may choose to attend the site visit with the contractor.

A.5 INFORMATION COLLECTED: AGENCY ACTIVITIES, COLLECTION METHODOLOGY, AND INFORMATION MANAGEMENT

A.5.a Agency Activities

A.5.a.i EPA and Contractor Activities¹⁰

Many of the EPA activities described here will be conducted by contractors with EPA oversight/technical direction. For example, EPA will oversee contractor development of the data collection methodology, and collection and analysis of assessment data. For purposes of describing Agency activities related to the 2007 DWINSA, contractor effort is not distinguished from EPA effort. Separate estimates for contractor burden and cost will be provided in Section A.6.c. In addition, Section B.1.c describes the contractor's role.

Up-Front Activities

The following pre-assessment activities will be conducted:

- Revise the data collection instrument/site visit protocol. EPA is revising data collection instrument for the large and medium system surveys, and a site visit protocol for the small CWSs, based on lessons learned during the 1999 and 2003 DWINSAs. This task will include developing cover letters and as well as other materials for state use.
- Train state participants. To ensure that participating state officials understand every aspect of the 2007 DWINSA, EPA will conduct regional training sessions. The training will help ensure consistent responses across the country, high response rates, and efficient use of state staff.
- Select 2007 DWISNA respondents. The Agency will draw samples for the 2007 DWINSA.
- Develop data system. EPA will update the data system, developed for the 2003 DWINSA, to store and analyze data. The system will produce the necessary statistical reports for EPA, Congress, and states. The system will also allow state offices access to the data.
- Send data collection instruments. Data collection instruments will be sent via FedEx to the selected systems.
- Develop an electronic data collection instrument. The Agency will develop an electronic reporting form that systems serving more than 100,000 people can use to report their systems' projects. It is anticipated that this form will reduce the burden for systems completing the form, states reviewing the form, and the Agency performing data entry.

Data Collection Activities

EPA will conduct the following activities during the data collection phase of the 2007 DWINSA:

¹⁰ In states and territories where the EPA region acts as the primacy agency, EPA regional officers will perform the activities identified in this ICR as "State Activities." This should have no net impact on burden; burden for EPA regions acting as primacy agencies will be reflected in estimates for states rather than for the Agency.

- Provide technical assistance. The Agency will maintain a helpline primarily to provide technical assistance to large and medium systems (unless the state prefers to do so). The helpline will help ensure consistent responses across the country.
- Perform site visits. To ensure consistency, the site visits to small CWSs will be conducted by an EPA contractor (accompanied by state or regional personnel, if they wish to participate). As indicated above, site visit teams will complete the large and medium system survey instrument during the site visits.
- Review returned data collection instruments. EPA will review the completed data collection instruments to ensure that all data are documented and reasonable.
- Maintain the data. EPA will enter DWINSA data into the data system and perform quality control/quality assurance checks of data entry.

A.5.b Collection Methodology and Management

This section discusses the steps that EPA has taken to ensure that the information being collected will be accurate, reliable, and retrievable. This methodology was developed using experience gained in conducting the previous DWINSAs. EPA has incorporated into this methodology comments and advice from EPA staff involved with those assessments.

Development of Data Collection Instrument

Appendix B contains the data collection instrument. EPA has developed the 2007 DWINSA approach and the data collection instrument with the assistance of a workgroup. As is explained in Section A.3.c, the committee includes EPA headquarter, EPA regional, and state representatives. The 2007 DWINSA approach and data collection instrument will be refined through testing with water systems and consultations with outside groups. EPA's basic approach and many of the refinements to it were based on experience in conducting the 1995, 1999, and 2003 DWINSAs. In addition, EPA conducted a peer review of the 2007 DWINSA policies and statistical approach. In developing the 2007 DWINSA, EPA will select a set of appropriate assessment objectives that are easily answerable by knowledgeable respondents. EPA's pre-test of the data collection instrument is described in Section B.3.a. Section B.2.c.ii describes the steps taken to ensure that the data collection instrument will be an effective tool for retrieving the information EPA needs to meet the 2007 DWINSA objectives.

Methodology for Large and Medium CWSs (Systems serving more than 3,300 people)

Most systems serving more than 50,000 people have CIPs or similar documents that summarize their needs. These systems, therefore, are generally able to provide accurate information on their needs, and for some needs, accurate estimates on the associated cost. A data collection instrument will be sent to each of these systems. Clarifying information for completing the data collection instrument will be available from the state or EPA. Also, systems that participated in the 2003 DWINSA will be able to obtain a copy of their 2003 DWINSA data from the state and upload that data to the electronic form. After uploading the data, the system can update the projects (i.e., change parameters, delete projects that have been started, and add new projects identified since 2003).

The experience of states that participated in the previous DWINSAs indicates that systems serving 3,301 – 50,000 people can reliably complete a data collection instrument, if technical support is available. Most

of these systems could provide reliable data on the needs, and a large portion could provide cost estimates for meeting those needs. The states will provide technical support to these systems by answering their questions. EPA will provide classroom training to states that may want to collect the needs for systems through site visits. EPA will also offer a helpline for EPA regional, state, and system personnel.

EPA will send the data collection instrument to the systems. Respondents will send the completed data collection instruments to the state. The state will review all data and provide a quality control/quality assurance function. Next, the state will forward data to EPA for review and data entry. EPA performs a second quality control/quality assurance check to ensure that all data are documented and reasonable. EPA will enter the data (for systems that did not use the electronic reporting form). If the state chooses, it may verify that the data have been entered into the data system. Projects or cost estimates that are not documented will be identified in the data system as lacking documentation. If the system or state does not provide documentation, the project or cost estimate will be deleted from the 2007 DWINSA.

Methodology for Small CWSs (Systems serving 3,300 and fewer people)

Based on state experience with past needs surveys, the 1995 and 1999 DWINSA, and other experience in dealing with small CWSs, EPA know it is unlikely that small CWSs can reliably complete the data collection instrument. Therefore, EPA will collect data from these systems through site visits. To ensure consistency, the site visits will be conducted by an EPA contractor accompanied by state or EPA regional personnel, if they wish to participate. To help reduce costs, systems will be clustered together by county or clusters of counties. Unlike previous DWINSAs, the statistical pull for small systems will not require that all states have at least one cluster of small systems. In addition, at least 25 percent of the sample selected will be in counties with high levels of arsenic to ensure the collection of information regarding small system needs related to compliance with the new Arsenic Standard.

Methodology for American Indian and Alaska Native Water Systems and NPNCWSs

Estimates for the American Indian and Alaska Native water systems and NPNCWSs will be based on needs reported in the 1999 DWINSA and updated to 2007 dollars.

Data Quality

It is crucial that the results of the DWINSA be as uniform as possible across the country. Toward this end, EPA will take the following steps:

- EPA will establish a uniform set of assumptions or criteria for state, EPA regions, EPA headquarters, and contractor staff to evaluate data submitted by systems.
- EPA will provide training to all those involved in the DWINSA to ensure that the assumptions and procedures are clear and understood.
- EPA will provide quality assurance reviews of each data collection instrument submitted to ensure compliance with DWINSA policies and accuracy of data.

Among the most important steps in quality assurance is training. EPA will provide training sessions for state and EPA regional officials involved in the DWINSA. The regional training sessions will be designed to enable state staff to review completed data collection instruments and respond to questions from systems on the data collection instrument. The training will emphasize the following elements:

- Identifying the capital improvements associated with source, transmission, storage, treatment, and distribution.
- Estimating and reviewing construction cost estimates.
- Completing the 2007 DWINSA data collection instrument.

EPA will develop materials for distribution to state personnel who are unable to attend regional training sessions.

In addition to the training sessions, EPA will provide support for a helpline for state and water system personnel. It is anticipated that the helpline will be used primarily to provide information to the EPA regions and states and that the states will provide technical support to the systems. However, the helpline will be available to systems in states that have chosen not to provide their own technical assistance. Helpline staff will refer questions that raise a policy or technical issue to EPA staff.

Data quality will be assured by implementing the following mechanisms throughout the gathering and processing phases of the information collection:

- Adequate documentation. EPA has requested documentation of needs and costs, when available, to ensure the accuracy and reliability of the data. Acceptable forms of documentation of needs and costs are listed on List 4 of the Lists of Codes. EPA will not accept needs or costs without adequate documentation. EPA will make it very clear to respondents that they are not expected to develop cost estimates for the purposes of the 2007 DWINSA. The costs of projects without a cost estimate will be modeled.
- Provide inventory data on the total length of pipe in the water system. This information will only be required for water systems that submit projects for rehabilitation or replacement of pipe that are not independently documented (i.e., planning document, sanitary survey, or leak and break records). It is expected that not all systems will need to provide this information. This information is necessary to allow EPA to determine that the need reported is reasonable for a 20-year timeframe.
- Receipt control. The primary objective of the receipt control system will be to ensure that completed forms submitted by respondents (or forwarded by states) are logged in promptly and given proper chain of custody. A second objective is to provide states with the data needed to monitor cumulative receipts by date to identify potential problems with the response rate. Such response rate problems could necessitate action. See Section B.2.c.ii for EPA's method for improving the response rate. States that receive data collection instruments from respondents will be trained in receipt control.
- Data review by states. EPA will rely on the states to help ensure data quality. States that receive 1-percent of the DWSRF allocation formula were given the option to "opt-out" of participating in certain aspects of the DWISNA. Thirteen states have chosen not to participate in the statistical portion of the survey (i.e., collecting data from systems serving 3,301 – 100,000 people). However, all states that have systems serving more than 100,000 people will participate in the census portion of the survey. It is anticipated that the majority of states will support EPA in this effort. EPA will ask the Association of State Drinking Water Administrators (ASDWA) to communicate with the state drinking water administrators to encourage their participation. Staff from Region 9 will perform the state activities for systems in the three Pacific Islands, staff from

EPA Region 3 will act as the state for the District of Columbia, and staff from EPA Region 8 will act as the state for Wyoming.

EPA believes that state participation is essential in ensuring nationally consistent results because the states have more frequent communications with systems and possess a better understanding of each particular system's needs. Therefore, state personnel will have the opportunity to review the information on the form, as well as any accompanying documentation. When necessary, the states will contact the system to ask for clarifying information. In addition, EPA has improved states' ability to submit modifications through the Web site.

- Data entry. For data collection instruments submitted in hard copy form, the EPA contractor will screen the completed data collection instruments for legibility, completeness, and internal consistency, prior to entry into the DWINSA database. Reviewers will also assign comment codes to projects to describe any changes made to the data collection instruments. Data from the data collection instruments will be keyed into the database only after they have passed the initial screening. As data are keyed, an automatic data entry program will provide reasonable bounds checking and data verification. The program will signal the data entry operator, if an entry is out of the allowable range or is an invalid entry. For data collection instruments submitted electronically, the EPA contractor will follow the same review procedures as those submitted in hard copy.
- Data systems. EPA developed a web-based database system for the 2003 DWINSA that will be used for the 2007 DWINSA. The web-based system includes a data entry interface that allows the Agency and its contractors to input data and allows states to access, download, verify, and suggest modifications to their data (www.DWNeeds.com). EPA will use a commercial "off the shelf" program, Microsoft Access, to manage the information. The data system will provide the following functions:
 - Data entry through the user interface or batch upload.
 - Data verification through bounds checking.
 - A password-protected data modification interface.
 - Data access for states for review and verification of their data.
 - Predefined summary and statistical reports.
- Cost reasonableness ranges. EPA will develop "cost reasonableness ranges" to help verify the accuracy of the data and identify projects for further review.

Public Access to Data

The Agency's policy is to make the fullest possible disclosure of information without unjustifiable expense or unnecessary delay to the requester. Once the final Report to Congress has been submitted, the public will be given access to assessment data in accordance with EPA's policies and procedures for Freedom of Information Act (FOIA) requests. However, as a matter of policy, EPA will neither disclose the identity of any respondent to this questionnaire, nor the identity of any participating water system. EPA will develop standard report formats for providing data to the public.

A.5.c Small Entity Flexibility

In designing the 2007 DWINSA methodology, EPA has taken small systems' relatively limited technical capabilities and financial resources into account. EPA's experience with the previous DWINSAs has shown that small systems lack the resources and technical ability to complete the data collection instruments. Small CWSs regulated by the states will be included in the 2007 DWINSA and assessed by site visitors. EPA will use the 1999 DWINSA results as the primary basis for the 2007 DWINSA estimates for NPNCWSs, and Alaska Native and American Indian water systems.

EPA anticipates that many small and medium systems will not be able to provide information on all needs and capital costs. For projects without a documented cost, EPA will model a cost.

A.5.d Collection Schedule

The current schedule assumes EPA will receive OMB approval for data collection by November 1, 2006. EPA will send data collection instruments to drinking water systems on January 18, 2007. All systems participating in the 2007 DWINSA will be asked to complete and return the data collection instruments within 1 month of receipt.

To facilitate efficient data entry at EPA headquarters, EPA will ask the states to submit data for one-third of the systems within 3 months after data collection begins, or by April 18, 2007. Data for two-thirds of the systems will be due within 6 months (by July 18, 2007), and all data will be due on October 18, 2007. Exhibit A-5-1 summarizes the major collection milestones.

Exhibit A-5-1 Collection Schedule

Task	Date
Information Collection Request Submitted to OMB	August 2006
EPA Selects Systems to be Included in State Samples	August 2006
States Submit to EPA Contact Information to be Included on Return FedEx Labels	September 2006
Training Sessions for States and Regions	September - November 2006
Mail Out of Data Collection Instruments to Selected Systems	January 18, 2007
Deadline Given to Systems to Return the Data Collection Instrument to States	February 18, 2007
1/3 of Sent Data Collection Instruments Returned by States to EPA	April 18, 2007
2/3 of Sent Data Collection Instruments Returned by States to EPA	July 18, 2007
All Sent Data Collection Instruments Returned by States to EPA	October 18, 2007
No New Projects Will Be Accepted by EPA	October 18, 2007
No New Information on Submitted Projects Will Be Accepted by EPA	December 18, 2007
All Information in the Data System Finalized	February 18, 2008
Report to Congress Due	February 2, 2009

A.6 ESTIMATING THE BURDEN AND COST OF THE COLLECTION

A.6.a Respondent Burden

A.6.a.i Burden to Public Water Systems

The annual CWS burden for the 2007 DWINSA is estimated to be approximately 5,689 hours (approximately a total of 22,757 hours). Exhibit A-6-14 shows the breakdown of the annual burden hours for large, medium, and small CWSs. The bases for the burden estimates are detailed below.

Systems Serving More Than 50,000 People

The respondent burden for the systems serving more than 50,000 people consists of systems' burden for completing the data collection instrument. EPA estimates that the total unit burden is 10.39 hours per system. Exhibit A-6-1 summarizes the unit burden, broken down by activity and labor category.

- Participate in informational phone call. Each informational call should last about 15 minutes (0.25 hours). EPA anticipates that management staff will take the call at half of the systems and technical staff will take the call at the other half. Thus, the unit burdens are 0.125 hours for management staff and 0.125 hours for technical staff.
- Read cover letter/data collection instructions. EPA made the following assumptions in estimating the burden for reviewing the cover letter and data collection instructions:
 - A manager will receive the 2007 DWINSA and read the cover letter. The estimated time for managers to review the cover letter is 30 minutes (0.50 hours).
 - Technical staff will read the cover letter and data collection instructions. EPA estimates that the burden for this activity is 1 hour per system.

Thus, the total unit burden is 1.5 hours per system [(0.50) + (1.0)].

- Collect and copy supporting documentation. Time required to review system files, and collect and copy supporting documentation will vary greatly. EPA estimates that it will take 1.5 hours at 30 percent of the systems, 2.5 hours at 30 percent of the systems, 4 hours at 30 percent of the systems, and 16 hours at 10 percent of the systems. Thus, the average time per system is as follows:

$$(1.5 \times 0.30) + (2.5 \times 0.30) + (4 \times 0.30) + (16 \times 0.10) = 4.0 \text{ hrs/system}$$

- Call for technical assistance. Many systems will call states for technical assistance. In developing the burden estimate for this activity, EPA made the following assumptions:
 - The number of requests for assistance will equal 100 percent of the number of systems. (This estimate accounts for the fact that some systems will call more than once, while some will not call at all.)

- Each call will be placed by technical staff.
- About 50 percent of the questions will be “straightforward” and require a single phone call averaging 15 minutes (0.25 hours).
- About 50 percent of the questions will require the state to perform research and call the system back. In this case, EPA estimates that the total burden for the two calls is 26 minutes (0.43 hours).

Thus, the total unit burden is 20.5 minutes (0.34 hours) per system $[0.50(0.25) + 0.50(0.43)]$.

- Complete data collection instrument. EPA estimates that technical staff will take 3 hours to complete the data collection instrument. This estimate is consistent with EPA’s experience with the previous DWINSAs. Management is expected to take 18 minutes (0.30 hours) to review the completed data collection instrument for accuracy. Clerical staff is anticipated to take 1 hour to provide support to the technical and managerial staff. Thus, the total unit burden is approximately 4.30 hours per system.

Exhibit A-6-1 Estimated Unit Burden for Systems Serving More Than 50,000 People

Activity	Estimated Burden (hours)			
	Management	Technical	Clerical	Total
Participate in informational phone call	0.125	0.125		0.25
Read cover letter/data collection instructions	0.50	1.00		1.50
Collect and copy supporting documentation		2.00	2.00	4.00
Call for technical assistance		0.34		0.34
Complete data collection instrument	0.30	3.00	1.00	4.30
TOTAL	0.93	6.47	3.00	10.39

Systems Serving 3,301 – 50,000 People

Exhibit A-6-2 shows the unit burden for 1,538 systems serving 3,301 – 50,000 people. EPA estimates that each of these systems will take a total of 6.78 hours to respond to the 2007 DWINSA.

- Participate in informational phone call. Each informational call should last about 15 minutes (0.25 hour). EPA anticipates that management staff will take the call at half of the systems and technical staff will take the call at the other half. Thus, the unit burdens are 0.125 hours for management staff and 0.125 hour for technical staff.
- Read cover letter/data collection instructions. EPA used the following assumptions to estimate the burden for reviewing the cover letter and data collection instructions:

- A manager will receive the 2007 DWINSA and read the cover letter. The estimated time for managers to review the cover letter is 30 minutes (0.50 hours).
- Technical staff will read the cover letter and data collection instructions. EPA estimates that the burden for this activity is 1 hour per system.

Thus, the total unit burden is 1.5 hours per system [(0.50) + (1.0)].

- Collect and copy supporting documentation. Systems serving 3,301 – 50,000 people typically have less documentation than larger CWSs. However, the time required to review system files, and collect and copy supporting documentation will vary greatly. EPA estimate that it will take 1.0 hour at 50 percent of the systems, 2.0 hour at 25 percent of the systems, and 4.0 hours at 25 percent of the systems. Thus, the average time per system is as follows:

$$(1.0 \times 0.5) + (2.0 \times 0.25) + (4.0 \times 0.25) = 2.0 \text{ hrs/system}$$

- Call for technical assistance. Many systems will call EPA or the contractor for technical assistance. In developing the burden estimate for this activity, EPA made the following assumptions:
 - The number of requests for assistance will equal 150 percent of the number of systems. (This estimate accounts for the fact that some systems will call more than once.)
 - Each call will be placed by technical staff.
 - About 50 percent of the questions will be “straightforward” and require a single phone call averaging 15 minutes (0.25 hours) in duration.
 - About 50 percent of the questions will require the state to perform research and call the system back. In this case, EPA estimates that the total burden for the two calls will be 30 minutes (0.50 hours).

Thus, the total unit burden is 34 minutes (0.56 hours) per system [1.5(.50)(0.25) + .50(0.50)].

- Complete data collection instrument. EPA estimates that technical staff will take 2 hours to complete the data collection instrument. This estimate is consistent with EPA experience with the previous DWINSAs. Management is expected to take 28 minutes (0.47 hours) to review the completed data collection instrument for accuracy. Thus, the total unit burden is approximately 2.47 hours per system.

Exhibit A-6-2 Estimated Unit Burden for Systems Serving 3,301 – 50,000 People

Activity	Estimated Burden (hours)			
	Management	Technical	Clerical	Total
Participate in informational phone call	0.125	0.125		0.25
Read cover letter/data collection instructions	0.50	1.00		1.50
Collect supporting documentation		1.00	1.00	2.00
Call for technical assistance		0.56		0.56
Complete data collection instrument	0.47	2.00		2.47
TOTAL	1.10	4.69	1.00	6.78

Small Systems (Systems serving 3,300 and fewer people)

EPA will conduct site visits at 600 small CWSs. Because EPA is conducting site visits to these CWSs, the burden imposed on the systems is small. EPA estimates that the unit burden to small CWSs averages 3.25 hours per system. Exhibit A-6-3 summarizes the burden for each activity.

- Participate in informational/scheduling telephone call. The telephone call to discuss and schedule the site visit should take approximately 15 minutes (0.25 hours). Most small CWSs are staffed by one technical person; therefore, the entire burden falls with the technical labor category.
- Accompany survey team/answer questions. EPA does not expect that small CWS personnel will accompany the survey team for the entire site visit; however, EPA anticipates that system staff will make themselves available to answer very basic questions about the system configuration. EPA estimates that the burden to assist the survey team is 2 hours for half of the small CWSs selected and 4 hours for the remaining 300 systems. Thus, the average burden per system is as follows:

$$[(2 \times 0.5) + (4 \times 0.5)] = 3.0 \text{ hrs /system}$$

Exhibit A-6-3 Estimated Unit Burden for Small CWSs

Activity	Estimated Burden (hours)			
	Management	Technical	Clerical	Total
Participate in informational phone call		0.25		0.25
Accompany data collection team/answer questions		3.00		3.00
TOTAL	0	3.25	0	3.25

A.6.a.ii Burden to Primacy Agencies

Participating states will play an important role in conducting the DWINSA—they will help EPA ensure that the 2007 DWINSA is administered consistently nationwide. Most state activities will either involve using and reviewing data directly or facilitating EPA’s use and review of data. For example, states will review SDWIS inventory information for the statistical sample and verify that it is correct. They will help ensure a high response rate by telephoning systems serving more than 3,300 people before the 2007 DWINSA mailout and by making reminder calls to the systems that have not returned their data collection instruments by a specified date. States will help ensure data quality by answering systems’ questions on the data collection instrument and by reviewing completed data collection instruments and accompanying documentation for completeness and accuracy.

Given varying time and resource constraints, some states will wish to participate in the 2007 DWINSA more fully than others. The burden and cost estimates presented below represent a level of participation that EPA believes will ensure nationally consistent results. EPA encourages all states to participate at least at this level. The unit burden estimates are consistent with what was found to be true in the 2003 DWINSA.

The reader should note that the burden will vary widely by state, even for the same set of activities. A state’s actual burden depends on the number of drinking water systems in the state, the size and sophistication of those systems, the extent to which the state goes beyond the minimum requirements for the 2007 DWINSA, and other factors. Exhibit A-6-4 summarizes the burden estimates for each of the activity categories.

Exhibit A-6-4 Overall State Burden Summary

Activity Category	Estimated Burden
Up-Front Activities	110 hours, plus 0.2 hours/system
State Burden for Systems Serving More Than 50,000 People Assessment	6.58 hours per system
State Burden for Systems Serving 3,301 – 50,000 People Assessment	6.25 hours per system
State Burden for Small CWS Assessment	0.5 hours per system

Up-Front Burden

This activity category includes the state “fixed burden” for helping EPA prepare for the 2007 DWINSA. The total burden for these activities is 110 hours per state, plus 0.2 hour per system assessed. Exhibit A-6-5 summarizes this burden.

- Participate in training and other pre-mailout efforts. The burden for this activity is estimated at 80 hours per state and is not expected to depend on the number of systems in the state.
- Help EPA verify SDWIS data. Based on state experience EPA estimates that verifying SDWIS data for systems in the sample will require approximately 12 minutes (0.2 hours) per system.
- Perform miscellaneous administrative activities. The burden for these activities should be 30 hours per state.

Exhibit A-6-5 State Unit Burden for Up-Front Activities

Activity	Estimated Burden
Participate in training and other pre-mailout activities	80 hours/state
Help EPA verify SDWIS data	0.2 hours/system
Perform miscellaneous administrative activities	30 hours/state
TOTAL	110 hours/state, plus 0.2 hours/system

State Burden for Systems Serving More Than 50,000 People

This section estimates the state burden for helping EPA conduct the 2007 DWINSA in systems serving more than 50,000 people assessment by providing technical assistance where needed, calling systems that do not return the data collection instrument on time, reviewing the completed data collection instrument and documentation, and discussing the results with EPA. Although most of these systems will be able to answer the questions on the data collection instrument, states will provide them with clarifying

information as necessary. The state burden for activities associated with the systems serving more than 50,000 people is summarized in Exhibit A-6-6, which follows the activity descriptions.

- Telephone systems to ensure participation. EPA estimate that this preliminary phone call will take about 15 minutes (0.25 hours) per system.
- Provide technical assistance. In developing a burden estimate for this analysis, EPA made the following assumptions:
 - The number of requests for technical assistance will equal 100 percent of the number of systems. (This estimate accounts for the fact that some systems make such requests more than once.)
 - Of those that do require technical assistance, about 50 percent of their questions will be “straightforward,” requiring only 15 minutes (0.25 hours) to answer.
 - About 25 percent of their questions will entail limited research and follow-up, requiring 30 minutes (0.50 hours) to answer, including time to call EPA with questions.
 - About 25 percent of their questions will require the state to perform some research, and will require 1.0 hours to answer.

Therefore, the state burden for providing technical assistance is estimated at about 30 minutes (0.50 hours) per request $[0.5(0.25) + 0.25(0.50) + 0.25(1.0)]$. This is an average. Some states may choose to provide a much higher or lower level of technical assistance than anticipated by EPA.

- Call back systems that do not return the data collection instrument by a certain date. EPA assumes that the number of these “reminder” calls will equal 100 percent of the systems. This assumes that most (but not all) will need at least one reminder call and some will need two or three. The average time for these calls is 20 minutes (0.33 hours) per system. This does not include answering technical questions, which is accounted for above. Rather, it includes locating the correct contact person and obtaining a brief report on the status of the 2007 DWINSA response.
- Review completed data collection instruments and documentation. The data collection instrument will be returned directly to the state for review. For some systems, this documentation is expected to be quite voluminous, and reviewing it will be the most burdensome part of the 2007 DWINSA. The time required for this review is difficult to estimate. States that generate their own documentation for the 2007 DWINSA or add projects for distribution or transmission projects are required to ensure that the total pipe inventory section on the 2007 DWINSA is completed. Based on discussions with the states concerning their level of effort in previous assessments, EPA estimates that, on average, states will take 5.0 hours to review each submission. This estimate includes the time required to make follow-up phone calls and gather additional information as necessary.
- Discuss results with EPA. To estimate the state burden for resolving questions on completed data collection instruments, EPA made the following assumptions:

- EPA will have questions on 50 percent of the completed data collection instruments. Some of these questions will actually apply to all systems.
- Each question will take the state 1 hour to resolve.

Therefore, the burden per system is 0.5 times 1 hour, or 0.5 hours per system [0.50(1.0)].

Exhibit A-6-6 State Unit Burden for Systems Serving More Than 50,000 People

Activity	Estimated Burden (hours per system)
Call to ensure participation	0.25
Provide Technical Assistance	0.50
Call back systems that do not return the data collection instrument by a certain date	0.33
Review completed assessment forms and documentation	5.00
Discuss results with EPA	0.50
TOTAL	6.58

State Burden for Systems Serving 3,301 – 50,000 People

This section estimates the state burden for helping EPA conduct the systems serving 3,301 – 50,000 people for the 2007 DWINSA by telephoning systems to ensure participation, calling back systems that did not return the data collection instrument on time, reviewing the completed data collection instrument and the accompanying documentation, and discussing the results with EPA. The state burden for activities associated with the systems serving 3,301 – 50,000 people is summarized in Exhibit A-6-7, which follows the activity descriptions.

- Telephone systems to ensure participation. EPA estimates that this preliminary phone call will take about 15 minutes (0.25 hours) per system.
- Provide technical assistance. In developing a burden estimate for this analysis, EPA made the following assumptions:
 - The number of requests for technical assistance will equal 150 percent of the number of systems. (This estimate accounts for the fact that some systems make such requests more than once while some will not call at all.)
 - Of those that do require technical assistance, about 50 percent of their questions will be “straightforward,” requiring only 15 minutes (0.25 hours) to answer.

- About 25 percent of their questions will entail limited research and follow-up, requiring 30 minutes (0.50 hours) to answer, including time to call EPA with questions.
- About 25 percent of their questions will require the state to perform some research, and will require 1.0 hours to answer.

Therefore, the state burden is estimated at about 45 minutes (0.75 hours) per request $[1.5(0.5(0.25) + 0.25(0.50) + 0.25(1.0))]$. This is an average. Some states may choose to provide a much higher or lower level of technical assistance than anticipated by EPA.

- Call back systems that do not return the data collection instrument by a certain date. It is assumed that the number of these “reminder” calls will equal 100 percent of the systems. This assumes that most (but not all) will need at least one reminder call and some will need two or three. The average time for these calls is 30 minutes (0.50 hours) per system. This does not include answering technical questions, which is accounted for above. Rather, it includes locating the correct contact person and obtaining a brief report on the status of the 2007 DWINSA response.
- Review completed data collection instruments and documentation. The data collection instrument will be returned directly to the state for review. For states, this is the most burdensome part of the 2007 DWINSA, and the burden for this review is difficult to estimate. States that generate their own documentation for the 2007 DWINSA or add projects for distribution or transmission projects are required to ensure that the total pipe inventory section on the 2007 DWINSA is completed. For this ICR, EPA assumes that this activity takes states an average of 4.25 hours per system.
- Discuss results with EPA. To estimate the state burden for resolving questions on completed data collection instruments, EPA made the following assumptions:
 - EPA will have questions on 50 percent of the completed data collection instruments. Some of these questions will actually apply to all systems.
 - Each question will take the state 1 hour to resolve.

Therefore, the burden per system is 0.5 times 1 hour, or 0.50 hours per system.

Exhibit A-6-7 State Unit Burden for Systems Serving 3,301 – 50,000 People

Activity	Estimated Burden (hours per system)
Telephone systems to ensure participation	0.25
Provide Technical Assistance	0.75
Call back systems that do not return the data collection instrument by a certain date	0.50
Review completed data collection instruments and documentation	4.25
Discuss results with EPA	0.50
TOTAL	6.25

State Burden for Small CWSs (Systems serving 3,300 and fewer people)

This section estimates state burden for briefing the EPA contractor on systems that will be visited. The state burden for small CWSs is summarized in Exhibit A-6-8, which follows the activity descriptions.

- Brief contractor conducting site visits. States should take about 30 minutes (0.50 hours) per system to brief the contractor on individual systems. States may choose to accompany the site visitor on the site visit, but it is not required.

Exhibit A-6-8 State Unit Burden for Small Systems

Activity	Estimated Burden (hours per system)
Brief contractor conducting site visits	0.50
TOTAL	0.50

A.6.b Respondent Costs

A.6.b.i Costs to Public Water Systems

Exhibit A-6-9 summarizes the burden and costs to water systems. Total costs are estimated at \$701,206, which consists solely of labor costs. There are no operation and maintenance (O&M) costs or capital costs associated with the collection.

PWS labor costs are based on the number of burden hours times an average hourly wage rate, including overhead. The average hourly wage rate is the rate taken from a 2003 EPA document entitled Labor Costs for National Drinking Water Rules. The quoted rate was \$26.05 in 2003 dollars for small systems and systems serving 3,301 – 50,000 people and \$31.26 in 2003 dollars for systems serving more than 50,000 people. This rate has been inflated to 2005 dollars using the Employment Cost Index. The inflated rate is

\$28.24 for small systems and systems serving 3,301 – 50,000 people and \$33.88 for systems serving more than 50,000 people.

Exhibit A-6-9 Total Burden and Cost to Water Systems

Respondent	Unit Burden (hours)			Total Responses	Total Hours	Hourly Rate	Total Cost
	Management	Technical	Clerical				
Systems Serving More Than 50,000 People	0.93	6.47	3.00	999	10,380	\$33.88	\$351,661
Systems Serving 3,301 – 50,000 People	1.10	4.69	1.00	1,538	10,428	\$28.24	\$294,477
Small CWSs	0.00	3.25	0.00	600	1,950	\$28.24	\$55,068
TOTAL	2.02	14.40	4.00	3,137	22,757		\$701,206

Note: The average burden per response is 7.25 hours (22,757/3,137). The average burden for those systems completing a data collection instrument is 8.20 hours [(10,380+10,428)/(999+1,538)]. Numbers may not add due to rounding.

A.6.b.ii Cost to States

Exhibit A-6-10 shows the annual costs to states. As discussed above, all states have committed to help EPA administer the 2007 DWINSA with at least the minimum of activities.¹¹ Based on EPA's projection that all states will participate in the DWINSA, the cost to states is \$840,155. EPA will assess a total of 3,137 systems in the states. The labor costs are based on an average full time equivalent (FTE) cost of \$75,088 including overhead, which equates to approximately \$36.10 per hour.¹² This rate, which has been inflated to year 2005 dollars, is based on the rate (\$65,255) suggested by the workgroup that developed the State Workload Model in 2001¹³ and is consistent with the rates used in ICRs recently developed by DWPD.

There are no O&M or capital costs for states under this ICR.

¹¹ Region 9 will perform state functions for the three Pacific Island territories, Region 3 will perform these functions for the District of Columbia, and Region 8 will perform these functions for Wyoming. For simplicity, however, EPA is including the burden for these states in this ICR.

¹² According to the *ICR Handbook*, an employee is paid an average of 2,080 hours in 1 year.

¹³ The State Workload Model is a spreadsheet model used by states/primacy agencies to estimate resource needs for implementation of the SDWA.

Exhibit A-6-10 Total Burden and Cost to States

Activity	Number of states/ Systems*	Unit Burden	Total Burden	Hourly Rate	Total Cost
Up-front	56	110 hours/state	6,160	\$36.10	\$222,376
	3,137	0.20 hours/system	627	\$36.10	\$22,635
State burden for systems serving more than 50,000 people assessment	999	6.58 hours/system	6,573	\$36.10	\$237,285
State burden for systems serving 3,301 – 50,000 people assessment	1,538	6.25 hours/system	9,613	\$36.10	\$347,029
State burden for small CWS assessment	600	0.50 hours/system	300	\$36.10	\$10,830
TOTAL			23,273		\$840,155

A.6.c Agency Burden and Cost

The Agency burden and cost reflects the burden and cost directly incurred by EPA headquarters and EPA regions, summarized in Exhibit A-6-11. EPA will also bear the cost of contractor activities. Exhibit A-6-11 details the burden/cost of contractor activities. Both exhibits distribute burden/costs among Fiscal Years 2006, 2007, 2008, and 2009, reflecting that Agency and contractor activities will vary substantially over the 4-year period.

EPA made the following assumptions in developing its estimate of Agency and contractor burden and cost¹⁴:

EPA Headquarters

- Over the 4-year period, EPA Headquarters will expend a total of 2.8 FTEs (e.g., an average of 0.7 FTEs per year over the 4 years). Assuming 2,080 hours per year, this equates to 5,824 hours.
- The average salary and benefits (i.e., personnel compensation and benefits [PC&B]) of the FTEs is at the GS 13, Step 5 level of \$140,262. Assuming 2,080 hours per year, this equates to \$67.43 per hour.

EPA Regional Offices

- Over the 4-year period, EPA Regions will expend a total of 1.4 FTEs (i.e., an average of 0.35 FTE per year). Assuming 2,080 hours per year, this equates to 2,912 hours.

¹⁴ Hourly rates are from U. S. Office of Personnel Management, 2006 General Schedule (GS) Locality Pay Tables and overhead rates are from *Information Collection Request for Public Water Supply Program*, December 20, 1993.

- The average salary and benefits (i.e., PC&B) of the 1.4 FTEs is at the GS 11, Step 5 level of \$99,862. Assuming 2,080 hours per year, this equates to \$48.01 per hour.

EPA Contractor(s)

- Over the 4-year period, the EPA contractor(s) will expend a total of 49,780 hours of direct labor.
- The EPA contractor(s) will provide this professional labor at a total hourly rate, including all applicable indirect costs, of \$68.00.

Based on these assumptions, EPA estimates that the total burden/cost to EPA of the DWINSA over the 4-year period is 58,516 hours and \$3,917,558. Exhibits A-6-11 and A-6-12 follow, providing greater detail.

Exhibit A-6-11 Burden/Cost to EPA (Excluding Contractor Activities)

Fiscal Year	EPA Headquarters		EPA Regions		Total	
	Hours	Cost	Hours	Cost	Total Hours	Total Cost
FY 2006	892	\$60,148	446	\$21,412	1,338	\$81,560
FY 2007	2,020	\$136,209	1,010	\$48,490	3,030	\$184,699
FY 2008	2,020	\$136,209	1,010	\$48,490	3,030	\$184,699
FY 2009	892	\$60,148	446	\$21,412	1,338	\$81,560
TOTAL	5,824	\$392,714	2,912	\$139,804	8,736	\$532,518

Exhibit A-6-12 Burden/Cost of Contractor Activities

Activities	FY 06		FY 07		FY 08		FY 09		Total	
	Hours	Cost	Hours	Cost	Hours	Cost	Hours	Cost	Hours	Cost
Planning	2,000	\$136,000	5,159	\$350,812	3,200	\$217,600	300	\$20,400	10,659	\$724,812
Assessment Design	200	\$13,600	0	\$0	0	\$0	0	\$0	200	\$13,600
Peer Review	280	\$19,040	0	\$0	0	\$0	0	\$0	280	\$19,040
Modeling	334	\$22,712	982	\$66,776	960	\$65,280	960	\$65,280	3,236	\$220,048
Data Base Development	862	\$58,616	1,000	\$68,000	1,000	\$68,000	1,000	\$68,000	3,862	\$262,616
Assessment Production	0	\$0	945	\$64,260	0	\$0	0	\$0	945	\$64,260
Site Visits	0	\$0	12,000	\$816,000	0	\$0	0	\$0	12,000	\$816,000
Data Analysis	0	\$0	9,629	\$654,772	2,533	\$172,244	0	\$0	12,162	\$827,016
Report Writing	0	\$0	208	\$14,144	1,459	\$99,212	700	\$47,600	2,367	\$160,956
Statistical Analysis	0	\$0	94	\$6,392	656	\$44,608	750	\$51,000	1,500	\$102,000
Tech Assistance	0	\$0	960	\$65,280	200	\$13,600	200	\$13,600	1,360	\$92,480
Training	389	\$26,452	820	\$55,760	0	\$0	0	\$0	1,209	\$82,212
TOTAL	4,065	\$276,420	31,797	\$2,162,196	10,008	\$680,544	3,910	\$265,880	49,780	\$3,385,040

A.6.d Estimating Respondent Universe and Total Burden and Costs

Respondents for this ICR include CWSs and states. This ICR estimates that the number of CWS respondents is 3,137. In addition to the CWS respondents, this ICR assumes 56 states (50 states plus the District of Columbia and U.S. Territories).¹⁵ Therefore, the total number of respondents is 3,193. The total costs and burden for these respondents are detailed in Exhibits A-6-9 and A-6-10.

A.6.e Bottom Line Burden Hours and Costs

Exhibit A-6-13 summarizes the bottom line burden hours and costs for CWSs and states for this collection. The total burden for CWSs and states is 46,030 hours at a cost of \$1,541,361.

Exhibit A-6-13 Bottom Line Respondent Burden

Respondent Type	Burden Hours	Total Cost
Community Water Systems	22,757	\$701,206
States	23,273	\$840,155
TOTAL	46,030	\$1,541,361

Exhibit A-6-14 summarizes the burden hours and costs for CWSs and states per year. It is estimated that the CWSs will complete the data collection instrument in 2007. It is estimated that states will incur half of the burden associated with the 2007 DWINSA in 2007. The remaining half is assumed to be incurred evenly in 2006 and 2008.

Exhibit A-6-14 Burden Hours and Costs for Respondents per Year

Respondent Type	Total Hour Burden (per year)				Total Cost (per year)			
	2006	2007	2008	2009	2006	2007	2008	2009
CWSs	0	22,757	0	0	\$0	\$701,206	\$0	\$0
States	5,818	11,637	5,818	0	\$210,039	\$420,077	\$210,039	\$0
TOTAL	5,818	34,394	5,818	0	\$210,039	\$1,121,283	\$210,039	\$0
Average per Respondent	1.82	10.77	1.82		\$66	\$351	\$66	

¹⁵ For several of these entities, primacy activities are actually implemented by EPA Regional offices. However, as a simplifying assumption, they are included with the states for respondent calculations under this ICR.

Exhibit A-6-15 summarizes the bottom line burden hours and costs for EPA for this collection. The total burden for EPA (including its contractor) is 58,516 hours at a cost of \$3,917,558.

Exhibit A-6-15 Bottom Line Burden Hours and Costs for EPA (including its contractor)

Respondent Type	Burden Hours	Total Costs
EPA	8,736	\$532,518
Contractor	49,780	\$3,385,040
TOTAL	58,516	\$3,917,558

Exhibit A-6-16 shows the bottom line hour and dollar burden estimate by the Information Collection (IC) Entities. IC Entities covered by this ICR include, publicly-owned CWSs, privately/investor owned CWSs, state owned CWSs, and state government agencies.

Exhibit A-6-16 Disaggregated Burden by Affected Information Collection (IC) Entities

		TOTAL NUMBER OF IC ENTITIES	BURDEN PER RESPONSE	TOTAL HOUR BURDEN	HOURLY RATE	TOTAL COST	COST PER RESPONSE
Water Systems Respondents							
Serving More than 50,000 People	Publicly Owned CWSs	837	10.39	9,008	\$ 33.88	\$305,195	\$ 352
	Private/Investor Owned CWSs	123		1,278		\$43,298	
	State Owned CWSs	9		94		\$3,168	
	<i>Subtotal</i>	999		10,380		\$351,661	
Serving 3,301 to 50,000 People	Publicly Owned CWSs	1,368	6.78	9,275	\$ 28.24	\$261,927	\$191
	Private/Investor Owned CWSs	145		983		\$27,763	
	State Owned CWSs	25		170		\$4,787	
	<i>Subtotal</i>	1,538		10,428		\$294,477	
Serving 3,300 and Fewer People	Publicly Owned CWSs	285	3.25	926	\$ 28.24	\$26,157	\$92
	Private/Investor Owned CWSs	305		991		\$27,993	
	State Owned CWSs	10		33		\$918	
	<i>Subtotal</i>	600		1,950		\$55,068	
State Government Respondents		56	415.59	23,273	\$ 36.10	\$840,155	\$15,003
Total Water System Respondents							
	Publicly Owned CWSs	2,520	7.62	19,209		\$593,280	\$235
	Private/Investor Owned CWSs	573	5.68	3,252		\$99,053	\$173
	State Owned CWSs	44	6.72	296		\$8,873	\$202
	<i>Total CWSs Subtotal</i>	3,137	7.25	22,757		\$701,206	\$224
Total Respondents		3,193	14.42	46,030		\$1,541,361	\$483

A.6.f Reasons for Change in Burden

This ICR does not modify an existing ICR.

A.6.g Burden Statement

The public reporting burden for collections included in this ICR is detailed above. The total public reporting burden over the 4-year length of the 2007 DWINSA is estimated to be 46,030 hours, of which 22,757 hours are attributable to CWSs and 23,273 hours to states. These estimates include time for gathering information as well as developing and maintaining records. Public reporting burden for this collection of information is estimated to average 14.42 hours per response.¹⁶ Respondent burden for completing the data collection instrument is 8.20 hours.¹⁷

Burden means the total time, effort, or financial resources expended by people to generate, maintain, retain, disclose, or provide information to or for a Federal agency. This includes the time needed to review instructions, adjust the existing ways to comply with any previously applicable instructions and requirements, train personnel to respond to the information collection request, search data sources, complete and review the collection of information, and transmit or otherwise disclose the information. An agency may not conduct or sponsor, and a person is not required to respond to, a request for information collection unless it displays a currently valid OMB control number. The OMB control numbers for EPA's regulations are listed in 40 CFR Part 9 and 48 CFR Chapter 15.

To comment on the Agency's need for this information, the accuracy of the provided burden estimates, and any suggested methods for minimizing respondent burden, including the use of automated collection techniques, EPA has established a public docket for this ICR under Docket ID No. EPA-HQ-OW-2006-0486 which is available for online viewing at www.regulations.gov, or in person at the public viewing at the Water Docket in the EPA Docket Center (EPA/DC), EPA West, Room B102, 1301 Constitution Ave., NW, Washington, DC. The EPA Docket Center Public Reading Room is open from 8:30 a.m. to 4:30 p.m., Monday through Friday, excluding legal holidays. The telephone number for the Reading Room is (202) 566-1744, and the telephone number for the Water Docket is (202) 566-2426. An electronic version of the public docket is available at www.regulations.gov. This site can be used to submit or view public comments, access the index listing of the contents of the public docket, and to access those documents in the public docket that are available electronically. Also, you can send comments to the Office of Information and Regulatory Affairs, Office of Management and Budget, 725 17th Street, NW, Washington, DC 20503, Attention: Desk Office for EPA. Please include the EPA Docket ID No. EPA-HQ-OW-2006-0486 in any correspondence.

¹⁶ For this ICR, the number of responses is calculated at 3,193 (999 large systems, 1,538 medium systems, 600 small systems, and 56 states and U.S. Territories). The burden per response is calculated as the total respondent burden (46,030) divided by the number of responses (3,193).

¹⁷ See Exhibit A-6-9.

This page intentionally left blank.

PART B OF THE SUPPORTING STATEMENT (FOR STATISTICAL SURVEYS)

INTRODUCTION TO PART B

The Environmental Protection Agency (EPA) proposes to conduct the following type of statistical survey for the 2007 Drinking Water Infrastructure Needs Survey and Assessment (DWINSA). EPA proposes a mail assessment of community water systems (CWSs) serving populations of more than 3,300. EPA will send site visitors to collect data from CWSs serving 3,300 or fewer people. EPA is proposing the same methodology for collecting data for CWSs serving more than 3,300 people, as was used in the 2003 DWINSA. EPA is also proposing the same approach used in 1999 to collect data from CWSs serving 3,300 or fewer people.

This page intentionally left blank.

B.1 SURVEY OBJECTIVES, KEY VARIABLES AND OTHER PRELIMINARIES

B.1.a Survey Objectives

The primary objective of the 2007 DWINSA is to collect information from CWSs on the infrastructure they need 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 need. EPA has established policies to ensure that the overarching goals of the survey are met. These policies are provided to the states and help EPA:

- Estimate the total national 20-year need.
- Estimate the total 20-year need for each participating state.
- Provide complete and accurate data to Congress.
- Provide a tool to fairly distribute DWSRF capitalization funds to states.
- Maintain the credibility of the DWINSA findings.

EPA proposes to collect information on the cost of systems' infrastructure needs; if cost data are not available from systems, 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" includes raw water source, transmission, source water treatment, storage, distribution, pumping stations, and other needs.

EPA will use the information from the DWINSA to project capital investment requirements of drinking water systems. The information will be used to allot Drinking Water State Revolving Fund (DWSRF) monies among states and as part of an allotment formula for the American Indian and Alaska Native DWSRF set-aside program.

EPA is proposing the same methodology as used in previous DWINSAs. Two changes were made for the 2007 DWINSA from the approach used in 2003. Data will be collected from systems serving 3,300 or fewer people. Data was not collected from these systems in the 2003 DWINSA. Data will be collected from a random sample of systems serving 50,001 – 100,000 people. These systems were selected with certainty in 2003. The sampling design will be discussed in detail below.

B.1.b Key Variables

Several key variables are available from the Safe Drinking Water Information System (SDWIS). To ensure accuracy, the 2007 DWINSA will verify these data by asking respondents to confirm existing information (pre-printed on the data collection instrument), or correct it. These variables include population served, total design capacity, number of service connections, primary source of supply, ownership type (private or public), and whether the system purchases water from, or sells water to, another public water system (PWS).

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; documentation of need and cost (if necessary); if the project is a new project or rehabilitation of 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; cost of the project; and date of the cost estimate. For most of these variables, respondents will choose the appropriate “documentation,” “type,” or “regulation or requirement” from a lists of codes.

The principal variable of interest is total projected capital needed for each CWS in the 2007 DWINSA for the time period 2007 – 2026. 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 EPA use to allocate 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) in this manner.

Information on type of need will be used to disaggregate total capital needs for EPA’s Report to Congress. Information on documentation of 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 will be described in greater detail later in this document.

B.1.c Statistical Approach

The 2007 DWINSA is being designed to achieve a desired level of precision for state-level estimates of total capital needs for medium and large CWSs. It also is being designed to estimate the total capital needs of small systems for the nation as a whole. EPA proposes a survey of a statistical sample to estimate total capital needs for CWSs serving populations of more than 3,300. This statistical approach minimizes burden while achieving the desired level of precision.

Medium and Large CWSs

The 2007 DWINSA design divides CWSs serving populations of more than 3,300 into two groups: large CWSs (serving populations of more than 50,000), and medium-sized systems (serving populations of 3,301 – 50,000). EPA proposes to sample with certainty systems serving more than 100,000 people. These systems have the largest capital needs, and they have the staff to respond efficiently to the 2007 DWINSA. EPA proposes to draw a random sample of medium systems (serving 3,301 – 100,000 people). This methodology can reduce burden and still achieve the DWINSA data quality objectives. To meet the state-level precision targets, EPA will first determine the total sample size for each state to meet the target level of precision. EPA will then allocate the sample to strata in order to maximize the efficiency of their design.

Small CWSs

The objective of the 2007 DWINSA is to develop state-level estimates of total capital needs for CWSs. For large and medium systems, as explained above, this objective is achieved by selecting samples that are allocated across various strata in the population to achieve an overall precision level for each state. Several barriers prevent us from developing state-level estimates for systems serving populations 3,300 and fewer:

- First, a mail survey is not an effective approach to collection of data from these small CWSs. State experience with mail surveys for small CWSs suggests that total non-response and item non-response would be very high with a mail survey. Also, states believe that the absence of knowledgeable respondents at small CWSs limits the general reliability of the responses. Therefore, the 2007 DWINSA workgroup recommended, and EPA agrees, that the best way to gather information from small CWSs is through site visits made by EPA contractors. This will minimize total non-response, eliminate item non-response, and significantly improve the reliability of data collected.
- Second, if EPA assumes that all data collected from small CWSs will require site visits, then the number of such visits is constrained by the budget allocated for the 2007 DWINSA. EPA's current budget provides for approximately 600 site visits. Including small CWSs in the state-level design proposed for the medium and large systems, however, would require approximately 22,000 site visits. Thus, the statistical design for medium and large systems cannot be applied to small CWSs.
- Given this dilemma, the EPA workgroup committee recommended that EPA adopt a different approach for small CWSs, one that focused on national-level estimates. The direct sample estimates of total capital needs at the national level will be used to infer the total capital needs for small CWSs in each state. The workgroup preferred this approach.

EPA is designing and conducting the 2007 DWINSA with the assistance of a contractor:

Contractor

The Cadmus Group, Inc.
57 Water Street
Watertown, MA 02472
(617) 673-7000

Contractor Roles

- **Technical oversight for all contractor activities**
- **Oversight of data collection instrument design and testing.**
- **Oversight of statistical sample design**
- **Training**
- **Mailings; logistics**
- **Technical support for respondents and states**
- **Model development**
- **Data processing**
- **Statistical sample design**

B.1.d Feasibility

The 2007 DWINSA data collection instrument has been designed with the capabilities of the typical respondent in mind. To fully assess feasibility, the Agency undertook the following steps. First, EPA convened a workgroup (see Section A.5.b) to comment on the proposed data collection and its feasibility. Second, EPA 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 a mail survey. The data collection instrument was pre-tested, as described in Section B.3.a.

The Agency recognizes that some medium CWSs (and a few large CWSs) may not have cost data or documentation of costs for some projects. In those cases, the 2007 DWINSA data collection instrument requests other readily-available information that EPA can use to model costs. EPA will make it very clear

to respondents that they are not expected to develop cost estimates for the purposes of the 2007 DWINSA. In addition, EPA (or states) will provide large and medium CWSs with a helpline to assist them complete the data collection instrument.

Unlike the medium and large systems, the DWINSA will not be self-administered by small CWSs; rather, EPA contractors, accompanied by state personnel if state personnel participate in this portion of the 2007 DWINSA, will visit the small CWSs. Prior to the visit, the contractors will have access to all state records on the system (e.g., the results of recent sanitary surveys and inspections). The contractors will spend approximately 3 hours with the system owner or operator, requesting information that will be helpful in estimating system infrastructure needs. The contractor will then conduct a physical inspection of the system to confirm information provided by the owner or operator.

The EPA contractor will focus attention on the capital needs associated with treatment of source water, transmission, storage, and distribution. Capital needs associated with treatment will be modeled using methods similar to those currently used by EPA in the development of economic analyses. (In these analyses, data on occurrence of contaminants and cost estimates for treatment of source water to remove contaminants yield the cost of compliance with regulations that require the removal of contaminants from finished water.)

Reliance on site visits to small CWSs was strongly recommended by the EPA workgroup to avoid problems that have faced every state survey of small CWS infrastructure needs:

- *Total non-response.* Since many systems have not clearly identified responsible parties, and since responsible parties often are unwilling to respond to data collection instruments, it is difficult to use a mail survey to obtain the necessary information. Working with participating state regulatory agencies and representatives of small CWSs should minimize non-response problems.
- *Item non-response.* System owners and operators often are not knowledgeable about the capital needs of their systems. Unlike larger systems, who may maintain CIPs, small CWSs lack information to answer questions. Since EPA contractor engineers will conduct site visits to gather data, item non-response should be eliminated.
- *Reliability.* State drinking water regulators are suspicious of information provided directly from owners or operators of small CWSs. Unlike larger systems, small CWSs usually do not have professional, certified operators. Instead, one is likely to meet mobile home park owners, volunteers from homeowners associations, and others who are not water supply professionals. State drinking water administrators clearly prefer the judgments of EPA contractor engineers, accompanied by their own staff, for reliable information on capital needs.

Finally, employing site visits will substantially reduce the burden on small CWSs. Total burden on the systems, on average, will be about 1 hour. Instead of completing a data collection instrument, the system owner or operator can answer questions asked by the visiting engineer. The approach was discussed with knowledgeable state drinking water regulators, as well as representatives of small CWSs, and all parties agreed that it was the best approach to achieve the desired results of the 2007 DWINSA.

Sufficient contract funds have been identified to complete the 2007 DWINSA.

The time frame for the 2007 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 of early 2009. 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, data collection instrument, pre-test, collection methods, and follow-up procedures.

Medium and Large CWSs

The sample design for the DWINSA is stratified random sampling within 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 – 10,000; 10,001 – 25,000; 25,001 – 50,000; 50,001 – 100,000; and populations of more than 100,000), and primary source of supply (surface and ground). Systems serving more than 100,000 people are selected with certainty. The size of each state's sample of systems serving populations of 3,301 – 100,000 is set to meet the DWINSA's data quality objectives.

EPA's precision target is to be 95 percent confident that the true need lies within an interval, the upper and lower bounds of which do not exceed 10 percent of the sample mean (or estimated need). Once the total size of the sample of systems serving more than 3,300 people has been determined for each state, the number of samples to be taken in each stratum within each state will be allocated in a manner that minimizes the variance of the estimated total capital costs. EPA will use a Neyman allocation to determine the number of systems to select from each stratum. The Neyman allocation is described in detail in Section B.2.b.ii.

Small CWSs

The 2007 DWINSA design for small CWSs, like that for medium and large systems, is stratified random sampling. The stratification variables for small CWSs are the same as those for other systems: size of population served and primary source of supply.

Unlike the medium and large systems, the design for small CWSs is driven by a significant budgetary constraints: EPA cannot afford to complete more than approximately 600 site visits. EPA's objective in sampling is to achieve the maximum level of precision on a national basis without exceeding that budgetary constraint. Precision targets will be discussed in Section B.2.c, below.

B.2.a Target Population and Coverage

The target population is CWSs in the nation. A CWS is a 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). The DWINSA is being designed to produce estimates of the capital need of medium and large systems for each state. It is being designed to produce estimates of the capital need of small systems for the nation as a whole.

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. The strata employed in the design are discussed in Section B.2.b.iii. Neyman allocation is used to efficiently allocate the sample of water systems among the strata.

B.2.b.i Sampling Frame

The sampling frame 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 Safe Drinking Water Act (SDWA) Amendments of 1996. The following information will be extracted from SDWIS for the statistical survey and verified by participating states:

- Name of system
- Contact person
- Address of system
- Population served
- Total design capacity
- Number of connections
- Primary source (surface water or ground water)
- PWS identification number (PWSID)
- Ownership type
- Consecutive system (i.e., does system purchase or sell water)

From these data, EPA will develop the frame from which EPA will (1) calculate summary statistics (e.g., number of systems per state in pre-defined strata) for use in calculating sample size, and (2) randomly choose systems within the design strata to take part in the 2007 DWINSA.

Justification for the Use of SDWIS

The following criteria are often used in assessing 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 medium and large system survey 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 2007 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, EPA has conducted audits of the quality of SDWIS data. As a result, EPA is aware of the problems with SDWIS. The audits, however, show that errors in classification of systems by strata proposed for the 2007 DWINSA are rare. The audits show that systems are misclassified by population or source in less than 1 percent of all cases.

To mitigate any potential problems with the sample frame, the 2007 DWINSA design anticipates substantial state involvement in the 2007 DWINSA process. Participating states, for example, will be checking the sample frame of systems that will be used to determine the final sample. In EPA's experience, states often have in-house data systems with very accurate data, particularly on medium- and large-sized CWSs. Even if these data are not transmitted to SDWIS, they are available to states and can be used by states to check the sample frame.

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 2007 DWINSA. As shown on this exhibit, the sampling design will be implemented to achieve state-level precision targets for medium and large CWSs. Precision targets are discussed in Section B.2.c.

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 targets for that state. The second step allocates the sample across the relevant strata in the state. The strata are described in section B.2.b.iii.

The first step in determining the sample size is calculating the total number of samples required at the state level to meet the precision requirements. The sample size is given by:

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

- Where:
- n_{0g} = the sample size (prior to the finite population correction)
 - N_{gh} = the total number of systems in the h^{th} stratum in the g^{th} state (taken from SDWIS)
 - S_{gh} = the standard deviation of the variable of interest for the h^{th} stratum in the g^{th} state (estimated using data from the 2003 DWINSA)
 - H = the number of strata defined in the sample design for the g^{th} state
 - V_g = the desired sampling variance for the total medium and large system 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 2003 DWINSA) and multiplying 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 (0.10 for the Assessment). Z_α is the value of a standard normal distribution for a confidence level of $1 - \alpha$, (1.96 for the Assessment).

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

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

The second step allocates the total sample to each of the strata EPA will randomly draw this number of samples from each of these strata. The Neyman allocation formula is used for the allocation:¹⁸

$$n_{gh} = n_g \left(\frac{N_{gh} S_{gh}}{\sum_{h=1}^H N_{gh} S_{gh}} \right)$$

(Because systems serving populations more than 100,000 are to be sampled with certainty, H is reduced by the number of large-system strata in the sample design for the large and medium systems.)

In order to implement these sample size and sample allocation equations, 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 2003 DWINSA.

Exhibit B-2-1 State Sample Sizes

STATE	TOTAL NUMBER OF MEDIUM AND LARGE SYSTEMS	ESTIMATED SAMPLE SIZE FOR MEDIUM AND LARGE SYSTEMS
Alaska	21	16
Alabama	354	137
Arkansas	174	82
American Samoa	1	1
Arizona	127	28
California	676	194
Colorado	159	39
Connecticut	60	46
District of Columbia	1	1
Delaware *	26	3
Florida	382	108
Georgia	224	57
Guam	3	3
Hawaii *	30	2
Iowa	135	40

¹⁸ 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.

STATE	TOTAL NUMBER OF MEDIUM AND LARGE SYSTEMS	ESTIMATED SAMPLE SIZE FOR MEDIUM AND LARGE SYSTEMS
Idaho *	45	1
Illinois	462	83
Indiana	213	100
Kansas	120	41
Kentucky	266	102
Louisiana	224	87
Massachusetts	247	57
Maryland	55	20
Maine *	35	1
Michigan	306	55
Minnesota	160	56
Missouri	207	82
Northern Mariana Islands	2	2
Mississippi	197	148
Montana *	34	1
North Carolina	272	47
North Dakota **	31	-
Nebraska	44	21
New Hampshire *	37	2
New Jersey	227	57
New Mexico *	60	1
Nevada	35	11
New York	355	56
Ohio	314	79
Oklahoma	160	53
Oregon	109	43
Pennsylvania	341	75
Puerto Rico	122	56
Rhode Island *	28	2
South Carolina	167	46
South Dakota *	42	1
Tennessee	282	114
Texas	968	109
Utah *	105	7
Virginia	164	45
Virgin Islands	3	3
Vermont	34	12

STATE	TOTAL NUMBER OF MEDIUM AND LARGE SYSTEMS	ESTIMATED SAMPLE SIZE FOR MEDIUM AND LARGE SYSTEMS
Washington	199	50
Wisconsin	177	51
West Virginia *	110	3
Wyoming **	27	-
Total	9,359	2,537

*Eleven states have chosen not to participate in the statistical portion of the survey (i.e., collecting data from systems serving 3,301 – 100,000 people). They will however participate in the census portion of the survey (i.e., collecting data from systems serving more than 100,000 people). The number in the “Estimated Sample Size for Medium and Large Systems” represents the total number of systems in the state that serve more than 100,000 people.

** Two states have chosen not to participate in the statistical portion of the survey (i.e., collecting data from systems serving 3,301 – 100,000 people). In addition, these states do not have any systems that serve more than 100,000 people.

Small CWSs

The total small system sample is set at 600 by available resources. EPA will allocate the sample among six strata to produce the most efficient estimate of small sample need, given this sample size. Section B.2.b.iii discusses the how the sample will be stratified. The sample for systems serving 3,300 or fewer people is allocated among source water and population-served strata using a Neyman allocation. Within each ground water stratum, the sample is divided proportionately between systems in states with and without a substantial occurrence of arsenic. This method was chosen because the data on the variance in system need by arsenic occurrence are not available.

B.2.b.iii Stratification Variables

The objective of stratification is to increase the efficiency of the sampling design (thereby reducing the number of samples required at any level of precision) by the creation of independent strata. 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) 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 strata variance is minimized). If the within-stratum variance is relatively small, then a precise estimate of the variable of interest can be obtained with a relatively small number of samples. Each of the strata estimates can be combined to obtain a precise estimate for the 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.

EPA’s drinking water programs have historically evaluated CWSs based on (1) size (number of persons served), and (2) primary source (ground water and surface water).¹⁹ Using total capital need information obtained from the 2003 DWINSA, EPA evaluated several classification schemes. This analysis showed that the stratification scheme selected for the 2007 DWINSA medium and large system sample (10 strata based on size and source) was reasonable. Some states may have a different number of strata; this accommodated using their data as it is currently organized. Varying strata will be permitted only when the 2007 DWINSA’s overall precision is not reduced. The proposed strata for medium and large systems are as follows:

¹⁹ For the purposes of the 2007 DWINSA, purchased surface water systems are included with ground water systems. This design yields lower within-stratum variance.

Size of Population Served	Source	Sample Methodologies
3,301 – 10,000	Ground	Random sample.
3,301 – 10,000	Surface	
10,001 – 25,000	Ground	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 group rather than two.
10,001 – 25,000	Surface	
25,001 – 50,000	Ground	
25,001 – 50,000	Surface	
50,001 – 100,000	Ground	Random sample
50,001 – 100,000	Surface	
More than 100,000	Ground	Sampled with certainty
More than 100,000	Surface	

EPA’s sample design for small CWSs is also stratified based on the size of the population served and the source water of the system. In addition to source and population served, at least 25 percent of the counties selected will be in counties with high levels of arsenic. This is to ensure the collection of data regarding necessary infrastructure needs for systems serving 3,300 or fewer people affected by the Arsenic Rule. In 2003, infrastructure costs related to the arsenic regulation were determined using the Economic Analysis for the final rule. The workgroup decided this information was outdated and new estimates are necessary.

The proposed strata are as follows:

Water Source	Population Served
Surface Water Systems	25 – 1,000
	1,001 – 3,300
Ground Water Systems	25 – 1,000
	1,001 – 3,300

B.2.b.iv Sampling Method

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

For systems serving 3,301 – 100,000 persons, all CWSs will be allocated to 10 strata, based on population served and primary source. The sample size for each stratum in each state will be determined by the sampling strategy outlined above. The sampling method will be an equal probability random sample within each stratum. Anticipating a level of non-response, EPA will over-sample to achieve the desired number of completed data collection instruments. Since the expected response rate for systems serving 3,301 – 100,000 persons is 90 percent, EPA will draw a sample of 2,857.

All CWSs serving populations of 3,300 or fewer will be allocated to six strata, based on population served, primary source, and arsenic occurrence. The sample size for each stratum will be determined by the sampling strategy outlined above. The sampling method will be a two-stage probability proportional

to size random sample within each stratum. Past response rates for these systems exceeded 90 percent. EPA will over sample to account for non-response, and will draw a sample of 600.²⁰

B.2.b.v Multi-Stage Sampling

To achieve the required precision, reduce the burden to small systems, and to keep costs down, a two-stage cluster sample will be used for systems serving fewer than 3,300 people. The use of a two-stage sample design will result in slightly reduced precision for the stratum-level estimates.

First-Stage Sample

All small CWSs will be assigned to a county (or county equivalent in jurisdictions that do not have counties). Data on all small CWSs will be sorted by county so that EPA can determine the number of systems, by strata, in each county. If a particular county does not contain the required number of systems (a minimum of 6 systems), it is grouped with an adjacent county; the combined county group is referred to as a county-cluster. The first-stage sample will be approximately 120 counties, selected with probability proportional to size, where size is a composite measure of the number of small systems in each county. This method ensures that counties with more CWSs serving 3,300 or fewer people have a greater probability of being selected.²¹

States will be given a SDWIS list of small CWSs in the county (or counties) selected in the first-stage sample for their jurisdictions, and EPA will ask states to verify that the systems on the list are active CWSs with populations of 3,300 and fewer and assigned to the appropriate county. If the number of systems in a county is large (e.g., 100 or more), EPA will select a sub-sample of the systems in that county to reduce the burden on the state. This review by the states will produce a clean sample frame for the second-stage sample.

Second-Stage Sample

In the second stage, a stratified random sample of five systems is drawn from each of the counties or county-clusters selected in the first stage of sampling.

B.2.c Precision Requirements

B.2.c.i Precision Targets

The sampling design for large and medium systems will be implemented at the state level. 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 10 percent of the total need for these systems. For example, if the total need for these systems in a state is estimated to be \$2 billion, EPA will be 95 percent confident that the actual total need is between \$1.8 billion and \$2.2 billion.

The size of the small system sample is driven by budget constraints, not precision targets. EPA estimates that the sample size of 600 will allow it to estimate the national capital need of these systems with a 95 percent confidence interval equal to plus or minus 15 percent of the national small systems need. This precision level will be less than the level for estimates developed for medium and large systems, but it

²⁰ For purposes of burden calculation, EPA assumes 100 percent response.

²¹ This method is based on Folsom, R.E., F.J Potter., and S.R. Williams, "Notes on a Composite Size Measure for Self-Weighting Samples in Multiple Domains," *American Statistical Association 1987 Proceedings of the Section on Survey Research Methods*, August, 1987, pp. 792-796.

will not materially reduce the overall precision for total cost estimates at the state level. Small CWS costs are a small portion of total system costs in each state. Thus, the lack of precision for these systems will not significantly reduce the overall precision of the state-level estimates.

B.2.c.ii Nonsampling Error

EPA has developed an assessment approach that will employ several quality assurance techniques to maximize response rates, response accuracy, and processing accuracy to minimize nonsampling error. A pre-test will supplement the experience of EPA and its contractor (The Cadmus Group, Inc.) in formulating a strategy to reduce non-sampling error.

Particular emphasis will be placed on maximizing response rates. Standard methods that have proved effective in other surveys of CWSs will be used, including the following:

- States will review the sample of systems to receive the mail data collection instrument and will ensure that the best person to receive the data collection instrument is determined in advance.
- EPA and the states will coordinate in the production of a cover letter for the 2007 DWINSA. EPA's opinion (shared by state drinking water administrators, trade associations, and PWSs) is that surveys on state letterhead will be better received than letters on EPA letterhead. Therefore, states can use state-level cover letters signed by a senior state official instead of the EPA letter.
- The data collection instrument design, content, and format have been reviewed by organizations representing CWSs. In addition, the data collection instrument design, content, and format were reviewed by states that participated in the 1995, 1999, and 2003 DWINSAs.
- The data collection instrument design, content, and format will be pre-tested to ensure that all questions are properly stated and can be answered by all systems in the mail survey.
- Items being asked are those that owners or operators of systems serving populations greater than 3,300 should know. EPA does not ask for items that require monitoring, research, or calculations on the part of the respondent.
- The data collection instrument design is limited to 12 pages. By limiting the information requested, EPA believe that the average respondent can complete the data collection instrument in approximately 4 hours.
- Toll-free phone numbers will be provided to help respondents with questions or problems. In addition, respondents will be encouraged to call state personnel who will be trained to answer questions.
- Pre-paid return envelopes will be provided to respondents to make returning the data collection instrument convenient.

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

- EPA expects complete coverage of the target population using SDWIS, supplemented by state agency review of all systems.
- Data will be 100 percent independently keyed and verified.

- The data collection instrument is pre-coded to improve accuracy by eliminating unnecessary processing steps.

Supplementing these standard methods, EPA proposes several unique steps to eliminate non-sampling error, which have been developed in concert with organizations representing the states and CWSs. These organizations believe that the 2007 DWINSA is important and that a high level of participation by all CWSs is essential to its success. Because of the substantial commitment being made by states and CWSs to the 2007 DWINSA, EPA believes that response rates will be higher than most surveys of similar respondents. To ensure success, states and organizations representing CWSs are taking the following steps.

- *Participation of the states.* Because the DWINSA will be used to allocate DWSRF funds to states, each state has a strong interest in achieving a high response rate. EPA believes that state participation will be a key factor in guaranteeing high response rates and low item non-response. State personnel who work with CWSs every day are in a strong position to encourage systems to complete the 2007 DWINSA form. These states have committed to assisting EPA in achieving a high response rate by participating in follow-up activities. The states also will be available for technical assistance for any system that has questions about the 2007 DWINSA. All states have already agreed to participate in the 2007 DWINSA.
- *Participation of Organizations Representing CWSs.* EPA anticipates public support of organizations representing CWSs. 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).

This support by the organizations representing the respondents for the 2007 DWINSA can be helpful in many ways to minimize non-sampling errors. For example,

- These associations are likely to agree to prepare a letter for each system in their membership, stressing the importance of the 2007 DWINSA of drinking water infrastructure needs. This letter, along with the letter from the states, should make systems more likely to respond.
 - In the past DWINSAs, the largest association representing CWSs serving populations greater than 3,300—AWWA— provide support of its national organization behind the DWINSA. To improve the response rate, the AWWA enlisted of the support of its state affiliates (called “Sections”) in telephone follow-up to encourage response. AWWA assisted in past DWINSAs to help achieve the overall response rate of 94 percent. EPA hopes to secure similar AWWA support for the 2007 DWINSA.
- *Communications Strategy.* EPA has developed a comprehensive communications strategy that will inform likely respondents of the need for their participation. This strategy includes articles in magazines, newsletters, and bulletins of all major organizations that represent (or communicate with) CWSs. This includes publications of all of the organizations mentioned above, plus the state and local affiliates of these organizations. The strategy is designed to develop widespread peer-group support for participation in the 2007 DWINSA.

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-printed on all data collection instruments. The respondent needs only to enter accurate information if any pre-printed information is not correct.

The 2007 DWINSA is based on matrices that request a list of capital projects that the system plans for the period 2007 through 2026. For each project listed, the system is asked to provide: type of need; documentation of need and cost (if necessary); if the project is for new infrastructure or rehabilitation of 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; cost of the project; and date of the cost estimate. For most of these variables, respondents will choose the appropriate “documentation,” “type of need,” or “regulation or requirement,” from the Lists of Codes. All matrices have 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.

This page intentionally left blank.

B.3 PRE-TESTS AND PILOT TEST

B.3.a Pre-tests

EPA conducted two pre-tests of the data collection instrument for the 2007 DWINSA. The 2007 DWINSA pre-tests were conducted by EPA's contractor, The Cadmus Group, Inc. The pre-tests gathered feedback on the effectiveness of the data collection instrument, highlighted imprecise, ambiguous, or redundant questions, and indicated where further inquiry is needed. A pre-test was held in both Maine (four participants) and Montana (three participants). These states were chosen because they are both "opt-out" states, and because most of their systems will not need to participate in the 2007 DWINSA. Also, the contractor conducting the pre-tests has offices in both these states and by conducting the pre-test in these states they were able to reduce costs. The names of the seven systems were provided to EPA by the state 2007 DWINSA contacts. Based on the comments received EPA made modifications to the data collection instrument.

B.3.b Pilot Test

To eliminate unnecessary burden on states and CWSs, it has been decided that no pilot test for the 2007 DWINSA will be conducted. A pilot test was conducted for the 1995 DWINSA and consisted of 60 CWSs from New York and Texas. The procedures for mailing the data collection instruments and collecting the data are the same as those used for the 1995, 1999, and 2003 DWINSAs. EPA believes these procedures are well tested and have proven to be successful; therefore, it is not necessary repeat this testing step.

This page intentionally left blank.

B.4 COLLECTION METHODS AND FOLLOW-UP

B.4.a Collection Method

The proposed collection method for medium and large systems is a mail survey. The study data collection instrument and lists of codes will be mailed to all systems in the sample. State drinking water agencies will begin follow-up if the mail 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 proposed collection method for small systems is to visit to each small system in the sample. An EPA contractor, accompanied by state personnel that choose to participate, will interview the owner or operator and fill in the data collection instrument for all costs except treatment costs. (Costs of treatment will be modeled, using methods similar to those used by the OGWDW for regulatory impact analyses for new regulations.)

B.4.b Survey Response and Follow-up

The target response rate (defined as the ratio of responses to eligible respondents) for the 2007 DWINSA is 90 percent. EPA realizes that this is an ambitious target, but EPA believes that there are special circumstances that warrant such a target. Also, overall response rates of 94, 97, and 96 percent were achieved in the 1995, 1999, and 2003 surveys, respectively. In the first three surveys, EPA conducted the following proposed activities to achieve that high response rate.

- *Support from the Respondent Population.* This is a national survey of infrastructure needs for drinking water systems. The medium and large systems, as well as all national organizations representing these systems, understand the importance of the DWINSA results. All national organizations have endorsed the DWINSA and have communicated the importance of a high response rate to their members. As discussed in Section B.2.c, organizations have provided access to their newsletters and magazines to publicize and endorse participation in the DWINSA. EPA will ask national organizations representing smaller CWSs (e.g., the National Rural Water Association (NRWA) and AWWA) to help communicate the importance of a high response rate to their members.
- *Follow-up by States and Respondent Peer Groups.* Since a majority of participating states have indicated their willingness to participate in follow-up activities, these procedures will be implemented by state personnel, most of whom are personally familiar with the respondents. Procedures that states will use include reminder letters and telephone follow-up. In states that elect not to participate in follow-up, the EPA contractor will conduct these activities. If the follow-up fails after three attempts (one reminder letter plus two telephone follow-ups), EPA will shift to a second approach: peer-group follow-up by members of a trade association, such as AWWA. Procedures to be used by the association include a reminder letter followed by telephone calls. Such involvement is likely to improve the 2007 DWINSA's response rate.
- *Recruitment by States and Respondent Peer Groups of Small Systems.* In participating states, scheduling of site visits will be conducted by state personnel, most of whom are personally familiar with the respondents. If state personnel cannot schedule a visit with a system in the sample, EPA will turn to respondent peer groups.

This page intentionally left blank.

B.5 ANALYZING AND REPORTING SURVEY RESULTS

B.5.a Data Preparation

State personnel will check all cost data and documentation to ensure that it is consistent with state and national standards. States will then send the completed and reviewed data collection instruments to EPA for a second round of review by EPA contractor staff.

Once data have been checked, the contractor will key and verify the data. Senior data entry staff will be used for the verification process to improve quality control. Editing will include automated logic and range checks and checks for missing data. Missing cost data will be modeled, using other information provided by the respondents on the data collection instrument. When modeling is insufficient, missing data will be imputed using the standard methods such as cell means and regression. The sample of CWSs will be weighted so that stratum estimates can be summed to prepare state-level estimates.

B.5.b Analysis

EPA will prepare a report that tabulates the results of the 2007 DWINSA and explains the precision of the state-level estimates of total capital needs. Examples of statistics that will be produced include:

- Eligible capital needs by state and by types of need.
- Total capital needs by state and by types of need.
- Total capital needs by domains within the total population, e.g., systems serving populations greater than 100,000.
- Mean and median statistics on total capital needs (by type of need) for systems of various sizes. (These data will be of particular interest to participating respondents who will receive a short summary of these statistics.)
- Standard errors calculated for key statistics.

The analysis will be similar to that done for previous DWINSAs.

B.5.c Reporting Results

The 2007 DWINSA results will be made available to EPA and the public through:

- A printed report that is submitted to Congress on drinking water infrastructure needs. This report will be distributed to all participants in the 2007 DWINSA and all interested offices at EPA.
- Micro-computer access to state data (each state can access only its own data).
- Micro-computer access to the entire database (EPA only).

A report containing all technical information (data collection instrument, sampling plan, response rates, and variances) will be prepared and distributed. Record layouts, codes, and complete file documentation will be developed for data users (both micro-computer and mainframe users).

Appendix A
Public Notice Required Prior to ICR Submission to OMB

This page intentionally left blank.

Appendix B
Data Collection Instrument and Lists of Codes

This page intentionally left blank.

Appendix C
Comments and Response to Comments Received on the First *Federal Register*
Notice

This page intentionally left blank.

In response to the first *Federal Register* document published June 5, 2006, EPA received one set of comments from the American Water Works Association (AWWA). The following is a summary of the comments and EPA's response.

AWWA suggested that EPA set objectives for the Needs Survey. More detail of the general objectives of the 2007 DWINSA has been added to the objectives discussion in the final ICR.

They also point out that although a precision target is set at 95% +/- 10%, this does not address accuracy of the data collected and ask that EPA report their accuracy. Although precision can be measured, the accuracy of each water system's response to the 2007 DWINSA is not quantifiable. EPA believes that the most accurate assessment of the drinking water infrastructure needs is to rely on the information and judgment of the owners and operators of water systems within the context of historic best engineering practices. EPA strives to further achieve the most accurate data possible by enlisting the assistance of the state coordinators to contact the system and work with them to ensure they list all required projects.

AWWA indicated that the policy that needs associated with recently promulgated rules not be reported on the questionnaires but be added to the national needs based on the regulatory impact analysis (RIA) will result in underestimating the need associated with the Long-Term 2 Enhanced Surface Water Treatment Rule (LT2ESWTR). EPA sought input on this issue from the state coordinators as well as a technical expert group comprised of AWWA and AMWA membership. The group and EPA agreed that although some large systems may be planning for investments associated with additional treatment for *Cryptosporidium*, the rule does not require source water monitoring to be completed for many years. As such, all systems cannot be expected to be aware of their infrastructure needs to meet the new ruling. However, systems will be aware of any need associated with covering or treating uncovered finished water reservoirs, and as this requirement is in the near term (April 2009), planning will likely have taken place by the time of the 2007 DWINSA. Therefore, there is a policy to allow systems to report needs associated with uncovered finished water reservoirs, but does not allow needs associated with additional treatment for *cryptosporidium*. The costs estimated in the EA for addressing uncovered reservoirs will, therefore, be separated out and not be added to the national needs; whereas the other EA cost estimates, including those for treatment works, will be added, in total, to the national base. The 2007 DWINSA's Report to Congress will note the approach and how it impacts the 2007 DWINSA results. We expect that by the time of the 2011 DWINSA, systems will have incorporated to a fuller extent their LT2ESWTR compliance needs into their infrastructure planning documents and individual projects for these types of needs can be accepted from surveyed systems. This approach worked out with the technical expert group was endorsed by the State and Regional Needs Survey Coordinators Workgroup and by an independent peer review of the 2007 DWINSA methodologies.

AWWA expressed concern that the original notice did not provide sufficient information for a respondent to address the burden estimate. They were also unable to assess the statistical viability of the assessment as the final samples were not included in Exhibit B-2-1. In the public review and comment process for ICRs, the initial review is only of the information provided in the *Federal Register* notice; primarily the intended purpose of the information collection effort and its estimates of burden. EPA provided a rough and incomplete draft of the actual ICR in the public docket only to provide the reviewer with some further background and context to the information provided in the actual Notice. In the next opportunity for public comment of this ICR, reviewers will be provided the full and complete ICR document for consideration..

AWWA was concern that the burden estimate was likely too low. EPA has reviewed the assumptions made regarding burden and believes they are sound. EPA's estimates of burden are based on the experience of three previous DWINSAs and the input of survey respondents and state coordinators during those efforts and after in reviews to determine the strengths and weaknesses of those DWINSAs. The burden estimate

was revised substantially from the estimate from the last DWINSA primarily taking into account greater efforts to ensure consistency across systems and states in determining long-term infrastructure rehabilitation and replacement needs.

AWWA indicated that it was not clear from the survey instrument that the agency is only concerned with needs associated with the SDWA compliance. In fact this is not the case. EPA is collecting data on capital improvement needs to continue to ensure adequate supplies of safe drinking water, a goal that includes, but is not limited to regulatory compliance. EPA does ask the respondent to identify whether a need is associated with a regulation. EPA will continue to make clear in the DWINSA's Report to Congress that it attempts to capture infrastructure needs for this broader goal.

AWWA also commented on several parts of the draft questionnaire (not a part of the original notice) that discuss system inventory. AWWA was concern that providing system inventory information would be too burdensome for many of the surveyed water system operators. It will be made further clear that in the final questionnaire that the inventory tables are simply provided as a helpful tool for the system to assess any asset investment needs that are not yet part of their formal planning documents.

AWWA also expressed concern that the Needs Survey should not be fielded at the same time as the Community Water System Survey (CWSS). Although EPA understands the concern with respect to the impact this may have on response rate, the decision to conduct these surveys concurrently was primarily to accommodate data collection for small systems. The survey design is based on EPA contractors collecting data in the field from 600 small systems. This was accomplished by combining the small system data collection efforts for both the 2007 DWINSA and the CWSS. Adequate cost savings were realized from this combination of efforts to allow EPA to collect data from 600 systems for both surveys. For systems serving 3300 to 100,000, the 2007 DWINSA and CWSS will attempt to mutually exclude selected systems from each other survey. Systems serving over 100,000 will be sampled by both surveys, but we believe these size systems are capable and interested in participating in both surveys.

AWWA was concerned that sampling only 400 small water systems would not be adequately representative of the large number of these sized systems. As noted above, the sampling frame has been expanded from 400 small systems to 600. It should also be noted that the 2007 DWINSA's objective for small systems is to obtain results statistically-significant only at the national level. For medium and larger systems, the objective is to obtain results statistically significant at the state as well as national levels. A relatively smaller number of small systems are needed to provide statistically-significant results only at the national level.