

U.S. Bureau of the Census  
Center for Economic Studies  
Research Data Centers  
Handbook for Researchers

January 2008



## Preface

Welcome to the Census Bureau's Research Data Centers, managed and operated by the Census Bureau's Center for Economic Studies (CES) and its RDC partners. We hope that your research project goes smoothly and is successful. To help accomplish this, we provide this manual so that you will know what to expect at each stage your RDC project.<sup>1</sup> You are responsible for knowing the rules, regulations, and policies explained in this manual.

*IMPORTANT NOTE: Your research project at a Census Research Data Center (RDC) affords you opportunities to carry out unique research arising from the ability to access and explore confidential micro records. But the opportunities come at a price. Your research will take place under a set of rules and limitations that will be considerably more constraining than those prevailing in typical research environments. An important purpose of this manual is to explain these constraints to you.*

*The constraints stem from the legal requirements to meet the conditions of your appointment as a researcher at the RDC -- particularly the requirement to maintain confidentiality of the underlying microdata files to which you have been granted access, and the requirement that your project benefit Census Bureau data programs.*

This manual is organized as follows.

- Chapter 1 explains the legal requirements to maintain confidentiality of the microdata at the RDCs and the legal basis for your obtaining access to the data. In some cases, these requirements come not only from the Census Bureau, but also from other agencies that provide data to the Census Bureau or sponsor surveys. These requirements underlie all the policies described in the rest of the manual.
- Chapter 2 explains the life cycle of an approved research project at CES and summarizes the policies and procedures you must adhere to at each stage – from your first contact with us until your final checkout, and beyond.
- Chapter 3 explains in detail the policies and procedures you must follow in obtaining release of the research output you wish to remove from the secure RDC laboratory. Most importantly, it lays out policies related to protecting the confidentiality of the data and outlines the procedures you must follow to obtain release of your research results.
- Several appendices provide supplemental material. Appendix A is a brief glossary of terms used in this manual. Appendix B provides a template for the Account and Data Request form you must use to request a computer account. Appendix C contains templates of the Request for Clearance of Research Output, which you must submit when requesting release of any research output from the RDC facility. Appendix D is a template of the User Provided Data form, which you must submit whenever you are requesting that data be uploaded to the RDC computer systems for use in your project. Appendix E provides summaries of and excerpts from the laws underlying the policies in this manual.

---

<sup>1</sup>. For purposes of this handbook, the term "Research Data Center (RDC)" includes all of the restricted-access facilities operated by the Census Bureau's Center for Economic Studies and its RDC partners, including the facility operated by CES at Census Bureau headquarters.

## Table of Contents

Preface.....	1
Chapter 1: Introduction - General Regulations and Policies .....	4
The Legal Requirement to Maintain Confidentiality of Data at an RDC .....	4
Conditions for Restricted Access to Microdata -- Your Special Sworn Status .....	5
Chapter 2: The Life Cycle of Research at the Research Data Centers .....	6
During Your Stay at the RDC .....	6
<b>Obtaining Access to the Secure RDC Office</b> .....	6
<b>Obtaining Access to RDC Computer Facilities</b> .....	6
<b>Useful Documents</b> .....	7
<b>Scheduling Lab Time</b> .....	7
<b>Annual Renewals: Special Sworn Status (SSS) and Data Security Training</b> .....	8
<b>Annual Project Progress Reports</b> .....	8
<b>Maintaining Physical, Computer, and Data Security</b> .....	8
<b>Maintaining Confidentiality and Obtaining Release of Research Output</b> .....	9
<b>Getting Support</b> .....	9
<b>Adding New Researchers to Project</b> .....	9
<b>Adding an Additional RDC site to the Project</b> .....	9
<b>Obtaining New Data or Carrying Out Different Research</b> .....	10
<b>Seminars and Conference Presentations</b> .....	10
<b>Producing CES Discussion Papers</b> .....	10
<b>Providing Papers and Research Reports</b> .....	10
<b>Requesting Extensions on Projects</b> .....	10
Exit Process .....	11
<b>Documenting Benefit to the Census Bureau</b> .....	11
<b>Documenting Your Research</b> .....	11
<b>Backing Up and Archiving Data and Programs</b> .....	11
<b>Returning Badge or Keys</b> .....	11
After Exit .....	11
<b>Returning for Follow-up Work</b> .....	11
<b>Exit Interview</b> .....	12
<b>Continue to Provide Papers</b> .....	12
Chapter 3: Data Access, Confidentiality, and Disclosure Analysis Policies and Procedures .....	13
General Policies on Maintaining Confidentiality of Data and Clearing Release of Research Output .....	13
Requesting and Tracking Clearance of Project Output .....	15
Clearing Research Results .....	16
Disclosure Analysis for Model Output .....	18
Disclosure Analysis for Tabular Output .....	19
Clearing Other Files – Programs and research notes .....	22
Decisions and Release of Output .....	23

---

Bibliography .....	24
Appendix A: Glossary .....	26
Appendix B: Clearance Request Memo.....	40
Appendix C: User-Provided Data Memo.....	44
Appendix D: Legal Information.....	45
Census Bureau Legislation – Title 13 U.S. Code .....	45
General Information.....	45
Protection of Confidential Information -- Sections 9 and 214.....	45
Authority for Temporary Staff (Special Sworn Status) – Section 23(c).....	46
Data Programs Authorized by Title 13 Chapter 5 .....	46
Other Title 13 Data Programs.....	48
Protecting Federal Tax Information (FTI) – Titles 26 and 18, U.S. Code.....	49

## Introduction - General Regulations and Policies

This chapter explains the legal requirements you must meet to maintain the confidentiality of microdata at the RDC, and the legal basis for your obtaining access to the data. You will see that in some cases the requirements are not only those of the Census Bureau, but also those of other agencies that provide data to the Census Bureau or sponsor surveys. These requirements underlie all the policies described in the rest of the manual.

### ***The Legal Requirement to Maintain Confidentiality of Data at an RDC***

The microdata underlying the Census Bureau's data programs are confidential under Title 13 of the U.S. Code (U.S.C.). These microdata were collected directly by the Census Bureau in various surveys and Censuses, and they include statistics that quantify the activities of individual business establishments, firms, households, or individuals. Disclosure of confidential information obtained from Census Bureau surveys is illegal under Title 13. Officially or otherwise, the Census Bureau may not release publicly any data that reveal the identity of individual respondents or information about them. The Bureau gives a pledge of confidentiality to the respondents, the purpose for which is concern for respondent privacy. In addition, there are practical reasons to ensure confidentiality: Publication (even inadvertent) of confidential microdata records would most likely reduce cooperation with the data collection programs.

To become a researcher at an RDC, the Census Bureau must grant you Special Sworn Status (SSS), as described below. Under your SSS, you are subject to the same legal penalties as regular Census Bureau employees for disclosure of confidential information. The penalties are a fine of up to \$250,000, imprisonment for up to five years, or both. Moreover, your obligation to maintain confidentiality of any microdata you have accessed continues for the rest of your life.

Data collected under Title 13 include virtually all Census Bureau microdata sets available at the RDCs. However, other agencies' laws may also be applicable, and in such cases you will be required to abide by the legal requirements and confidentiality procedures of *both* agencies. For example:

- Your project may be approved to use administrative data that another agency supplies to the Census Bureau under legal agreements.
  - A primary example is Federal Tax Information (FTI) the Internal Revenue Service provides to the Census Bureau for use in carrying out programs authorized under Chapter 5, Title 13 U.S.C. These programs are the Economic and Decennial Censuses and their related survey programs. The relevant laws are Titles 18 and 26 U.S.C. The IRS provides the Census Bureau with certain data. Appendix E provides a brochure that summarizes requirements that must be met in carrying out projects that use FTI. It also provides references to more detailed information. Note that *all* researchers at an RDC must undergo training in the proper use and protection of FTI, whether or not their projects use FTI, because in an RDC researchers might inadvertently see FTI.

- Your project may be approved to use confidential microdata collected by the Census Bureau under Title 13 for another agency. An example is the Current Population Survey, which is collected by the Census Bureau under Title 13 for the Bureau of Labor Statistics. Researchers may legally use these data for statistical and research purposes only, as is stated in Title 42, U.S.C., Section 3789g.

We will inform you of any additional requirements.

### **Conditions for Restricted Access to Microdata -- Your Special Sworn Status**

The Census Bureau has granted you Special Sworn Status (SSS) to carry out a specific research project that has been approved by the Census Bureau. As a researcher with SSS, you are given restricted access to the specific data required for your approved project. Under Title 13, these data may be used only for *statistical purposes*; that is, to generate an aggregate description of a group of persons or businesses. Data collected for statistical purposes may not be used to make decisions regarding a specific individual or business. For more information, see Appendix A on terminology. SSS is authorized by Title 13 U.S.C. 23 (c) "to assist the Bureau of the Census in performing the work authorized by this title." The Census Bureau may provide SSS to an individual in the following cases:

- When an individual has expertise or specialized knowledge that can contribute to the accomplishment of Census Bureau projects or activities or engages in a joint project with the Census Bureau;
- When an individual is employed by an agency/organization performing a service for the Census Bureau under contract or providing information to the Census Bureau for statistical purposes;
- When Federal law requires an individual to audit, inspect, or investigate Census Bureau activities.

*It is important that you understand that the Census Bureau's use of researchers with SSS is to further its agenda. The discretionary authority to swear in researchers with SSS rests with the Census Bureau. No individual has a "right" to appointment nor is the Census Bureau obligated to appoint a person who may believe the Census Bureau needs his or her expertise. In approving your research project, the Census Bureau has made the judgment that your activities can contribute to the success of Census Bureau data programs or lead to improvements in them. We have found that granting access to microdata by qualified researchers for specific projects results in mutual benefits to the Census Bureau, the researchers, and their scientific fields, and provides information useful to decision makers.*

## **Chapter 1: The Life Cycle of Research at the Research Data Centers**

This chapter provides a summary and overview of what you can expect in doing research at the Center for Economic Studies or the Research Data Centers. One of the purposes of the chapter is to give you an understanding of certain perhaps unaccustomed constraints on your activities, particularly if you are used to carrying out research in an academic environment. As explained in the previous chapter, these constraints arise mostly from the two legal requirements - to protect confidentiality of the data and to produce benefits to Census Bureau data programs.

### **During Your Stay at the RDC**

The following summarizes policies governing your residence at the RDC: renewing SSS and training, maintaining security, maintaining confidentiality and obtaining release of research output, obtaining new data or carrying out different research, progress reports, and giving seminars and conference presentations.

### **Obtaining Access to the Secure RDC Office**

The RDC office is tied to a security system that meets Census Bureau specifications (e.g., electronic card key, combination lock, alarm system). You will be given the proper procedures for entry (e.g., your own key code, card, lock combination, and/or door key). Your RDC Administrator will guide you through the security and access procedures particular to your specific RDC.

You will be given a federal Census Bureau badge that will give you access to the main RDC door. This is the property of the US Government. As stated on the back of the badge, “This credential is the property of the United States Government. Counterfeiting, altering, or misusing violates section 499, Title 18 of the U.S. Code.” If you lose your badge, immediately notify your RDC Administrator. Moreover, at the end of the project (or once you stop working on a project), you will be responsible for returning your badge.

### **Obtaining Access to RDC Computer Facilities**

The standard computing equipment at RDCs are X-Terminals (thin-client computing devices) that are used to communicate with secure servers housed in the Census Computer Center in Bowie, MD. Through the thin client device, you will have remote access — only from within the RDC secure facility — to the CES servers containing your project data and files. All data storage and data processing are performed on these servers. No Title 13 or Title 26 data files are physically on site at any RDC.

Thin Client Computing for the RDCs has the following characteristics:

- Data are stored in a format determined by CES to be best suited to support research activities (e.g., as SAS datasets).
- A limited range of statistical software is available for researcher usage.
- Access to the central servers is accomplished at the RDC via an x-terminal. All data process work is conducted on the central servers, and no actual data processing is performed at an RDC thin client machine.

- All thin-client machines are currently running RedHat Linux. No portable media devices (e.g., floppy, CD/DVD, or USB drives) are available.
- The RDC computer network is isolated from other networks, which means that within the laboratory you will have no access to email or the World Wide Web.
- You will be accountable for your computer use through the use of passwords and system logs. Please refer to the Thin Client User Guide for password guidelines.

Each researcher will get an individual project account on the server. By now, you should have already received an Account and Data Request Form from CES, signed it and sent it back to CES. Your RDC administrator will coordinate your account activation with the CES data staff. ***Please note that your project's duration will officially begin when the first person on your project logs into his/her project account for the first time.***

CES will issue you a Thin Client Users Guide to help you become oriented to the system.

## Useful Documents

You will receive a packet of materials, including

- A copy of this manual;
- Thin Client User Guide;
- An electronic template of the memo you must submit to provide for releases. (We call this the *clearance request form* or *clearance request memo*. Chapter 4 describes this document in detail and the process of obtaining clearance. (A paper copy is provided in Appendix C of this handbook)
- An electronic template of the User-Provided Data Memo which must accompany any outside data you would like to bring into the lab (data must be part of your approved project, or required to help you obtain the benefits promised in your approved project). Please note: many frequently used public-use datasets are already loaded onto the RDC servers under the /public directory. (A paper copy is provided in Appendix D of this handbook.)

Other useful documents available in the RDC may include data documentation, survey forms, software manuals, and/or SAS or Stata programming books. These materials will vary across RDCs. Ask your RDC Administrator what is available at your RDC. Some data documentation is available in the directories for the corresponding datasets on the RDC server. Many useful working papers and data documentation can also be found online at the CES website at [www.ces.census.gov](http://www.ces.census.gov).

## Scheduling Lab Time

At some RDCs, you may need to schedule your laboratory time. In your proposal, you will have specified the length of term and some pattern of use.

- If your lab use will be intermittent, you should schedule your visits well in advance. If you do not, it may cause problems in scheduling other researchers and could result in your not having access to a lab terminal at your desired times.
- It is especially important for you to be aware of when the term of your agreement will expire. You

can check on this date by looking at your project record in the CES online management system. Your RDC Administrator will also be able to provide you with this information.

*IMPORTANT NOTE: We will give priority to researchers who adhere to their schedules and schedule visits in advance.*

## **Annual Renewals: Special Sworn Status (SSS) and Data Security Training**

It is required that your SSS be renewed every year, and this renewal coincides with the badge expiration date on your Census badge. Once your SSS has been renewed, you will receive a new badge with a new expiration date. Please notify your RDC Administrator 30 days in advance that your badge is going to expire. You will not be able to enter the RDC without your new badge.

The following Awareness Training modules need to be renewed at the beginning of each fiscal year. These should be completed in late fall (October or November) or when notified by your RDC Administrator that the new fiscal year training is available. The CES online management system will also send reminders to renew each of the following training modules thirty days before they are due. Please remember to keep your e-mail address up to date in the management system so that you can receive these reminders and other important notices.

- Title 13 Awareness Training (web-based)
- Title 26 Awareness Training (web-based)
- Information Technology Security Awareness Training (web-based).

Awareness Training for Rules of Behavior for Information Technology Systems may also need to be renewed on an annual basis.

## **Annual Project Progress Reports**

You must submit annual progress reports on your project to your RDC administrator. The report should describe the progress that is made towards achieving the benefits to the Census Bureau proposed in your research project. This report should also include any new benefits that may have been realized as a result of your work on your RDC project. Your RDC Administrator will provide the template.

## **Maintaining Physical, Computer, and Data Security**

The RDC lab has both physical security and computer security procedures, and a related set of policies, that you must follow. The security and related policies stem from the need to meet the conditions under which you have been granted access to the data for your project and the need to protect the confidentiality of the data. The official CES policies on Office Security, Computer Security, Data Access, and Confidentiality are listed in the Account and Data Request Form that you signed and sent in to CES. A copy of this form is included in this Handbook as Appendix B.

The RDC office has a security system that meets Census Bureau specifications (e.g., electronic card key). Your RDC administrator will give you the proper procedures for entry. Your RDC Administrator should

also review with you the security policies specific to that RDC, including what to do and whom to call in an emergency, what to do if you accidentally set off an alarm, building evacuation routes and procedures, etc.

## **Maintaining Confidentiality and Obtaining Release of Research Output**

Chapter 3 details CES policies on maintaining confidentiality of the data. It also provides instructions for obtaining release of research output. CES policies regarding disclosure analysis are also listed in the Account and Data Request Form (Appendix B).

## **Getting Support**

You may have data or subject matter questions during your project. **Please begin by addressing questions to your RDC administrator.** The administrator may be able to answer your question or may put you in touch with a CES or other Census Bureau employee who can answer the questions. We believe that having contact with CES and Census Bureau experts can provide benefits both to your project and to the Census Bureau's data programs. However, we need to manage this process so that you obtain the support you need in a way that does not overburden the staff, and so that we can maintain a record of support given and of your suggestions for improvements to Census Bureau data programs.

Also, even though you will have access to the lab during non-business hours (including holidays), support during these times will be limited. This will be true during the winter as well – if the government is shut down, support will be limited.

## **Adding New Researchers to Project**

The Principal Investigator on a project can request addition of new researchers to the project. The Principal Investigator on the project must submit a written request to the RDC Administrator. New researchers must have accounts on the CES website and be added to the project record. The second step is for them to apply for Special Sworn Status and take all the necessary training. Your RDC administrator will coordinate this.

## **Adding an Additional RDC site to the Project**

If one of the researchers on your project lives or works closer to a different RDC than the rest of the project team (e.g., due to that person moving or due to having/adding a co-author in another town), then you can request adding an additional RDC lab as a secondary work site for your project. Please note that the project must show that there is a “substantial presence” at the secondary site, meaning that one of the project researchers will be accessing the project at the secondary site for a significant period of time. Secondary sites will **NOT** be approved for brief periods of access, such as a project researcher wanting to work at a secondary site for two weeks while on a vacation or research trip. The RDC Administrator at the project's primary site is responsible for administration of the project. **Moreover, fees may apply at each RDC.**

The PI on the project needs to send the request to the RDC administrator (RDCA) at the primary site in order to add a new site. The RDCA at the primary site will notify the RDCA at the new site. The Executive Directors at both the primary site and the new site need to approve of the addition (including

any potential fees). If the request is approved, the local RDCAs will make the necessary arrangements for researcher access.

### **Obtaining New Data or Carrying Out Different Research**

Under your Special Sworn Status you may carry out only your approved research project using the data specified in your approved research proposal and your project agreement. If you wish to conduct research using a different/additional Census data set, you must submit a new proposal to be sent through the regular review process. The same applies to a new line of research that is not covered in your approved research proposal. If you have a new idea, please discuss it with your RDC Administrator. The RDCA will work with CES to determine if the idea falls within the scope of your existing project. Note that CES will not release output that does not fit in the current scope of the project.

### **Seminars and Conference Presentations**

We encourage you to present your work in seminars and at conferences, including seminars in your local area. However, all output presented to persons without SSS status must undergo disclosure review and clearance. Any presentations or reports must contain the standard disclaimer, shown on page 12.

### **Producing CES Discussion Papers**

You are required to submit at least one research paper produced as part of the project for inclusion in the CES Discussion Paper series (or any other appropriate Census Bureau series.) *Inclusion in such series does not preclude inclusion in any other paper series* (e.g., university or NBER working paper series.) For submission procedures, contact your RDC administrator

### **Providing Papers and Research Reports**

You must provide copies to CES (via your RDC administrator) of all published research papers and reports resulting from your RDC research project. These are crucial parts of your project records.

**IMPORTANT NOTE:** You will be allowed to start a new project only when you have provided us with all available project reports and papers from any previous projects.

### **Requesting Extensions on Projects**

Extensions are strongly discouraged and are granted only in rare circumstances. CES considers these requests on a case-by-case basis. You must request a project extension at least 90 days in advance of your project's end date. Your extension request must contain a rationale that focuses on why the additional time is needed to provide the proposed benefits to the Census Bureau. In addition, you must submit a document that summarizes the benefits that have been realized so far in your project. See the next paragraph for further explanation.

## **Exit Process**

### **Documenting Benefit to the Census Bureau**

As explained previously, as a researcher with Special Sworn Status, you are carrying out a project that aids the Census Bureau in its mission. You described these benefits in your proposal and benefits statement (also called a Predominant Purpose Statement). Therefore, you will be required to submit a Post-Project Certification (PPC) summarizing the key findings pertinent to Census Bureau Chapter 5, Title 13 programs, including whether/how the accomplished benefits differed from those proposed and describing their potential application. Often this PPC is accompanied by a technical memorandum, which explains the benefits in detail. This technical memorandum may contain confidential information that will not be released to the public.

*IMPORTANT NOTE: We will not release your final research output before you have provided this document. Moreover, you will not be allowed to return as a researcher unless you submit it.*

### **Documenting Your Research**

You must provide documentation of data sets and programs for any output you use in a report, memo, or paper -- whether you remove them from the Bureau or not. For a variety of reasons, it is important that CES and perhaps other researchers (including yourself) be able to reproduce any results that appear in print. Therefore, for any tabulation or model output you take out, you must provide the program, the log (if the statistical software produces a log), and the output. We will ask you for this information as part of the process of clearing your research output for release.

Moreover, and very importantly, you should trace the evolution of your data sets from the initial data sets we provide you (and that you may bring to the project) to the data sets that give your research results and the statistical output generated. This trail should include program and data set names, and should be sufficiently detailed that CES staff members could reproduce your work. See Chapter 3 for more details.

### **Backing Up and Archiving Data and Programs**

The CES project (thin client) servers are backed up regularly. If you lose a file and wish to have it restored, please contact your RDC administrator.

Your research files will be archived at the end of your project. The current CES data archive police will keep your research files for seven years. However, any files in your home directory (/home) will not be archived. Be sure to move these to your project directory (/rdcprojects).

### **Returning Badge or Keys**

We will require you to return your Census badge and other issued keys, where applicable, when you exit or whenever you will not be on site for a significant amount of time. Any codes you were given for a door or for an alarm will be disabled at that time.

## **After Exit**

### **Returning for Follow-up Work**

Our policy is to allow researchers to return to the primary RDC without additional charge for reasonable

amounts of follow-up work in response to a journal referees' comments for a "revise & resubmit." The typical situation involves carrying out additional model estimations. As a guide, a "reasonable" amount of work:

- Uses no new data or research samples;
- Asks for release of a small amount of output that is very similar to output previously released, and otherwise does not cause disclosure problems;
- Is carried out in less than 3 weeks.

Note: Currently, to access the RDC lab for follow-up work after your project term expires, you must submit the following:

1. A letter requesting re-admission for follow-up work
2. The letter from the journal editor and the referees' comments
3. The Post-Project Certification (PPC) document
4. A statement of new benefits to the Bureau from the follow-up work

The PPC for the original project must be approved before the request for reactivation of the project will be reviewed by CES. In addition, projects accessing Federal Tax Information (i.e., economic data) will require IRS approval of both the PPC and the reactivation request.

*Please consult with your RDC administrator well in advance to determine the requirements.*

### **Exit Interview**

At the end of your project, you must go through an exit interview. At this time you will be reminded of your obligations, including the fact that the Oath of Confidentiality is a lifetime obligation.

Note: Currently, to carry out follow-up work (other than that required for a "revise & resubmit") for projects using Federal Tax Information after your project term expires, you must submit a new proposal under the standard proposal review process. *Please consult with your RDC administrator well in advance to determine the requirements.*

### **Continue to Provide Papers**

You must provide us copies of **all** published research papers and reports resulting from your RDC research project, even if it is long after your project ends. These are crucial parts your project records. As stated previously, we strongly encourage you to submit research papers for inclusion in the CES Discussion paper series; this does not preclude inclusion in any other working paper series.

## Chapter 2: Data Access, Confidentiality, and Disclosure Analysis Policies and Procedures

This chapter presents policies and procedures you must follow in protecting the confidentiality of Census Bureau data supplied for your project, and in clearing your research results for release.

NOTE: This chapter gives policies and procedures concerned with protecting confidentiality of data supplied to you by the Census Bureau under Title 13, U.S.C., and possibly also protected by Title 26 U.S.C. (Federal Tax Information provided by IRS). The chapter does not discuss protecting data supplied by the National Center for Health Statistics (NCHS) and the Agency for Health Care Research and Quality (AHRQ).<sup>2</sup> Please contact these agencies for specific information.

The first section of this chapter gives general policies relating to maintaining confidentiality and obtaining release of your research output. The second section describes confidentiality protection and the need to plan for having your research output cleared for release and how to document your clearance requests. The following sections give specific procedures and disclosure rules covering research output generated from economic microdata (business establishments and firms) and from demographic microdata (households and individuals).

### General Policies on Maintaining Confidentiality of Data and Clearing Release of Research Output

- *Absolutely no data and/or research results may leave the facilities without being examined and approved by the CES Disclosure Officer or a designated RDC administrator.* This applies especially to confidential information on individuals, households, establishments, and firms – including microdata; sensitive identifiers such as names and addresses; and the presence or absence of an establishment, firm, household, or individual person in a survey.
  - *CES releases approved research output only electronically, almost always via email; we do not release output in paper form.*
  - You must not discuss individual respondents in any manner -- in your research papers; via email, phone, or fax; or in casual conversation -- with persons outside Census Bureau facilities. This includes other project researchers with Special Sworn Status.
- The Census Bureau considers it important not only to avoid disclosure of confidential information, but also to avoid the *perception* of disclosure. (One reason: people or businesses will be less likely to participate in Census Bureau data collections if they *perceive* that their confidential data is being/will be disclosed, whether or not there is any *actual* disclosure.)
- All RDC printers are configured to place the following watermark on all printouts: “Disclosure Prohibited - Title 13 U.S.C. and Title 26 U.S.C.” (*Note:* If the label or watermark does not appear, please report this to your RDC administrator immediately.)

---

<sup>2</sup> An exception: The Census Bureau carries out the Medical Expenditure Panel Survey – Insurance Component (MEPS-IC) for AHRQ under Title 13 U.S.C. This survey is supplied by the Census Bureau and is covered here.

*IMPORTANT NOTE: Responsibilities for Protecting Confidential Census Data, including that containing Federal Tax Information owned by the Internal Revenue Service (IRS)*

*You should be aware of the following policies. For more detail, see Title 26, Appendix B, U.S. Code.*

- *IRS data may be used only for purposes authorized by the Internal Revenue Code. Similarly, confidential Census data may be used only for purposes authorized by the U.S. Census Bureau.*

*It is your responsibility to protect the data:*

- *Use data as authorized, but only for job-related purposes.*
- *Casual browsing of confidential data is illegal.*
- *Label all confidential data (see above for labeling requirements).*
- *Secure the area at all times, making sure to lock the door if you are the last to leave.*
- *Store confidential data in a secure enclosed cabinet when not in use.*
- *Ensure that listings and computer screens with confidential data are never visible to visitors or other non-authorized users.*
- *Dispose of all confidential data properly.*

- The creation of any printout with FTI must be recorded in one of the CES FTI Printout and Disposal Logs. Ask your RDC administrator where the log is located. The destruction (shredding) of any printout with FTI must also be logged. NOTE: only printouts with actual FTI need to be recorded in the logs. FTI printouts are those printouts *for economic and SSA projects only* that would not pass disclosure (even if you do not submit it for disclosure). If you are in doubt, log it.
- If you are using a data set for which official publications have not yet been released, you must not release any research results without consulting with CES. (Note: this is a relatively rare event, since usually CES obtains internal microdata sets for use at RDCs only after the Census Bureau has released the planned survey publications or public use files.)
- We require projects to emphasize model output rather than tabular output. In most cases, tabular output that is removed from the Center will consist of small one- or two-dimensional tables of variables that describe the samples that appear in the models.
- We STRONGLY DISCOURAGE the release of "intermediate output", i.e., output that you know will not appear in a publication. (Note: we do not limit the production of such output for examination inside the RDC, only its release.) Intermediate output often consists of detailed tables of preliminary descriptive statistics, or large numbers of similarly specified regression models, often based on 'thin' samples. Intermediate output increases disclosure risk for several reasons: releasing similar tables based on changing samples (adding or dropping small numbers of observations) causes complementary disclosure problems, since comparison of the released data tables may reveal information on who is in the sample as well as their characteristics. Moreover, intermediate output is difficult to track, and it can be difficult to associate with particular project publications (or even projects).

- You must use a standard disclaimer on all research papers and reports that indicates that your views/results do not represent the opinions or views of the Census Bureau. This requirement is imposed because your research products are not considered official Census Bureau data products unless they go through an official Census Bureau review to ensure that they meet certain Bureau standards.
- You must provide documentation of data sets and programs for any output you use in a report, memo, or paper.
- The parameters in certain disclosure rules themselves are confidential, and to reveal them is considered a disclosure. This policy is explained in the Disclosure Analysis section of this chapter.
- You must meet with the RDC administrator at the beginning of your project to discuss confidentiality and security policies and procedures.
- *You must account in your project planning for the time it takes to conduct disclosure analysis -- we cannot make rushed clearance decisions for events such as conference deadlines or the end of your project. Inform your RDC administrator as soon as you know you will need release of research output, and the approximate date you will need the output. You must allow sufficient time – up to three weeks – for a reviewer (usually an RDC administrator or the CES Disclosure Officer) to carry out disclosure analysis, and you should be prepared to discuss your request with your the reviewer. Review time can be longer in certain cases; e.g., if the output has to go to the Census Bureau’s Disclosure Review Board (DRB).*
- *If you work interactively, as early as possible, with the reviewer of your clearance request, you will minimize learning time on both sides, avoid costly misunderstandings, and speed approval of your clearance requests.*

*DISCLAIMER: The research in this (paper, report, presentation, etc.) was conducted while the authors were Special Sworn Status researchers of the U.S. Census Bureau at the \_\_\_\_\_ Research Data Center. Research results and conclusions expressed are those of the authors and do not necessarily reflect the views of the Census Bureau. This (paper, report, presentation, etc.) has been screened to insure that no confidential data are revealed.*

Note: The disclaimer may be modified to fit the number of authors and to include disclaimer information regarding other organizations.

## Requesting and Tracking Clearance of Project Output<sup>3</sup>

This section describes the process of removing project output from the RDC, including the roles of the various parties – you the researcher, the RDC administrator (including lead administrator), the CES Disclosure Officer and (sometimes) the Census Bureau’s Disclosure Review Board (DRB).

The clearance process involves two general steps, discussed in the sections below:

- Inform your RDC administrator that you are planning a request to release project output, and the approximate date you will need release of the output. Discuss the pending output with your RDC administrator, in light of the requirements spelled out below. In particular, ask the administrator (possibly consulting with the CES Disclosure Officer) whether the Disclosure Officer or the Census Bureau’s Disclosure Review Board (DRB) needs to review the output.

The DRB must approve release of any output from Census Bureau household censuses or surveys when any output cell (table or regression coefficient) is based on fewer than 75 unweighted unique individuals or when the output is presented for a geographic level lower than the state level. The CES Disclosure officer may also take any output to the DRB at his or her discretion, particularly large amounts of tabular output.

- Provide the materials for release, in the proper places with proper organization - including any required supporting materials. You will be asked to provide materials that allow us to document all approved releases (or *clearance*) of your research output. The documentation must indicate that the output was produced as part of your approved project, and it must allow us to document that published output does not disclose confidential information. To accomplish this, you must follow the process described below. There are two general types of project output:
  - Research results -- e.g., models, tables.
  - “Other” -- programs and (rarely) other documents such as research notes. “Other” documents do not require the same sets of forms as data/output and their release typically can be expedited.
- Request clearance of your output formally, via a module in the CES project management system. [Note: That module is still under development. Until it is put in place, please initiate requests through your RDC administrator. You need to discuss your clearance requests with your RDC administrator anyway, as described below.]

### **Clearing Research Results**

Clearing research results requires preparation and interaction with your RDC administrator and possibly the Disclosure Officer. When you have research results you wish to remove from the RDC, please do the following:

1. Make an appointment with your Administrator to discuss the output to discuss the clearance

---

<sup>3</sup> The glossary defines some confidentiality- and disclosure-related terms that may be unfamiliar to you. These terms are italicized the first time they appear in the sections below.

request. . (The Administrator may refer you to the Disclosure Officer.) The appointment should be made well enough in advance to allow for a thorough analysis of the results to be removed from the RDC. As a general rule, the appointment should be made at least a month in advance of the desired removal date.

2. Create a *dated clearance request subdirectory* for this request, e.g., rdc05:/rdcprojects/ch00000/disclosure/MMDDYYYY. This directory should contain the following:
  - A *clearance request memorandum*. (Also referred to as clearance request form.) We will provide you a template. This memo should contain
    - i) Names and descriptions of the *research output files* submitted for removal, and the programs that produced them.
    - ii) Names and descriptions of the *disclosure analysis files* corresponding to each output file, and the programs that produced them.
    - iii) A description of how the requested output fits into your approved research proposal.
    - iii) A description of the research sample—selection criteria and how the research sample differs from the samples underlying survey publications or other samples you have used.
    - iv) A glossary of variable names and descriptions.

In the management system, you will be given the option to upload an .rtf document (Clearance Request Document).<sup>4</sup> [Until the management system is available for this, you may email the document to your local RDC administrator, who will make sure that the Lead Administrator, and the CES Disclosure Officer receive it. ]
  - A directory called “output”, containing your output files (i.e., the files you want released).
  - A directory called “support,” containing *disclosure analysis files* that show the required *disclosure analysis statistics*. The specific statistics needed vary with the output and are described below, in the sections on models and tabulations.
3. When you have completed all this, please enter your request into the management system as described below.
4. Notes on the process of clearing research output:
  - a. Be prepared to discuss the clearance request with the person performing disclosure analysis. This may take more than one meeting, and may require you to make changes in your output to satisfy clearance requirements.
  - b. Provide properly documented programs, descriptions of data files, research output files and disclosure analysis files.

---

<sup>4</sup> Most researchers will create these in the lab and leave them in the supporting documents folder in the dated disclosure directory. However, some researchers fill them in outside of the RDC. This would allow them to upload the form.

- i) Many file formats are acceptable: most will consist of SAS or Stata programs, logs, and listings, sometimes including spreadsheets or data sets containing tables, parameter estimates, and the like.
  - ii) In properly documented output, *all variables should be LABELED and all variables should be in a consistent format*—e.g., identical variables (such as industry, IND) should have the same names and the same formats (numeric or alphanumeric) across data sets. All tables and output (research output or disclosure analysis output) should have clear *titles* indicating what the output is and indicating whether the analysis uses weighted or unweighted data (and what the weight variable is). The programs that produced the material are part of the documentation.
- c. For projects using the *same* samples over and over again (and hence having the same disclosure information), it is not necessary to copy the files with disclosure analysis statistics into the dated clearance request subdirectory each time, as long as you refer to the proper file(s) in the clearance request memorandum.
  - d. Working interactively with the reviewer will facilitate learning on both sides, avoid costly misunderstandings, and speed the review process. The clearer and more complete your request is, the more quickly it will clear.
  - e. The Disclosure Officer (or other designated CES employee) will download cleared files from your disclosure directory and will send them to you via e-mail. This e-mail also serves as a record of what files have been removed from the RDC.

## Disclosure Analysis for Model Output

1. For model output, disclosure is usually not a problem if the sample contains a sufficient number of observations—establishments, firms, households, or individuals—to start. The exceptions usually involve categorical variables. In a regression with dichotomous (0,1) variables, for example, the dummy may sometimes take values of 1 for plants belonging to a single firm or known small group of firms. These are disclosures because they identify which plants belong to that specific firm(s). For this reason, we ask you to make the following calculations for each and all of these regression coefficient “cells” for which you request to release information (e.g. a coefficient or descriptive statistics).
  - a. Number of establishments, firms, households, or individuals in the regression.
  - b. If the model has a continuous variable on the left hand side and if you are reporting indicator (dummy) variables on the right hand side, report the number of establishments, firms, households, or individuals in each dummy category.
  - c. If you have interaction terms, make these calculations for each of the possible categories in the cross classification (e.g., state by industry).
  - d. If the model has a discrete left hand side variable, calculate the numbers of observations for a *cross classification* of left hand side by right hand side discrete (dummy) variables. Here, “observations” refers to establishments and firms for economic (business) data, and unique individuals for demographic (household) data).
  - e. Note: You only need to make this calculation for dummy variables that you report. E.g., if you use dummies as “controls” or “fixed effects” for establishments, firms, households,

individuals, or states, but do not report them, you do not need to make these calculations for these dummies.

**IMPORTANT NOTE:** You need to make the above calculations *for the observations that actually appear in the model*. Sometimes estimation procedures in SAS or Stata automatically drop observations during an estimation procedure because of missing values of one of the variables. Make sure those same observations are also excluded from the sample used to calculate disclosure review statistics for that set of regression results or other statistics. Your samples may include different observations from model to model because of this. If this is the case, *we need a separate set of disclosure review statistics for each sample*. (Hint: a different N between regressions is a sign that this is occurring. Please make sure the observation count for your disclosure review statistics matches that for the output you are requesting)

2. If we find problems or estimates that do not meet our criteria for release, we may ask you to alter your specification. Alternatively, we may ask you not to report the values of the affected coefficients (as in the fixed effects mentioned above), or we may ask you to report that the coefficient is in a certain range (e.g., positive and significant). We will work with you to ensure, as much as possible, that your research results remain meaningful after respecification.
3. There can be special cases that don't fit this mold. Please discuss these in advance with the person performing disclosure analysis.

### Disclosure Analysis for Tabular Output

Tabular output, often called "descriptive statistics", is summary information consisting of, for example, counts, totals and basic distribution moments. We ask that you limit your tabular output to the minimum necessary to describe the sample used in your models, and how it might compare to the underlying population of interest. Disclosure criteria for economic (establishment and firm) output differ from the criteria for demographic (household and individual) output, so we discuss them separately below.

### Tables on Business Establishments and Firms (Economic Data)

1. For the *entire table as a whole, and all table cells individually*, provide the following information:
  - a. Number of establishments.
  - b. Number of firms.
2. If the table cell contains only plant or firm counts or employment statistics, the above is sufficient.
3. If the table contains payroll or other production or cost data, we require measures of whether one firm or a very small number of firms dominates the cell. These details vary by survey. For many surveys, this disclosure statistic is calculated based on value of shipments or another *key variable*. The disclosure statistic is also calculated for certain variables that are not closely correlated with the key variable (e.g., "lumpy" variables like investment). Some surveys require separate disclosure analysis on all continuous variables.

Table 1 on the next page summarizes the current primary disclosure rules for economic data. See the notes to the table for a description of the "Standard Rule" (SR). We also require the number of

firms to be at least a certain number, greater than  $n$ . The parameters of these rules are confidential, but we will provide them to you. We also will provide programs you may use to calculate the statistics we need.

4. Tables of summary statistics (means, variances, and related statistics) intended to describe the distributions of variables present certain special issues:
  - a. Tables may not contain minima, maxima, medians, or modes of a variable, or other statistics which are data from individual observations. (To illustrate, to report that the maximum value of variable  $X$  is 79 is to disclose that at least one observation has  $X=79$ .)
  - b. Care must be taken in presenting other statistics intended to describe the shape of distributions. We prefer that you present general statistics on distributional shape (e.g., skewness, kurtosis measures) rather than showing anything closely related to individual observations (e.g., quantiles). If you feel you need to present quantiles, we will give you certain procedures that, while still useful, will produce “fuzzy” estimates of quantiles that contain some uncertainty.
  - c. If we find problems or estimates that do not meet our criteria for release, we will do one of three things:
    - i) Ask you to *collapse*—that is, combine—certain cells. This will avoid disclosure problems at the expense of output detail, and is the preferred course of action.
    - ii) *Suppress* (e.g., replace by "D") the numbers in the affected cells. Note, in addition, we will probably need to suppress other cells from which someone could come sufficiently close to reconstructing the data in the suppressed cell -- this is called *complementary disclosure*. Complementary disclosure is by far the more difficult, time-consuming part of the disclosure analysis exercise.
    - iii) Take the output to the DRB. Their decisions are final.

**Table 1: Primary Disclosure Rules for Economic Data (Business Establishments and Firms)**

(See Notes at end of table)

Data Set	Primary Disclosure Rules	Comment
Census of Manufactures and Annual Survey of Manufactures (LRD)	<ol style="list-style-type: none"> <li>1. Continuous variables: SR (see table notes) on value of shipments and capital investment.</li> <li>2. Employment counts require a firm count only.</li> </ol>	<ol style="list-style-type: none"> <li>1. TVS is a “key variable” used to cover disclosure for variables considered correlated with it (i.e., most variables). Investment is a "lumpy" variable that is not considered highly correlated with shipments so disclosure is done separately for it. If TVS is not included in the study, then shipments or payroll are acceptable substitutes.</li> <li>2. The ASM and other surveys use unweighted numerator, weighted denominator in calculating shares. Our rules are more conservative. But most of studies using the ASM use unweighted data, so our procedures are appropriate.</li> </ol>
Economic Census – sectors outside manufacturing	<ol style="list-style-type: none"> <li>1. Continuous variables – SR on sales measure.</li> <li>2. Employment – firm count only</li> </ol>	Sales is a key variable.
Pollution abatement Costs & Expend. (PACE)	SR on every continuous variable	PACE variables not considered correlated; PACE expenditures can be lumpy, with some groups of plants known to have only certain types of expenditures..
Manuf. Energy Cons. (MECS)	Same as PACE	Similar to PACE
Surv. Of Manuf. Tech (SMT)	Firm count	SMT variables not considered very sensitive by themselves; no disclosure in publications.
Research & Dev. (R&D)	SR on every variable	Similar to PACE
Quarterly Financial Reports (QFR)	SR on all variables	
National Employer Survey (NES)	Continuous variables: plant count Discrete variables: Smaller plant count	Rules formulated in first survey year (1994). Survey considered less sensitive than other economic surveys, because there are no certainty strata, geography is limited, and most variables are discrete responses to questions about types of training or hiring practices. We may increase numbers & make disclosure firm-based.
Medical Expenditure Panel Survey - Insurance Component (MEPS-IC)	SR - total employment	
Survey . of Bus. Owners (SBO)	Firm count.	
Longitudinal Business Database (LBD)		Payroll and employment are Key Variables

Notes on Table 1:

- SR means “Standard Rule” (actually standard rules) – (a) An (n,k) rule such that the top n firms in each cell may account for no more than k percent of the variable (e.g., value of shipments) in each cell. Census Bureau economic (business) programs used this rule before 1992. (b) a p% rule, which requires that the sum of the smallest n (n>2) firms must account for less than p% of the value for the largest firm. Most Census Bureau programs have used this rule since 1992. (c) A further firm count that is higher than n in either the (n,k) or p% rules. References: FCSM (2005), chapter 4, sec. B.; Hundepool et al. (2007) section. 4.2.
- The table does not give rule parameters – i.e., n and k in the (n,k) rule, n and p in the p% rule, or the firm counts. Your RDC administrator will provide those you need to know. *The parameters are confidential; see "IMPORTANT NOTE" below.*
- The above rules are for primary disclosures. In addition, we will very likely need to suppress additional cells to ensure that it is not possible to calculate the values of suppressed cells using relationships derived from other cells.
- When two surveys are merged, use the more conservative rules. E.g., for the NES, use LRD rules for continuous variables when NES is merged to LRD.
- You may not remove data for a cells suppressed in the publications. We may ask you to examine publications to confirm this.

**IMPORTANT NOTE:** *To repeat: the rule parameters and the disclosure programs that contain those parameters are confidential.*

## **Tables on Households and Individuals (Demographic Data)**

To the extent possible, all tabulations from demographic survey or decennial census data must be created from a public-use version of the data file.

Tabulations created from the nonpublic data files, will be at the state level or above, and any cell must contain data from at least 75 unique individuals (unweighted). In these cases, the number and size of tabulations generated must be very limited. The Census Bureau's Disclosure Review Board (DRB) must review any tabulations that do not meet these criteria.

In cases where the researcher has linked third-party microdata or geographic-based aggregate data with the Title 13 confidential data and wishes to run tabulations, the RDC administrator will inform the Disclosure Officer. The DRB will, at its discretion, review the proposed tabulations prior to their removal from the center.

The DRB meets every Monday (with a few exceptions; e.g., holidays). The Disclosure Officer, who is a DRB member, will attempt to bring your request to the DRB within two meetings; and you will need to submit an easy-to-follow summary of what your requested output describes, how that output relates to your overall approved project, the issues in the output requiring DRB review, and why you think your output does not contain a potential disclosure problem. You should include how you will use your output to make particular points. This will help the DRB suggest ways to respecify your output, if needed, so that it does not present a disclosure risk, but still gets your intended message across. If you are using terminology or statistics that would not be familiar to those outside your field, you may be asked to write a description of those terms or statistics that someone not in your field can easily understand. These additional documents are meant to make it easier for the DRB (composed of people knowledgeable in statistics, but from varied backgrounds) to understand your output and make decisions on it.

For further reference on disclosure analysis rules for tables on households and individuals, please see the disclosure analysis rules for special tabulations for the 2000 Decennial and ACS datasets. These rules are posted on the Census website:

2000 Decennial: <http://www.census.gov/population/www/cen2000/sptabs/disclosure.html>

ACS: [http://www.census.gov/acs/www/Products/spec\\_tabs/drb\\_rules.htm](http://www.census.gov/acs/www/Products/spec_tabs/drb_rules.htm)

## ***Clearing Other Files – Programs and research notes***

Program files often present little risk, so we can usually release them fairly easily; the same is true for general project notes. However, we still must review them, and they can present disclosure risks if you are not careful. So, please minimize the number of requests, and the size (number of programs or length of notes) in each request. Please follow these guidelines to eliminate risk and make it easy for us to release them:

- Do not include code that edits (or otherwise refers to) individual observations via their identifiers (as opposed to groups), or that includes unreleased information such as quantiles. Use separate programs or “include” statements for this code.
- Use “macros” or include statement for other code that appears in many programs; e.g., data manipulation statements that prepare data for analysis. CES can usually release the macros, and the reviewer needs to look at this code only once. (With planning, this practice also can speed up

your coding in general.)

- Do not include notes in a program that describe “sensitive” information; e.g., the relationship between an individual and Social Security Numbers. When in doubt, CES will remove such code.

We can clear research notes if they are brief and do not include confidential information such as notes on specific observations you may have “cleaned,” or numbers that CES has not cleared for release. Again, we must look at all such notes before clearing them, so keep them as simple as possible.

To submit a request, first create a dated directory under your project’s “disclosure” directory and copy the files you want released into it. You do not need to provide a clearance request memo.

### ***Decisions and Release of Output***

When the reviewer decides whether a file (or group of files) is approved for clearance, the management system will send you (the originator of the disclosure request) an e-mail informing you. The Disclosure Officer or the reviewer will send you the output files via a separate email, as a zipped attachment.

If you disagree with decisions on clearance of research output, you may discuss these decisions with the reviewer and (if appropriate) the CES Disclosure Officer, who may take them to the DRB. The DRB’s decisions are final.

## Bibliography

The following is a short annotated bibliography giving references on many of the topics discussed in this manual, as well as certain other topics related to privacy, confidentiality, and data access. At the end we list some useful related web sites.

Doyle, P., J. Lane, J. Theeuwes, and L. Zayatz (eds.). *Confidentiality, Disclosure, and Data Access: Theory and Practical Applications for Statistical Agencies*. Amsterdam: Elsevier Science B.V., 2001.

Duncan, G.T., T. B. Jabine, and V.A. deWolf (eds). *Private Lives and Public Policies*. Washington, D.C., National Academy Press, 1993. An excellent reference on privacy, confidentiality, and data access issues related to the government statistical system.

Federal Committee on Statistical Methodology (FCSM), "Report on Statistical Disclosure Limitation Methodology." Statistical Policy Working Paper 22. Statistical Policy Office, Office of Information and Regulatory Affairs, Office of Management and Budget, May 1994 (Second Version 2005). Available at <http://www.fcsm.gov/working-papers/spwp22.html>.

Hundepool, A., J. Domingo-Ferrer, L. Franconi, S. Giessing, R. Lenz, J. Longhurst, E.S. Nordholt, G. Seri, and P. De Wolf. "*Handbook on Statistical Disclosure Control*." Version 1.0.1, March 2007. Available at [http://neon.vb.cbs.nl/cenex/CENEX-SDC\\_Handbook.pdf](http://neon.vb.cbs.nl/cenex/CENEX-SDC_Handbook.pdf)

Interagency Confidentiality and Data Access Group. "Checklist on Disclosure Potential of Proposed Data Releases." Interagency Confidentiality and Data Access Group, an interest group of the Federal Committee on Statistical Methodology, 1999. Available on FCSM web site below.

Jabine, T., "Procedures for Restricted Access." *Journal of Official Statistics*, 9(2). 537-590, 1993. An excellent and oft-cited background paper used in preparing Duncan, Jabine and deWolf.

Jewett, R. "Disclosure Analysis for the 1992 Economic Census." A manual describing the disclosure programs used for the 1992 Economic Census. The manual and a version of the programs are available on the FCSM web site cited below.

Willenborg, L., and de Wall, T. *Statistical Disclosure Control in Practice*. Lecture Notes in Statistics 111. New York: Springer-Verlag, 1996.

\_\_\_\_\_. *Elements of Statistical Disclosure Control*. Lecture Notes in Statistics 155. New York: Springer-Verlag, 2001. The two books by Willenbourg and de Wall are general references, perhaps a bit more technical than the Eurostat manual.

Web sites of interest:

- Federal Committee on Statistical Methodology: <http://www.fcsm.gov>. Contains information on FCSM activities, and links to information on confidentiality and disclosure analysis, including conferences and working groups.
- American Statistical Association Committee on Privacy and confidentiality home page: <http://www.amstat.org/comm/cmtepc/> Contains information on committee activities, and links to information on confidentiality and disclosure analysis, including conferences and working groups.
- Searchable U.S. Code website: <http://uscode.house.gov/lawrevisioncounsel.shtml>.

## Appendix A: Glossary

This appendix is a glossary of terms that can be divided into two general groups: terms related to confidentiality and terms related to the specific procedures for clearing research output at Census Bureau Research Data Centers. Please note many of the terms in the first group have multiple definitions; we have given only one here. For good discussions, see Duncan, Jabine, and deWolf (1993) pp. 22-24, and the web site of the American Statistical Association Committee on Privacy and Confidentiality.)

*Administrative records* - systems of records collected and/or maintained by government agencies for the purpose of administering programs – i.e., for *administrative purposes*. Administrative records are distinct from the systems of information collected exclusively for *statistical purposes*, such as those that the Census Bureau produces under authority of Title 13 U.S. Code.

*Business establishment* - a location at which business is carried out. Usually has a permanent plant number (PPN) associated with it.

*Business firm* - one or more establishments under common ownership. In our files, usually identified by the "ID" or "CFN" (Census file number).

*Clearance request memorandum* - a memo, contained in the clearance request subdirectory, describing a request to clear (release) research output.

*Clearing of research output* - the approval to remove research output from the secure RDC facility. Only the RDC administrator or other authorized Census Bureau employees may clear research output.

*Confidentiality* - the legal status that indicates that some body of data or information is sensitive and must be protected against loss, misuse, destruction, or unintentional change.

*Dated clearance request subdirectory* - a directory, under the project clearance directory, in which you place the material for each clearance request: the clearance request memorandum, the research output (and the programs that produced it), and the disclosure output (and the programs that produced it).

*Disclosure* – the inappropriate attribution of information to a data subject, whether an individual or an organization. Disclosure occurs when a data subject is identified from a released file (*identity disclosure*), sensitive information about a data subject is revealed through the released file (*attribute disclosure*), or the released data make it possible to determine the value of the characteristic of an individual more accurately than otherwise would have been possible (*inferential disclosure*). Comments: (1) Disclosure of sensitive data does not have to be exact; disclosure is considered to occur if a characteristic of an individual (e.g., income) is estimated sufficiently closely. (2) Revealing the presence or absence of a unit in a data collection typically is considered to be a disclosure.

*Disclosure analysis* - the act of ensuring that research output does not disclose confidential information about individual business establishments, business firms, households, or individuals.

*Disclosure output (disclosure analysis output)* - certain output, in addition to the research output, that provides the information we need to ensure that your research output does not disclose confidential information.

*Disclosure output files (disclosure files)* - computer files containing disclosure output.

*Disclosure statistics (disclosure analysis/review statistics)* - certain statistics that enable us to determine whether or not the research output can be cleared. The statistics include information on the numbers of observations (establishments, firms, households, or individuals) in all cells of a table or all categories in models; in some cases they include measures of the possible dominance of a cell total by a small number of observations (usually firms).

*Household* - all persons who occupy a particular housing unit as their usual residence; or who live there at the time of the interview and have no usual residence. The usual residence is the place where the person lives and sleeps most of the time. This place is not necessarily the same as a legal residence, voting residence, or domicile.

*Primary Disclosure Rules for Establishment and Firm Data* - rules for determining when a cell discloses confidential information. The two main rules: (1)  $(n,k)$  rule -- a cell discloses confidential information if its  $n$  largest respondents contribute more than  $k$  percent of the cell value. (3)  $p$ -percent rule -- a cell discloses confidential information if the second largest respondent can use the cell value to estimate the largest respondent's contribution to within  $p$  percent.

*Project* - the research specified in an approved research proposal.

*Project clearance directory* - a directory (one for each project) on your computer in that contains overall information about clearance activity for the project.

*Research output* - output you would like to remove from the Research Data Center. This will consist of output from statistical models (e.g., regression and other econometric models), together with a limited amount of supporting descriptive statistics (e.g., means and variances of variables in your research sample). Note: Output may not be released that shows data for individual observations. This currently includes residual plots; e.g., plots of residuals against the dependent variable. (This policy regarding plots is under review and may be changed.)

*Research output files* - computer files containing research output, usually generated with statistical packages such as SAS or Stata. The output often includes tables that summarize the coefficient estimates from groups of models, and supporting tables of sample means and variances.

*Statistical purpose* - the motivation for the collection of information to generate an aggregate description of a group of persons or businesses. Data collected for statistical purposes may not be used to make decisions regarding a specific individual or business, although the information may result in policy changes that could result in benefits or costs to persons or businesses. Examples: investigating geographical factors related to crime, the relation between welfare benefits and educational levels of welfare recipients. Stands in contrast to data collected for *administrative purposes*, which may be used to make decisions regarding individual persons or businesses - e.g., to determine eligibility for welfare benefits.

**Account and Data Request Form**  
(Lists official CES office and computer policies)

CENTER FOR ECONOMIC STUDIES  
RESEARCH PROJECT COMPUTER ACCOUNT AND DATA REQUEST

I. Description of Work

Researcher Name:

Project Number:

Title:

Location of Research:

Starting Date:

Ending Date:

Duration of Work:

Principal Researcher:

Detailed Description of Work:

The U.S. Census Bureau's Center for Economic Studies (CES) and its network of Census Research Data Centers (RDCs) support and encourage research activity using Census Bureau microdata to improve Census Bureau programs and inform public decision makers. As an integral part of its ongoing research program to improve Census Bureau programs, the Center for Economic Studies will provide support for a research project described in this document.

The CES research program - carried out with both CES staff and researchers from outside the Census Bureau (called research associates) - is centered on empirical analysis of confidential data collected in the Bureau's regular survey and census programs. The research associates carry out research projects using secure CES facilities, at CES headquarters in Suitland, MD and at several Research Data Centers.

Through a network of Census Research Data Centers (RDCs) and the Suitland, MD facility, the Census Bureau grants researchers with approved research projects restricted access to non-publicly available data files in secure environments. These environments ensure that the confidentiality of the data are maintained, but at the same time, provide limited access to data that are not a regular part of the Census Bureau data publication program. These research projects benefit the Census Bureau's programs. Access to non-publically available data files is a privilege, not an entitlement. This privileged access carries with it specific obligations and responsibilities, as specified in this document. This activity is authorized under Title 13 U.S.C. Sections 6, 8, and 23c.

Project Description: (See Attachment A: Proposal)

Benefits to the Census Bureau: (See Attachment B: Benefits Statement)

## II. Center for Economic Studies Furnished Equipment, Software, Data and Services

The Center for Economic Studies (CES) shall provide the following facilities, and support towards the execution of this work during the specified period:

1. A physical work space that is suitable to perform research oriented tasks.
2. Access to a thin-client computing device that is used to communicate with the CES data and compute server in a graphical interface mode.
3. Remote access -- only from within the RDC secure facility -- to the CES data and compute server, with sufficient dedicated data storage space to support the work as defined within this account and data request.
4. Usual and reasonable computer support. Researcher may consult with CES staff on computer-related topics to better understand the computing environment in which this work is conducted.
5. Usual and reasonable data support. Researcher may consult with CES staff in order to better understand the data that CES has provided to the researcher. CES will make every reasonable effort to provide researcher with documentation for each data set to which they are provided access.

The Center for Economic Studies shall provide the following computer software towards the execution of this work during the specified period:

1. The UNIX operating system, with all usual system tools, and utilities.
2. A graphical user environment, such as CDE, GNOME or KDE, configured to support the requirements of the researcher.
3. The Open Office automation package.
4. The following statistical/econometric software packages:
  - GAUSS
  - SAS
  - STATA
  - MATLAB
5. The following additional commercial software package:
  - HLM 5.04 for Unix
6. Based upon the estimated data and computing requirements for your project, you will be allocated sufficient disk storage to complete the work as defined in this document. Requests for additional work space must be submitted in writing to the RDC Administrator for forwarding to

CES Headquarters for approval.

The Center for Economic Studies shall provide to the researcher direct access to the following analytical datasets. These data will be provided in a SAS Dataset format. Any conversion of these SAS datasets to other formats shall be carried out by the researcher, and not by CES staff. Note that requests for additional data will require modification to this statement of work, or a new approved research proposal at the discretion of CES.

<Data set list>

The researcher acknowledges and agrees that Census Bureau data are collected for the purpose of producing published aggregate statistics. Use of these data for other purposes is of variable reliability. All survey and census results contain measurement error and may contain sampling error. Available information about these potential errors are provided or referenced in the appropriate publications. The Census Bureau recommends that data users incorporate this information into their analyses as these errors could impact inferences. Researchers analyzing the data in order to create their own estimates are responsible for their validity and shall not cite the Census Bureau as the source of these derived estimates, but only as the source of the underlying core data.

Researcher shall obtain and provide to CES any necessary approvals for the use of non-Census data used at CES from CES and from the data producers, as appropriate (See Attachment C: [Data Permissions for Researcher Provided Data](#)). Researcher shall provide these written approvals to designated CES staff. Without written authorization, non-Census Bureau data may not be loaded onto CES computer systems. The following data sets will be provided by the researcher for use in the project described in this document.

<Data set list>

### III. Protection of Confidential Information Provided to the Researcher

The researcher who participates in this project shall abide by all the following conditions, in order to guarantee the confidentiality of any and all data provided to them during the execution of the work defined within this account and data request and its attachments. All researchers will adhere to the following rules of behavior. These rules clearly delineate the responsibilities for all individuals with access to confidential Center for Economic Studies data. These rules will be enforced through sanctions commensurate with the severity of the infraction, including the possibility of criminal prosecution and the termination of the project at the discretion of CES.

All researchers who have access to confidential information in this project shall obtain Special Sworn Status (SSS) from the Census Bureau before work may begin. The Census Bureau will handle the administrative details of granting the researcher SSS. As individuals with SSS, the project researcher agrees to conform to all applicable confidentiality and security provisions of the Bureau of the Census and the Department of Commerce. In particular, these provisions state that no one may disclose any information that would allow identification of Census Bureau data on any individual survey respondent (business firm, business establishment, household, or individual). Furthermore, the data supplied may be used for statistical purposes only. Disclosure of confidential information is prohibited by law (Title 13 U.S.C. section 214) and is punishable by a fine of up to \$250,000 (Title 18 U.S.C. section 3559 and 3571) or a prison term of up to five years (or both) (Title 13, U.S.C., Section 214, as amended; Title 18, U.S.C., Sections 3559 and

3571).

[Once a researcher provides CES with all the information required to process an application for SSS, the Census Bureau requires a minimum of two weeks to approve the application. Researcher may not start or participate in a project without SSS approval.

After obtaining Special Sworn Status, but prior to accessing any data, a researcher must take the Title 13 Awareness Training and be certified for Title 13 Data access. This training takes place at the RDC lab under the supervision of the RDC Administrator. This certification process is repeated annually in March for the life of the project. ] This section needs updating. For example, all training must be taken prior to submission of SSS applications.

All researchers shall be provided with a Researcher Handbook. This handbook contains regulations and policies that researchers shall follow. Researcher acknowledges and agrees that violation of any rules or regulations contained in this handbook shall be enforced as would any other provision in this account and data request, including possible termination of the project at the discretion of CES.

#### Office Security Policies

1. Researcher agrees to keep security clearance and entrance information (identification, key code, card, lock combination, door key) current and secure. The researcher agrees not to share or lend security clearance and entrance information to anyone else.
2. Researcher agrees to keep the lab door locked and to ensure that is locked whenever they are the last person to leave the lab at any time.
3. Researcher agree to sign in every day that they work on the project at the lab.
4. Researcher must not allow persons without SSS status into the lab without prior approval by a designated CES staff member. In addition, all visitors will be escorted by designated CES staff members at all times.
5. Researcher agrees to place all output containing confidential information that they wish to discard into the designated receptacle(s), and only into these receptacles.
6. Researcher must not discuss confidential information over the telephone, by fax, or in e-mail.
7. Designated CES staff members (or other Census Bureau employees, as determined by CES) must examine all materials (especially research output) that any researcher wishes to remove from secure Census facilities, to prevent disclosure of confidential information. The researcher agrees to submit all materials to designated CES staff members for review prior to removing these materials from the lab. The researcher agrees that they will not remove research output before obtaining explicit approval from a designated CES staff member.

#### Computer Security Policies

1. Researcher acknowledges and agrees that all their computing activities will be monitored by

CES staff and that information from the monitoring programs will be accumulated and saved in a database. This information will be used to identify security violations, and to assess system usage.

2. Researcher must not include passwords in any computer files, such as the body of an electronic mail message or in a login script.
3. Researcher must not write their computer access password(s) down and must not give their password(s) to any other users.
4. Researcher who suspects that any password has been compromised agree to change the affected password immediately.
5. CES Computer Staff personnel may employ password-checking software to determine the "guessability" of user passwords. If you are informed by a CES staff member that your password has been cracked, you will be required to change it immediately.
6. You will not be able to load any software, "home-grown" programs, or data onto the lab machines. Programs and data must be uploaded into your project directory by CES staff.
7. ZIP-type drivers (or other large removable storage devices) are prohibited in the lab.
8. Laptop computers and any other portable computing devices that include an internal or external modem (e.g., Palm Pilot) are prohibited in the lab. Researcher agrees to never bring a laptop or other portable computing device that incorporates a modem into the lab.
9. Researcher agrees to schedule all visits to the lab. Researcher acknowledges and agrees that access to the lab will be denied unless a reservation has been made.

#### Data Access Policies and Confidentiality

1. Researcher is granted access only to the data authorized for the approved research project. Researcher agrees to access only those data specified in this computer account and data request and its attachments. Requests for new Census Bureau data will require researcher to submit a new research proposal to CES for review under standard project review procedures.
2. Researcher may supply data sets from outside the Census Bureau for use in the approved research project, and only for use in the approved project. If a researcher-supplied data set is from public data sources, the researcher must submit written documentation stating the source(s) of the data and confirming that the data are public. The researcher is responsible for documenting any possible restrictions on the use of the data set. If the data set is from a non-public source (e.g., an administrative data set), the researcher must supply to CES a copy of written authorization to use the data set, including any possible restrictions on the use of the data set. The researcher will document any possible restrictions on the use of any data set the researcher brings to the project. Researcher(also agrees to provide CES with data set documentation. Researcher acknowledges and agrees that CES shall be allowed to keep a copy of any data set the researcher brings to the project, that CES may make the data set available to other researchers for their approved projects.

3. Researcher agrees not to ask other researchers at the lab for access to their data sets, and agree not to give other researchers access to their data sets.
4. Researcher agrees to perform only analyses that support the work specified in this statement of work and its attachments. Researcher agrees not to carry out analyses for any colleagues or other researchers who do not have approved research projects.
5. Researcher acknowledges and agrees that research data files will not contain sensitive identifiers such as name and address, unless explicitly provided for in this statement of work and its attachments. If access to sensitive identifiers is required, a CES staff person will oversee the access. Social Security Numbers, as person identifiers, will never be provided to the researcher.
6. Researcher acknowledges and agrees that absolutely no data and/or research results shall leave the facilities without being examined and approved by a designated CES employee. Researcher agrees to have all materials examined by a designated CES staff member prior to removing those materials from the lab.
7. Researcher acknowledges and agrees that the Census Bureau considers it important not only to avoid disclosure of confidential information, but also to avoid the perception of disclosure. Researcher agrees to avoid any perception of disclosing sensitive information concerning the data they were provided.
8. Researcher acknowledges and agrees that all printed material, screen displays, and other media containing data that have not been cleared for release must contain the following label(s):  
  
"Disclosure Prohibited-Title 13 and Title 26 U.S.C." (if FTI data).
9. Researcher agrees not to discuss individual survey respondents in any manner in research papers, in email, telephone, or fax; or in casual conversation with persons outside the Census Bureau facilities. In addition, researcher must not discuss individual respondents' data with any other researcher not affiliated with their project.
10. If researcher is using a data set for which official publications have not yet been released, they agree not to release any research results without consulting with CES staff.
11. Researcher agrees to use the following standard disclaimer on all research papers and reports:  
"This paper (chapter) reports the results of research and analysis undertaken while the author(s) was a research affiliate at the Center for Economic Studies at the U.S. Census Bureau. It has undergone a Census Bureau review more limited in scope than that given to official Census Bureau publications. Research results and conclusions expressed are those of the author and do not necessarily indicate concurrence by the Census Bureau. It has been screened to insure that no confidential information is revealed."

The disclaimer may be modified to accommodate the number of researchers and other disclaimer information (e.g., regarding the researcher's sponsor or employer). In some cases, CES will require a somewhat different disclaimer. CES will inform the researcher in such cases.

12. Researcher acknowledges and agrees that any research data sets produced as part of this project can be made available to CES or to other researchers under approved research proposals, at the discretion of the Census Bureau.

## Disclosure Analysis

Disclosure analysis is a process by which CES staff ensure that no confidential information concerning survey respondents are included in any information that is allowed to leave CES facilities. Designated CES staff members shall review all research output for inadvertent disclosure of confidential information prior to release of the output. To aid in this disclosure process, each researcher shall maintain documentation of all data sets and programs that are created as a result of the work described in this document.

The process of disclosure analysis takes time and advance planning. Immediately upon beginning the project, the researcher agrees to meet with at least one designated CES staff member to discuss the specific output the researcher wishes to remove from the Census Bureau and the process of disclosure analysis. These discussions will cover all information the researcher must provide to CES staff for disclosure analysis, and the specific form in which the information must be provided (e.g., as SAS data sets or printouts, with variable and output labeling).

1. Researcher acknowledges and agrees that release of some analytical results specified in this work plan may not be feasible due to problems identified during disclosure analysis. In particular, the researcher acknowledges and agrees that it may not be feasible to release any specific tabular descriptive statistics. Researcher acknowledges and agrees that CES decisions are final regarding the form and content of any research results released to the public.
2. The researcher agrees to submit supplementary information, in addition to the research output they wish to remove, to the CES staff members who carry out the disclosure analysis. The researcher agrees to supply all such supplementary output in the form requested.
3. The researcher agrees to allow sufficient lead time for CES staff to perform disclosure analysis. The amount of time needed will depend on the specific results to be removed and the form in which the required information is submitted to CES staff. To aid CES staff members in carrying out disclosure analysis, the researcher agrees to meet with these staff members at one or more agreed-upon times, specified well in advance.
4. Due to the complexity of disclosure analysis, any delays in providing the required information or in meeting with CES staff until the project is completed or nearly completed may greatly delay or even prevent the researcher from receiving any output. Researcher acknowledges and agrees that the timing of the release of any research outputs is determined by CES staff.
5. For purposes of this account and data request, the parameters of rules used in disclosure analysis are themselves confidential under Title 13, and to reveal them is considered a disclosure. Researcher acknowledges and agrees that the disclosure rules shall not be discussed by the researcher with any persons other than CES staff members.
6. Before the first release of research output, the researcher agrees to provide a statement of benefits to Census Bureau programs resulting from the work, as described in the approved project proposal. CES will determine whether the statement is sufficient to release the research output.
7. Before the final release of research output, the researcher must provide a technical memorandum summarizing the key findings pertinent to Census Bureau programs, including

whether/how the accomplished benefits differed from those proposed and describing their potential application. At its discretion, CES may require further statements of benefits before other releases of research output.

8. In the event of litigation based on this research, Census Bureau and U.S. Government confidentiality and disclosure avoidance rules cannot be waived. Moreover, Census Bureau personnel would not be made available to testify regarding the data and analysis the researcher develops.

#### IV. Project Results

The researcher agrees to provide final end products at the completion of all activities described in Section I. of this account and data request. In addition, the researcher agrees to maintain project documentation that describes in detail all researcher activities, all researcher created data sets, all researcher created programs, and all researcher created outputs (such as tabulations, analytical results, and listings). More specifically, the researcher agrees to provide to CES the following project results:

1. Progress reports on the project as requested. These reports will document the activities accomplished to date and provide the designated CES staff member access to all aforementioned detailed project documentation.
2. Any and all research papers produced as part of the project, for inclusion in the CES Discussion Paper Series (or any other appropriate Census Bureau series, as designated by CES staff). These research papers may also appear in other paper series (e.g., university or National Bureau of Economic Research working paper series). The researcher agrees to provide CES with a copy of the paper as it appears in other paper series.
3. Copies of any published reports, journal articles, book chapters, slides from conference presentations (and the like) resulting from this project should be submitted to CES. Any publications must include a disclaimer and acknowledgment, as previously described.
4. A report on the data work conducted as part of the project. This report will document the main data sets used and created, and the programs used to accomplish this work. The report will contain enough detail to allow CES staff members to replicate the work.
5. A report containing information sufficient to allow other competent researcher(s) to reconstruct the research results. This will be a record of data set names (with descriptions of their contents), program names (with descriptions of what the programs do), and interrelationships among them. The principal researcher(s) will provide such information to CES staff in machine-readable and printed form.
6. A statement that explains how the research project provided a benefit to the Census Bureau. A technical memorandum summarizing the key findings pertinent to Census Bureau (particularly, Chapter 5, Title 13) programs, including whether/how the accomplished benefits differed from those proposed and describing their potential application.
7. A non-technical summary of the research results.  
Researcher acknowledges and agrees that failure to comply with the terms stated above

regarding research results shall result in the termination of all data access, and that CES will reject all subsequent project proposals until the researcher provides the results to CES. In particular, researcher acknowledges and agrees that CES, at its discretion, will not release research output until the researcher provides required documentation of benefits to Census Bureau programs, as specified above.

## V. Project Schedule

### 1. Supremacy of Census Bureau and Data Supplier Administrative Requirements

This computer account and data request does not supersede the administrative procedures and requirements of the Census Bureau and other data providing agencies. During the course of this project, changes may occur in administrative procedures and/or requirements of the Census Bureau and/or data providing agencies governing the use of the data provided to the researcher. In addition, changes may occur in the availability of data under both the governing MOU's in existence as of the date of this account and data request, and related Memorandum of Understanding (MOU) that come into existence during the course of the project. In such cases, relevant portions of this document may no longer be valid and shall be superseded by new language in order to accommodate administrative requirements of the Census Bureau and/or data providing agencies.

The project period will extend from the effective date MMM DD, YYYY through MMM DD, YYYY

I certify that I have read this computer account and data request, understand my responsibilities, and agree to the terms specified above:

---

Signature of Researcher Date

---

FOR INTERNAL USE ONLY

SERVER:

USERID:

USER NAME:

PRIMARY GROUP:

ACCOUNT EXPIRATION DATE:

DATE REQUEST FILLED:

NAME OF PERSON FILLING REQUEST:

This section needs updating. Anu receives the ADRF, usually as an attachment to email. The signed last page of the ADRF can be uploaded to the management system by the RDC administrator as a scanned file.

Send signed request, if using Federal Express or other delivery service, to:

James Lessard or Mark Mildorf  
Center for Economic Studies  
U.S. Census Bureau  
8905 Presidential Parkway  
Room 211, WP2  
Upper Marlboro, MD 20772-2653

If using U.S. Postal Service, send to (expect 2-3 weeks delivery time):

James Lessard or Mark Mildorf  
Center for Economic Studies  
U.S. Census Bureau  
4700 Silver Hill Road Stop Code 6300  
Washington, D.C. 20233-6300

## Appendix B: Clearance Request Memo

### REQUEST FOR CLEARANCE OF RESEARCH OUTPUT Center for Economic Studies and Research Data Centers

\*\*\*\*\*

\* This area will be filled out by CES

\* Project:

\* Clearance Request Number:

\* Date:

\* Submitted by:

\* Submitted to:

\* Cleared for release:

\* Cleared by:

\*\*\*\*\*

#### 1. GENERAL INFORMATION

- a. Name of this request's subdirectory under the project's main clearance directory:
  
  
  
  
  
  
  
  
  
  
- b. Please describe the outputs you wish to clear (Attach extra paper if you need more space):
  
  
  
  
  
  
  
  
  
  
- c. Please state how the outputs are part of the research project as approved (Note: If these outputs are described in your proposal, merely refer us there.)
  
  
  
  
  
  
  
  
  
  
- d. Please indicate how you expect the output to be presented (Check all that apply):  
  
 Journal paper  
 Working paper (Don't forget about the CES Discussion Paper Series)  
 Dissertation  
 Book chapter  
 Presentation at a conference  
 Report (e.g., put out by policy organization)  
 Memo for internal use  
 Supporting or intermediate output not to be published in any of the above  
 Other (please specify):

## 2A. DESCRIPTIONS OF RESEARCH SAMPLES:

Describe your Research sample(s) or "cuts" of data used in research output. For each sample, please describe your selection criteria and how the research sample differs from the samples underlying survey publications or other samples you have used. Take as much space as you need for each; add samples as needed.

SAMPLE 1:

SAMPLE 2:

SAMPLE 3:

## 2B. RELATIONSHIP BETWEEN SAMPLES

Describe how your samples relate to each other (e.g., if you have two samples, is one a subsample of another?) In the cases of samples and subsamples, there is an implicit third sample, the difference between the two. Please describe this sample above. We probably will need to examine any implicit samples as well.

## 2C. RELATIONSHIP TO OTHER PUBLICATIONS

Describe how your samples may relate to similar samples from other projects or from survey publications. (e.g., how your sample of an industry in the LRD differs from the Census of Manufactures or Annual Survey of Manufactures files in the LRD).

### 3. OUTPUT FILES

For each research output file to be removed, please enter the following information.

- a. File name (e.g., output.lst)
- b. Description of file (e.g., tables relating to \_\_ or models of \_\_)
- c. Program that produced the file (e.g., output.sas or model.do)
- d. The Research Sample Number that underlies the file
- e. The disclosure analysis file's name - the file that contains supporting statistics
- f. The name of the program that generated the disclosure analysis file (e.g., output\_disc.sas).

Note, if the disclosure information is produced in the same program as the research output to be removed, please cut and paste this into a new file.

- g. Results (we enter this) - Whether the file was cleared or not.
- h. Comments - any relevant information you or the person who clears the file may wish to add.

FILE NUMBER: 1

FILE NAME:

FILE DESCRIPTION:

RESEARCH OUTPUT PROGRAM:

RESEARCH SAMPLE NUMBER:

DISCLOSURE ANALYSIS FILE NAME:

DISCLOSURE ANALYSIS PROGRAM:

RESULTS:

COMMENT:

FILE NUMBER: 2

FILE NAME:

FILE DESCRIPTION:

RESEARCH OUTPUT PROGRAM:

RESEARCH SAMPLE NUMBER:

DISCLOSURE ANALYSIS FILE NAME:

DISCLOSURE ANALYSIS PROGRAM:

RESULTS:

COMMENT:

#### 4. VARIABLE DEFINITIONS

Please include all variables in the research output you wish to remove and in supporting disclosure output.

VARIABLE NAME:  
DEFINITION:  
SOURCE:  
COMMENTS:

VARIABLE NAME:  
DEFINITION:  
SOURCE:  
COMMENTS:

VARIABLE NAME:  
DEFINITION:  
SOURCE:  
COMMENTS:

VARIABLE NAME:  
DEFINITION:  
SOURCE:  
COMMENTS:

VARIABLE NAME:  
DEFINITION:  
SOURCE:  
COMMENTS:

#### 5. OTHER INFORMATION OR COMMENTS:

Please enter any further information you feel is relevant below.

## Appendix C: User-Provided Data Memo

Date:  
Project ID:  
Researcher Name:

### Physical Information (check media being used):

Media:

- \_\_\_\_\_ CD (ISO9660)
- \_\_\_\_\_ floppy disk (3 ½ inch, 1.44 MB drives)
- \_\_\_\_\_ DVD
- \_\_\_\_\_ DLT tape (GNU tar, give blocking factor)
- \_\_\_\_\_ other (specify)<sup>5</sup>

### File Information

File Name:  
File Size (bytes):  
Number of records:  
List of Variables (if SAS, attach 'proc contents'; if ascii, attach record layout):

Format:

- \_\_\_\_\_ ascii
  - Record End Deliminator:
    - \_\_\_\_\_ carriage return
    - \_\_\_\_\_ line feed
    - \_\_\_\_\_ other (specify)<sup>6</sup>
  - \_\_\_\_\_ SAS (v8 with sas7bdat extention)
  - \_\_\_\_\_ rtf text
  - \_\_\_\_\_ ascii text
  - \_\_\_\_\_ Other (specify)<sup>7</sup>

### Permissions

- \_\_\_\_\_ Downloaded from web without charge or restrictions, give specific URL.
- \_\_\_\_\_ Purchased. Provide copy of receipt and documentation verifying if public use or proprietary.
- \_\_\_\_\_ Proprietary. Provide written permission from source to use data in the approved project (letter or memo identifying file and giving permissions and/or any restrictions on use, e.g. retention data, etc.)

Description: (1-2 sentence description of what is in the file)

---

<sup>5</sup> CES strongly recommends Researcher provide files on CD, or DVD. Consult with your RDC Administrator before providing alternative media.

<sup>6</sup> Consult with your RDC Administrator if other delimiter is used.

<sup>7</sup> CES strongly recommends Researcher provide SAS or ascii files. Consult with your RDC Administrator before providing alternative format.

## **Appendix D: Legal Information**

This appendix provides information (summaries and quotes) on the laws underlying the policies spelled out in this researcher handbook. The most important law is Title 13 U.S. Code for the Census Bureau. Titles 18 and 26 for the Internal Revenue Service (IRS) is also extremely important to the Census Bureau because IRS federal tax data are used in many Census Bureau data programs – chiefly in constructing and maintaining sampling frames. The IRS legislation imposes certain restrictions on this use.

### ***Census Bureau Legislation – Title 13 U.S. Code***

#### **General Information**

Title 13 authorizes the Census Bureau to carry out specific data collection programs. It also requires the Census Bureau to maintain the confidentiality of the data the Census Bureau collects, and it authorizes the Bureau to use "temporary staff" to "assist the Bureau in carrying out the work authorized by this title."

Several chapters of Title 13 authorize the Census Bureau to collect data. This is important because Under IRS law (Title 26, section 6103), Federal tax data may be used only in programs authorized under Chapter 5 and to carry out "related statistical activities authorized by law." The Decennial Census and the Economic Census and related survey programs are authorized under Title 13 Chapter 5. An important exception is the Quarterly Financial Reports (QFR), which is authorized under Title 13 Chapter 3, but is specifically authorized to use Federal tax data under Title 26 section 6103 as a "related statistical activity authorized by law."

The following sections provide relevant sections from Title 13, along with brief summaries.

#### **Protection of Confidential Information -- Sections 9 and 214**

##### *Summary:*

With the exception of data from censuses or interim surveys of governments, information collected and/or compiled under authority of Title 13, is confidential. The following restrictions apply:

- Data may be used only for statistical purposes
- No publication may permit identification of data for a particular business or individual
- Only sworn employees may examine individual reports
- Temporary staff (i.e., persons with Census Bureau Special sworn Status) may assist in performing work authorized by Title 13 but only if those persons are sworn to observe the confidentiality provisions in Section 9 of the statute.
- Conclusion: All Researchers at RDCs must assist the Census Bureau in performing work authorized by Title 13, and all must be sworn to observe the confidentiality provisions of Title 13, Section 9.

##### *Law:*

CHAPTER 1—ADMINISTRATION

SUBCHAPTER I— GENERAL PROVISIONS

**Sec. 9. Information as confidential; exception**

- (a) Neither the Secretary, nor any other officer or employee of the Department of Commerce or bureau or agency thereof, or local government census liaison may, except as provided in section 8 or 16 or chapter 10 of this title or section 210 of the Departments of Commerce, Justice, and State, the Judiciary, and Related Agencies Appropriations Act, 1998.(1)
- (1) use the information furnished under the provisions of this title for any purpose other than the statistical purposes for which it is supplied; or
  - (2) make any publication whereby the data furnished by any particular establishment or individual under this title can be identified; or
  - (3) permit anyone other than the sworn officers and employees of the Department or bureau or agency thereof to examine the individual reports. No department, bureau, agency, officer, or employee of the Government, except the Secretary in carrying out the purposes of this title, shall require, for any reason, copies of census reports which have been retained by any such establishment or individual. Copies of census reports which have been so retained shall be immune from legal process, and shall not, without the consent of the individual or establishment concerned, be admitted as evidence or used for any purpose in any action, suit, or other judicial or administrative proceeding.
- (b) The provisions of subsection (a) of this section relating to the confidential treatment of data for particular individuals and establishments, shall not apply to the censuses of governments provided for by subchapter III of chapter 5 of this title, nor to interim current data provided for by subchapter IV of chapter 5 of this title as to the subjects covered by censuses of governments, with respect to any information obtained therefor that is compiled from, or customarily provided in, public records.

**Sec. 214. Wrongful disclosure of information**

Whoever, being or having been an employee or staff member referred to in subchapter II of chapter 1 of this title, having taken and subscribed the oath of office, or having sworn to observe the limitations imposed by section 9 of this title, or whoever, being or having been a census liaison within the meaning of section 16(2) of this title, publishes or communicates any information, the disclosure of which is prohibited under the provisions of section 9 of this title, and which comes into his possession by reason of his being employed (or otherwise providing services) under the provisions of this title, shall be fined not more than \$5,000<sup>8</sup> or imprisoned not more than 5 years, or both.

**Authority for Temporary Staff (Special Sworn Status) – Section 23(c)**

*Law:*

CHAPTER 1—ADMINISTRATION  
SUBCHAPTER I— OFFICERS AND EMPLOYEES

**Section 23. Additional officers and employees (excerpts)**

- (b) The Secretary may utilize temporary staff, including employees of Federal, State, or local agencies or instrumentalities, and employees of private organizations to assist the Bureau in performing the work authorized by this title, but only if such temporary staff is sworn to observe the limitations imposed by section 9 of this title. Authorizations For Data Collection Programs – Chapter 5

**Data Programs Authorized by Title 13 Chapter 5**

*Summary:*

The following data collection programs are authorized by Title 13 Chapter 5.

- Quinquennial Economic (Business) Censuses (Sec. 131)
- Decennial Census of Population and Housing (Sec. 141)
- Quinquennial Census of Governments (Sec. 161)
- Intercensal Population Estimates (Sec. 181)
- Current Survey Programs - on the subjects covered by Title 13 Census programs (Sec. 182)

*Law:*

---

<sup>8</sup> Note: The fine has been increased to \$250,000, as stated elsewhere in this handbook.

*Economic (Business) Censuses:*

CHAPTER 5

SUBCHAPTER I - MANUFACTURES, MINERAL INDUSTRIES, AND OTHER BUSINESSES

**Sec. 131. Collection And Publication; Five-Year Periods**

The Secretary shall take, compile, and publish censuses of manufactures, of mineral industries, and of other businesses, including the distributive trades, service establishments, and transportation (exclusive of means of transportation for which statistics are required by law to be filed with, and are compiled and published by, a designated regulatory body), in the year 1964, then in the year 1968, and every fifth year thereafter, and each such census shall relate to the year immediately preceding the taking thereof.(9)

*Decennial Census Of Population And Housing:*

CHAPTER 5

SUBCHAPTER II - POPULATION, HOUSING, AGRICULTURE, IRRIGATION, AND UNEMPLOYMENT

**Sec. 141. Population And Other Census Information.**

- (a) The Secretary shall, in the year 1980 and every 10 years thereafter, take a decennial census of population as of the first day of April of such year, which date shall be known as the "decennial census date", in such form and content as he may determine, including the use of sampling procedures and special surveys. In connection with any such census, the Secretary is authorized to obtain such other census information as necessary.

*Census of Governments*

CHAPTER 5

SUBCHAPTER III - GOVERNMENTS

**Sec. 161. Quinquennial Censuses; Inclusion Of Certain Data**

The Secretary shall take, compile, and publish for the year 1957 and for every fifth year thereafter a census of governments. Each such census shall include, but shall not be limited to, data on taxes and tax valuations, governmental receipts, expenditures, indebtedness, and employees of States, counties, cities, and other governmental units.

**Sec. 163. Authority Of Other Agencies**

This subchapter does not revoke or impair the authority of any other Federal agency with respect to the collection or release of information.

CHAPTER 5

SUBCHAPTER IV - INTERIM CURRENT DATA

*Intercensal Population Estimates*

**Sec. 181. Population**

- (a) During the intervals between each census of population required under section 141 of this title, the Secretary, to the extent feasible, shall annually produce and publish for each State, county, and local unit of general purpose government which has a population of fifty thousand or more, current data on total population and population characteristics and, to the extent feasible, shall biennially produce and publish for other local units of general purpose government current data on total population. Such data shall be produced and published for each State, county, and other local unit of general purpose government for which data is compiled in the most recent census of population taken under section 141 of this title. Such data may be produced by means of sampling or other methods, which the Secretary determines will produce current, comprehensive, and reliable data.

*Survey Programs*

**Sec. 182. Surveys**

The Secretary may make surveys deemed necessary to furnish annual and other interim current data on the subjects covered by the censuses provided for in this title.

## Other Title 13 Data Programs

### *Summary:*

The following data programs are authorized by other parts of Title 13.

- Quarterly Financial Reports (Chapter 3)
- Foreign Commerce and Trade Statistics (Chapter 9)

### *Law:*

TITLE 13

Chapter 3

Subchapter IV - Quarterly Financial Statistics

#### **Sec. 91. Collection and publication**

(a) The Secretary shall collect and publish quarterly financial statistics of business operations, organization, practices, management, and relation to other businesses, including data on sales, expenses, profits, assets, liabilities, stockholders' equity, and related accounts generally used by businesses in income statements, balance sheets, and other measures of financial condition.

(b) Except to the extent determined otherwise by the Secretary on the basis of changed circumstances, the nature of statistics collected and published under this section, and the manner of the collection and publication of such statistics, shall conform to the quarterly financial reporting program carried out by the Federal Trade Commission before the effective date of this section under section 6(b) of the Federal Trade Commission Act.

(c) For purposes of section 6103(j)(1) of the Internal Revenue Code of 1986, the conducting of the quarterly financial report program under this section shall be considered as the conducting of a related statistical activity authorized by law.

TITLE 13

Chapter 9 - COLLECTION AND PUBLICATION OF FOREIGN COMMERCE AND TRADE STATISTICS

#### **Sec. 301. Collection and publication**

- (a) The Secretary is authorized to collect information from all persons exporting from, or importing into, the United States and the noncontiguous areas over which the United States exercises sovereignty, jurisdiction, or control, and from all persons engaged in trade between the United States and such noncontiguous areas and between those areas, or from the owners, or operators of carriers engaged in such foreign commerce or trade, and shall compile and publish such information pertaining to exports, imports, trade, and transportation relating thereto, as he deems necessary or appropriate to enable him to foster, promote, develop, and further the commerce, domestic and foreign, of the United States and for other lawful purposes.(18)

***Protecting Federal Tax Information (FTI) – Titles 26 and 18, U.S. Code***

This section contains a brochure (reformatted for inclusion in this handbook) that describes the need to safeguard and protect FTI. It summarizes the policies, the underlying laws, and punishments for violations.

Before gaining access to any data set containing FTI, you will be required to undergo awareness training and sign a statement indicating your understanding of and willingness to follow the policies that govern your restricted access to FTI.

## Quick Reference Guide

# Safeguarding & Protecting Federal Tax Returns & Return Information

The Census Bureau obtains Federal Tax Returns and Return Information (FTI) and related “Official Use Only” data from the Internal Revenue Service (IRS) for statistical purposes. As a recipient of these data, the Census Bureau must ensure that it establishes and maintains to the satisfaction of the IRS certain safeguards designed to prevent unauthorized access to and use of FTI, as well as to protect the confidentiality of that information. Failure to adequately protect these data can result in severe penalties, including fines, imprisonment, civil suits, and loss of your job.

## Definitions

All FTI and “Official Use Only” data should be treated as confidential, which means disclosure is restricted. **Disclosure** is the making known of FTI or “Official Use Only” data in any manner to anyone. There are three types of disclosures:

- **Authorized disclosure** – disclosing confidential data to other employees with a work-related “need to know.” This is legal.
- **Willful unauthorized disclosure** - knowingly providing confidential data to anyone who is not entitled to have them. This is illegal.
- **Inadvertent unauthorized disclosure** – mistakenly giving confidential data to a person who shouldn’t have it. Penalties can apply under Title 13.

**Willful unauthorized access** or inspection of FTI is simply looking at, or browsing through, data that are not needed to complete required tasks. This is illegal.

An **authorized use of FTI** at the Census Bureau is one that both Census and IRS officials have approved through the formal review and approval process.

## Important Facts

- ✓ All tax data, including names and addresses, are equally sensitive and are subject to IRS safeguard requirements and penalties.
- ✓ Files that contain both Federal Tax Returns and Return Information (FTI) and non-FTI data are referred to as “commingled” data sets and are subject to the same IRS safeguard requirements as FTI, no matter how limited the use of FTI.
- ✓ Each individual is bound by these Federal laws for life, regardless of his or her employment status.
- ✓ FTI data used at the Census Bureau are subject to the privacy and confidentiality provision of Title 26 Internal Revenue Code, Title 13 Census, and Title 18 Crimes and Criminal Procedure.
- ✓ Unauthorized disclosure of FTI is illegal. Only provide FTI to those with an approved, work-related “need to know.”
- ✓ Casual browsing of FTI for which you do not have an authorized “need to know” is illegal. For example, unless you have a “need to know,” you cannot inspect tax records of your neighbor or favorite TV personality.
- ✓ FTI must be secured at all times.
- ✓ No matter the media, FTI must be labeled as such.
- ✓ Listings and computer screens displaying FTI must not be visible to visitors or other non-authorized users.

## Additional Information

For more detailed information, visit:

### **IRS Publication 1075**

<http://ftp.fedworld.gov/pub/irs-pdf/p1075.pdf>

### **Administrative Records Handbook**

<http://cww.census.gov/po/>

### ***IT Security Office***

<http://cww.census.gov/itso/>

