

# FSRDC Research Proposal Guidelines

April 23, 2020

Persons wishing to conduct research on restricted Census Bureau data at a Federal Statistical Research Data Center (FSRDC) must submit a research proposal to the Census Bureau. The following guidelines describe the research proposal submission process.

Each FSRDC is a joint venture between federal statistical agencies (including the Census Bureau) and leading research institutions. All FSRDCs charge project fees for cost recovery. Some institutions maintain affiliations with their local FSRDC that cover fees for their faculty, staff, and/or students. Researchers will want to discuss the fees with the FSRDC Administrator and/or Executive Director for the FSRDC(s) they intend to use.

After you and your administrator have developed your proposal and submitted it, the proposal will undergo Census review. Researchers on approved projects must also complete a background investigation, including fingerprinting. It can take a number of months to complete these steps before you might expect to begin work.

Census has also partnered with other federal statistical agencies including the Agency for Health Care Research and Quality (AHRQ), the National Center for Healthcare Statistics (NCHS), the Bureau of Economic Analysis (BEA), and the Bureau of Labor Statistics (BLS) to host projects in the FSRDCs. Proposals to use data from other federal agencies are submitted to these agencies (see Appendix A for web links).

## Table of Contents

I.	The Proposal Process .....	1
A.	Preliminary Proposal Development .....	1
B.	Final Proposal Submission.....	1
C.	Proposal Content Guidelines.....	2
1.	Project Description .....	2
a)	Introduction.....	2
b)	Methodology.....	3
c)	Data Sources .....	4
d)	Project Output and Disclosure Risk Section.....	6
e)	Project Duration and Funding Source.....	6
2.	Benefits Statement.....	8
3.	Abstract.....	11
II.	Census Review.....	12
A.	Proposal Review Process .....	12
B.	Post Census Approval.....	14
	Attachment 1: Proposal Registration Form.....	16
	Instructions for Proposal Registration Form.....	17
	Example Proposal Registration Form .....	19
	Attachment 2: Predominant Purpose Statement .....	21
	Attachment 3 - Benefits to Census Bureau Programs.....	23
	Appendix A: FSRDC Data.....	24
	Appendix B: Useful Links .....	25

# **I. The Proposal Process**

## ***A. Preliminary Proposal Development***

Researchers who wish to conduct research at a Federal Statistical Research Data Center (FSRDC) should first contact the FSRDC administrator at the primary center where the research will be conducted (see list of [FSRDC contacts](#)). The researcher should discuss the proposed project with the administrator to determine whether the research fits with the Bureau's mandate, is feasible, is likely to provide benefits to Census Bureau programs under Title 13 of the U.S. Code, is designed to produce model-based statistical estimates, and that the research output will not pose undue risk of disclosing information about survey respondents. In order to assist in this assessment, the administrator may request a 2 to 3 page proposal overview from the researcher. This overview should include a summary of the research questions, the internal Census datasets to be requested, and a general idea of the project's methodology and output.

Once the FSRDC administrator and the researcher have agreed it is sensible to proceed, the researcher submits to the administrator a completed Attachment 1: Proposal Registration Form, which is then reviewed by the administrator and submitted to the Census Bureau.

The Census Bureau encourages researchers to assess carefully the time period over which they request access. Requests for extensions beyond the specified end date undergo careful scrutiny, must be justified, and generally are granted for circumstances beyond the control of the researchers (e.g. unexpected illness). A typical proposal duration is 36 to 60 months.

The proposal should include the PI's names as well as the names of all researchers, including RAs, associated with the project. The Census Bureau can host foreign citizens in the FSRDCs if they have lived in the U.S. for 3 years. Researchers that wish to work in the FSRDC that do not currently reside in the U.S. or have not lived in the U.S. for 3 years should discuss their circumstances with the FSRDC administrator early on. Doctoral students submitting a proposal for dissertation work are asked to include their primary advisor as a co-principal investigator. The Census Bureau recommends that the advisor also apply for Special Sworn Status (see below).

## ***B. Final Proposal Submission***

The final proposal consists of three separate documents in rtf (rich text) or doc/docx (MS-Word) format: project description, benefits statement, and abstract. These documents are submitted to the Census Bureau through your FSRDC administrator. It is expected that the proposal development process will involve multiple iterations of the proposal and, in particular, the benefit statement (called the Predominant Purpose Statement, PPS) prepared in conjunction with your FSRDC administrator before the final proposal is submitted. Your administrator is a resource for assistance in clarifying a range of potential issues, including topics related to feasibility and Census benefits.

## ***C. Proposal Content Guidelines***

The three document files should conform to the following specifications.

### **1. Project Description**

This document should describe the research question(s), methodology (including model equations to be estimated, how model variables will be measured, and hypotheses to be tested), Census and non-Census datasets to be used, expected outcomes, and a list of references cited. The proposal should be written so that it can be understood by a competent social scientist who is not necessarily a specialist in the field or on the topic within the field. Though the benefits to Census are presented in a separate proposal document (discussed in the next section), your proposal should discuss how your research benefits the Census Bureau.

Please keep in mind that the audience for a Census proposal is not the same audience that researchers typically address when writing research proposals. The Census Bureau has its own set of expectations and requirements, and so researchers often find that they must make significant revisions to preexisting proposals that were written for other purposes (e.g., funding grants). The FSRDC proposal should:

- Contain a title and the names of all researchers at the top of the first page.
- Include appropriate headings and subheadings throughout the document to assist reviewers in following the proposal narrative.
- Not include any appendices unless approved in advance by the FSRDC administrator.
- Be limited to no more than fifteen (15) single-spaced pages or thirty (30) double-spaced pages (inclusive of references).
- Use twelve point font and one-inch margins. Times New Roman or Calibri font are preferred.
- Include page numbers.
- Be written in revisable form text: rtf, doc, or docx format.<sup>1</sup>

The full proposal should contain the following sections; however, researchers may add additional sections as necessary. The table at the end of this section summarizes the mandatory components in each section of the proposal.

#### **a) Introduction**

This section should motivate interest in your project and provide an overview of the benefits of your project to the Census Bureau. As part of this overview, provide an overall description of your research questions and a brief discussion of the relevant literature. A full literature review is not necessary; some references are fine. Include a list of all internal Census Bureau datasets that your project will request. Discuss why you need non-public data and briefly describe what you will do with the data. The introduction should also state that the primary purpose of this research is to benefit the U.S. Census Bureau. It is also helpful, where appropriate, to connect the

---

<sup>1</sup> Some data partners request revisable form documents in order to perform electronic comparison in case they receive multiple versions.

proposal research to activities and initiatives performed by the Census Bureau.

All research must have a statistical purpose. In particular, Title 13 specifically prohibits the use of census micro-data for legal or regulatory purpose.<sup>2</sup> Please do not include analysis in your proposal that has or can be perceived to have a regulatory purpose. Proposals are also not expected to make specific policy recommendations.<sup>3</sup>

## **b) Methodology**

This section describes your empirical specification, the statistical equations you will estimate, what key variables are needed, how the data will be used, and how all datasets fit together.

It is important in this section to describe not only the methodology of your research questions but also the methodology of other benefits your project will provide to the Census Bureau. For example, if your project will evaluate the quality of certain variables in the restricted Census dataset, you should discuss in detail how you will evaluate these variables and what types of results you will report to Census.

The requested Census data needs to clearly map into the methodology section. Even though “Census Provided Data” is a separate section of the proposal (below), there should be a very clear connection between the methodology and the data. The purpose for each requested dataset and its role in the analysis should be clear. Many researchers overlook the “data crafting” aspect of the methodology. You need to be as detailed as possible in describing how measures will be constructed, and how the data will fit together. This description does not need to be a separate section or separate discussion per se, but it needs to be discussed where appropriate. Generally, this is accomplished by discussing how variables are constructed and/or which datasets will be used to construct those variables.

It is important to describe how, and at what unit of observation or aggregation datasets will be linked. For example, if one is linking the Annual Survey of Manufacturers (ASM), collected at the establishment level, to the R&D survey, which is firm based, the research methodology should incorporate this discussion. Also, linking survey data (e.g., the ASM and R&D are both surveys) to other survey data may lead to small sample sizes (and therefore potential disclosure issues) if the survey samples do not significantly overlap. For demographic data (e.g., linking the SIPP to externally provided contextual variables) detailing the level of geography at which this link will be performed is helpful.

Generally speaking, economic (firm and establishment) data sets share a common set of identifiers and can be linked together. The identifiers for demographic (household and person) census and survey data sets are most often designed to be used within the survey year and are

---

<sup>2</sup> Title 13, Chapter 1, Subchapter 1, section 9, Information as confidential. “Neither the Secretary, nor any other officer or employee of the Department of Commerce or bureau or agency thereof... may... use the information furnished under the provisions of this title for any purpose other than the statistical purposes for which it is supplied; or make any publication whereby the data furnished by any particular establishment or individual under this title can be identified... 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.”

<sup>3</sup> In addition, for proposals using establishment and firm data, we recommend avoidance of the word “policy” altogether. For example, substitute “decision maker” for “policy maker.”

not longitudinal, or linkable across demographic surveys. However, many demographic datasets can be linked together using Protected Identification Keys (PIKs); ask your RDC administrator for details. Record level linkage for economic data can be done using the Business Register. Record level linkage between researcher-provided data and demographic data must be performed by internal Census staff.<sup>4</sup> Note that linking cannot generate complete matches of an internal dataset to an external one; this would reveal existence of specific respondents in the Census data, which is a violation of Title 13 protections.<sup>5</sup>

It is important to convey to the reviewers that you have read the documentation that is publicly available about the datasets you are requesting and have thought through some of the implications of the data in relation to your research approach. A list of restricted-use data commonly used by approved projects in the FSRDCs is included in the links given in Appendix A. Often three sources of information about datasets are helpful: survey questionnaires, Census web pages detailing survey and sample design, and working papers that Center for Economic Studies and FSRDC researchers have written using the data (See Appendix B). Also, ask your FSRDC administrator to provide additional information or answer questions.

The unit(s) of observation for the analysis and the groups of units for which you will carry out analysis are both important. For each, be as specific as possible. Is the unit of observation an establishment, firm, household, or person? Alternatively, will your unit of observation be at a more aggregated level (e.g., Census tract or county)? Will you use some combination of units? Additionally, for what groups of units (what "levels") will you specify your models? Will you estimate models at (say) national or sub-national levels; if so, what levels? In addition, will you run separate models by gender, race, household structure, industry, etc.? If you have external or public-use data, how big are the samples that you will analyze? Information that describes the units of analysis and the groups of units/levels of samples you will be using for your analysis helps Census reviewers assess both the feasibility of your project and the risk of disclosing confidential information. For more on the latter, see section (d) following, and II.A below.

Write out the equations you plan to estimate. You do not need to give the exact functional form you plan to use, but write out in as much detail as possible the general equation and relevant variables, and talk about how the data will fit into the equation. You do not need to list every specification, nor every variable you will be using from each dataset, but discuss how the left and right hand sides of each equation will be measured, and if they are dichotomous, categorical, continuous, etc. Specify the dataset(s) from which these measures will come and how they will be constructed.

### **c) Data Sources**

Data sources should be divided into two sections: "Census Provided Data" and "Researcher Provided Data." (See below tables for examples)

---

<sup>4</sup> Circulation within Census of personally identifiable information (PII) such as name, address and SSN data is intentionally very limited, and excluded from the FSRDCs. Record level linkage is performed at Census by internal specialists within only a few departments. The researcher must pay the cost of having PIKs assigned to individuals in researcher-provided data.

<sup>5</sup> Furthermore, for business data with FTI, a complete match results in a violation of the IRS "fact-of-filing" information protection.

In the “Census Provided Data” section, specify the exact data you are requesting. Please consult the online tables in the links in Appendix A to see the data available through the FSRDCs and the years available for each dataset. If a dataset is listed as a separate entity in the table, it needs to be listed separately in your proposal. As examples, “LEHD Data” is insufficient information, instead list the requested datasets individually (e.g., EHF, ICF, etc.); the “Economic Census” includes eight datasets, list individually those requested; for the Longitudinal Research Database (LRD) please instead list the ASM, CMF and LBD.

For each requested dataset, include the years of data needed, a short description of how the dataset will be used, and a justification of the need for the restricted use version of the dataset. If you are requesting years of data not yet released, detail these with an “if available” caveat (e.g. CMF 2002, 2007, 2012, and if available 2017). The years of requested data need to be clearly justified, particularly for proposals requesting a long time series. If only datasets containing Federal Tax Information (FTI) are requested, note that no non-FTI data are requested.

It is straightforward for the Census Bureau to provide the standard FSRDC datasets that are listed in the links in Appendix A for an approved project. Data collected by the Census Bureau that is not listed is generally unavailable, not prepared for FSRDC use, or procedurally problematic in some way. If a need for non-standard data exists, discuss this with your FSRDC administrator, and incorporate this into your timing plans as requests for non-standard datasets considerably lengthen proposal evaluation.

If you plan to bring in non-FSRDC data to the project, list these datasets in a separate section called “Researcher Provided Data.” Please include the source for your external data, the years of the data, a short description of how the dataset will be used, and a justification of the need for the researcher-provided data. This list should be as comprehensive as possible. Future requests for researcher-provided data not listed in your proposal may require additional justification and approval. Note: public use Census data are considered external data (sourced from public web sites).

After the list of data, you should include a paragraph summarizing how all the data fit together, and how they will be linked and their expected overlap (if applicable).

**Census Provided Data – Datasets and Years**

<b>Dataset</b>	<b>Years</b>	<b>Reason/Use</b>
Annual Survey of Manufacturers (ASM)	2007-2017 (if available)	
Census of Manufactures	2007, 2012, 2017 (if available)	

**Researcher Provided Data – Datasets and Years**

<b>Dataset</b>	<b>Source</b>	<b>Years</b>	<b>Reason/Use</b>
Compustat	S&P (WRDS)	2007-2017	

#### **d) Project Output and Disclosure Risk Section**

In this section, you must describe the output you expect to ask to remove from the FSRDC. This description is important in assessing both the substance of the proposal and the risk of disclosing confidential information. Output from the FSRDCs must emphasize modeling output (regression coefficients, standard errors, and the like). Summary statistics (variable means, etc.) are allowed, but only to the extent that they support modeling output. For example, the kinds of output expected to be released are regressions and tables as found in an article in a peer-reviewed academic journal. You should make the point that your project will emphasize modeling output. Researchers who desire disclosure of large volumes of tabular output should request a special tabulation from the Census Bureau<sup>6</sup> rather than requesting access to micro data at the FSRDC.

Be clear in describing the unit(s) of analysis you will use and the groups of units ("levels") for which you will create research results. Are you only using certain industries or demographic groups? Your work will be cleared for release only if it passes certain disclosure criteria, including minimum sample size restrictions. For example, for a regression, the underlying sample must be large enough, as well as those subsamples as defined by indicator variables.<sup>7</sup> Particular care and advanced planning should be given if proposed estimates will be produced for "sub-state" geographies. Depending on the population of sub-state areas, noise injection may be required.

Note that including indicator variables in your analysis creates additional "cuts" at the sample (i.e., cross tabs), but only if you report the estimated coefficients associated with the indicator variables. Because of this potential issue, we encourage including statements in your proposal, as appropriate, for example: "Specifications will include detailed industry controls; however we will not release these fixed effect coefficients, as all we need is to note in our results tables that we included them."

Disclosure risk is a complex topic, and you will need to discuss your proposed output with your FSRDC administrator as you develop your proposal. See below Section II.A "Confidentiality protection and disclosure avoidance review requirements" for more specifics on what we look for in assessing proposals for disclosure risk.

Finally, in this section describe the deliverables you expect to produce, such as technical memoranda, Center for Economic Studies (CES) working papers, and the like.

#### **e) Project Duration and Funding Source**

This section should state the proposed duration of the project and expected intensity of use (e.g., 15 hours per week). List sources of funding for the FSRDC lab fees, as well as funding for the research project overall.

---

<sup>6</sup> Demographic tabulations: <https://www.census.gov/programs-surveys/popest/data/special-tab.html>

More about custom tabulations: [https://www.census.gov/about/policies/privacy/statistical\\_safeguards/custom-tabulations.html](https://www.census.gov/about/policies/privacy/statistical_safeguards/custom-tabulations.html)

<sup>7</sup> If you are using LEHD data, output must have been estimated using a sample that includes at least three states. The following language for LEHD proposals should be used: "We understand disclosed results using LEHD data must have been generated using multiple states."

Researchers often discover that it takes longer than anticipated to become familiar with the data and computing environment. For most projects, duration of at least 36 or 60 months is recommended. All projects are required to submit a yearly progress report/interim benefits statement.

## Summary of FSRDC Proposal Structure and Key Components

### NOTES:

- A. Proposal should not exceed 15 single-spaced pages.
- B. Refer to the FSRDC Proposal Guidelines for how to complete each section.

PROPOSAL STRUCTURE (Under each section are listed the mandatory components):

### 1. Introduction ( 1-3 pages)

- Overview of benefits to Census Bureau.
- Description of research questions, brief discussion of relevant literature and research.
- Basic overview of research plan, including a brief summary of all internal Census Bureau datasets requested.

### 2. Methodology (8-9 pages)

- Overview of empirical methodology.
- Subsections describing detailed empirical specifications, estimation methodology, key variables, and the specifics of how the data will be used in estimation.
- Subsection describing methodology used in providing benefits to Census Bureau.

### 3. Data (1-3 pages)

#### 3.1 Census provided data – datasets, years

- List of all dataset, including years.
- Description of planned usage.
- Summary of justification and need for restricted use data.

#### 3.2 Researcher Provided data – datasets, years

- List of all researcher-provided datasets, including years.
- Description of planned usage and need for researcher provided data.

#### 3.3. Data Linkages and Justification

- Description of how data will be linked together.

### 4. Project Output and Disclosure Risk (1-3 pages)

- Description of expected model-based output from the methodology.
- Description of expected supporting tabular output, if any.
- Description of products expected from the project including working papers and technical memos.
- Discussion of any possible disclosure risks and mitigation strategy.

### 5. Project Duration and Funding Source (less than one page)

- Discussion of needed timeframe, including any travel requirements for accessing FSRDC and how those requirements may affect work on project.
- Statement of funding source.

### 6. References

## 2. Benefits Statement

The benefits statement is contained in the Predominant Purpose Statement (PPS). Proposals must demonstrate that the research is likely to provide one or more Title 13 benefits to the Census Bureau. A research project must demonstrate that its predominant purpose is to benefit Census Bureau programs. If a project has as its predominant purpose one, or any combination, of the following criteria, then it will be considered to have the predominant purpose of increasing the utility of Title 13, Chapter 5 data. Researchers should consult an FSRDC administrator for more information. Any proposals seeking to use datasets that contain Federal Tax Information (FTI), generally speaking the establishment and firm based surveys and census, must also be reviewed by the Internal Revenue Service to ensure that the predominant purpose of the research is to contribute to Census Bureau programs under Title 13, Chapter 5 of the U.S. Code. The following are a list of approved Title 13 benefits.

The first four criteria apply only to projects that do not include Federal Tax Information (FTI). Projects requesting only demographic data (data for persons and/or households) can include these first four criteria.

1. Evaluating concepts and practices underlying Census Bureau statistical data collection and dissemination practices, including consideration of continued relevance and appropriateness of past Census Bureau procedures to changing economic and social circumstances;
2. Analyzing demographic and social or economic processes that affect Census Bureau programs, especially those that evaluate or hold promise of improving the quality of products issued by the Census Bureau;
3. Developing means of increasing the utility of Census Bureau data for analyzing public programs, public policy, and/or demographic, economic, or social conditions; and
4. Conducting or facilitating census and survey data collection, processing or dissemination, including through activities such as administrative support, information technology support, program oversight, or auditing under appropriate legal authority.

The remaining nine criteria apply to all projects.

5. Understanding and/or improving the quality of data produced through a Title 13, Chapter 5 survey, census, or estimate;
6. Leading to new or improved methodology to collect, measure, or tabulate a Title 13, Chapter 5 survey, census, or estimate;
7. Enhancing the data collected in a Title 13, Chapter 5 survey or census. For example: improving imputations for non-response; developing links across time or entities for data gathered in censuses and surveys authorized by Title 13, Chapter 5;
8. Identifying the limitations of, or improving, the underlying Business Register, Master Address File, and industrial and geographical classification schemes used to collect the data;
9. Identifying shortcomings of current data, collection programs and/or documenting new data collection needs;

10. Constructing, verifying, or improving the sampling frame for a census or survey authorized under Title 13, Chapter;
11. Preparing estimates of population and characteristics of population as authorized under Title 13, Chapter 5;
12. Developing a methodology for estimating non-response to a census or survey authorized under Title 13, Chapter 5;
13. Developing statistical weights for a survey authorized under Title 13, Chapter 5.

The PPS should address clearly and specifically how the project will provide one or more of the Title 13 benefits listed above. It has no length limitation, however clear and concise presentation is critical. Most commonly, researchers have written the benefits to address two criteria, Criterion 11 and one other.

Most FSRDC proposals include as a benefit preparing estimates of population and characteristics of population as authorized under Title 13, Chapter 5 (criterion 11). This is because, if one thinks of the Census survey life cycle, loosely speaking, as a series of four steps: planning, conduct, processing, and dissemination, then your research estimates are a key form of information dissemination. FSRDC-based research leverages unique skills of experts in a joint venture to generate new estimates of the population of individuals and firms to complement the variety of Census-produced publications, and further informs Census analysts of if and how the data can support meaningful understanding of the underlying processes of our economy. In writing benefits for criterion 11, include all of the analyses you plan to perform in summary form.<sup>8</sup>

Your access to the internal microdata is covered under legal statute because of your role as a consultant to assist Census with its work.<sup>9</sup> From a practical perspective, the implementation of this has included a norm for FSRDC proposals to meet at least two of the criteria listed above. Thus, a second benefit is proposed, most commonly an exercise connected to survey planning and processing activities at Census, for example an analysis about the data itself. The ideal benefit analysis leverages the same data preparation work as is intended for the estimates of population.

A good Census benefit accomplishes three elements: (i) a connection to Census program activity,

#### *Census Benefit Example 1*

An example of a Census benefit is to benchmark Census collected data to that of an outside source, to understand the quality of data produced through a Title 13, Chapter 5 survey. By linking Census establishment micro data to external sources of merger and acquisition activity as reported by firms to other federal agencies, projects have investigated the extent to which known acquisition events as identified in the external data are reflected in the Census business register. The terms of an acquisition can vary considerably, and these differences may be related to differences in reporting behavior. Cases of mismatch were explored and documented in a technical memorandum to Census Bureau staff, focusing specifically on the ability of the register to capture ownership dynamics.

---

<sup>8</sup> For some federal agencies, including the IRS, the PPS is the document and project scope they review and approve.

<sup>9</sup> Title 13, Chapter 1, Subchapter II, Section 23 states “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 [Information as confidential] of this title.”

(ii) articulates a specific analysis exercise, its purpose and approach, that can be clearly understood by reviewers and Census staff, and (iii) is focused, generally on only one of the above criteria. Discuss what might be found, why it could be informative, and, in particular, some specifics on how the analysis will be approached.

Related to (i) above, a benefit analysis that can be shown to have a relationship to activities Census is working on or potentially thinking about strengthens its relevance. Census program staff describe much about their activities on the Census web pages, easily accessible through searches.<sup>10</sup> Note that contributions that benefit other researchers, including Census researchers, are welcome (e.g., documentation contributions, or methods contributions that are of potential interest to academic researchers); however, these might not necessarily have application to Census activities.

The PPS document is formatted in a specific way. See Attachment 2 for the standard template. In the section entitled “Criteria”, specify which of the 13 benefit criteria listed above are relevant to your project. Then enter text under each relevant criterion describing your project’s benefits. Note that it is important not to change the format of the document, nor edit the paragraphs that are pre-printed there. Your FSRDC administrator will provide a project-specific template and assist you with the formatting and content of the PPS.

In the section “Requested Datasets” you should provide information on the datasets you will use in your project. If your project includes FTI datasets, then that section is filled out as follows: (i) enter each FTI dataset and the years, (ii) enter the FTI variables included in the FTI datasets (these are a standardized set for each data file – your administrator will help you with this), and (iii) if your project asks for a long time series, include a paragraph at the beginning of the section justifying the request.

An example for the Census of Manufactures is as follows:

**Census of Manufactures (CMF), 2002, 2007, 2012 and if available 2017**

FTI:	Description:
ST	State
IND	Industrial Classification (SIC/NAICS)
LFO	Legal Form of Organization

*Census Benefit Example 2*

An example of a Census benefit centers on the transition of the Decennial Census long form to the American Community Survey (ACS). This change provides advantages, such as ongoing updated information more frequently than once every ten years, however also provides some issues that may require adjustment to population estimates. Projects have investigated the effect of changes in question design, comparing data from the same question in cases of different wording, and seasonality issues due to the change from data collected in April of the decennial year to ACS continuously collected data for variables such as journey time to work and the commuting transportation mode.

<sup>10</sup> Proposals requesting LEHD data are expected to include benefit to the LEHD program.

CFN	Census File Number (includes EIN)
EIN	Employer Identification Number
COU	County
SW	Salaries and Wages (IRS Payroll)
TE	Total Employment
TVS	Shipments (IRS Receipts)

Under the section “Census Bureau Datasets”, list any internal, restricted Census Bureau datasets that your project will use, along with the years of data. If only FTI datasets are requested, note that no non-FTI data are requested. Under the section “User Provided Datasets,” list both external proprietary datasets that you will use as well as public use datasets (including public use Census datasets). If you will be linking any of the “User Provided Datasets” to the FTI datasets, describe at what level that linkage will occur (e.g., establishment level, Census tract level, etc.)

If your project includes FTI datasets, then in the “Need and Use of FTI” section, you must list every FTI variable in these datasets and describe why your project needs that variable and how it will be used. An example is provided below.

<b>FTI</b>	<b>Datasets requested that include the FTI</b>	<b>Description</b>
CFN	ASM, CCN, CFI, CMF, CMI, CRT, CSR, CUT, CWH, LBD	Census File Number (Contains EIN)
This field is used to link establishments across Census datasets within a year. For 2001 and earlier years, the CFN is also used to recover the Census firm identifier, embedded within the CFN. These links will be performed to construct a panel of establishments incorporating variables from the ASM, LBD and Economic Census files.		

The benefits document/PPS should not be repeated in the proposal although the proposal should clearly describe the benefits the project will provide. In addition, the PPS and the proposal should not reference each other; they are stand-alone documents. The two documents together with the abstract constitute the proposal submission. In addition, the PPS should be in executive summary form: no equations, no footnotes, and no references/bibliography. Finally, the PPS must be written in the third-person.

Appendix B includes some helpful links to source documents that can provide input to drafting benefit statements. It is important to discuss proposed benefits to Census with your administrator.

### **3. Abstract**

This document should be 150-250 words in length, written in 1st person, reference major datasets and key methods, and discuss expected findings. It should capture the essence of the project proposal similar to what would be expected of a scientific journal publication. At the top of the first page, the project title and the names of all researchers must be stated.

## **II. Census Review**

### ***A. Proposal Review Process***

The Census Bureau accepts proposals for review at any time during the year. Proposals are reviewed both by Census Bureau staff and external experts on subject matter, datasets and disclosure risk. Relevant data sponsors and data custodians also review proposals that request certain datasets, for example other federal and state agencies. See the sponsor column in the data tables in the links in Appendix A. For example, proposals seeking to use datasets that contain FTI must also be reviewed by the Internal Revenue Service. LEHD data are sponsored by the 50 states; many states require state review of proposals that request use of their data.

Research proposals submitted to the Census Bureau are reviewed based on five major criteria.

#### ***Benefit to Census Bureau programs under Title 13***

While each of the criteria is important in determining whether to approve a proposal, the project's potential to benefit Census Bureau programs carries the greatest weight.

#### ***Scientific merit***

This criterion relates to the project's likelihood of contributing to existing knowledge. When available, evidence that a Federal-funding agency such as NSF or NIH has approved the proposed research constitutes one indication of scientific merit.

#### ***Clear need for non-public data***

The proposal should demonstrate the need for and importance of non-public data. The proposal should explain why publicly available data sources are not sufficient to meet the proposal's objectives.

#### ***Feasibility***

Feasibility criteria include data availability and statute consistency. Internal Census reviewers will be focused on the ability of the data to support the proposed research, and will be interested in the proposal's discussions of measures and how the researchers plan to combine the data. The proposal must show that the research can be conducted successfully given the methodology and requested data. Here, researchers need to convince reviewers that they sufficiently understand the data they are requesting (unit of measurement, available measures, etc.), that their research questions can be answered in a disclosure-compliant way, that they have thought through and anticipated potential sample size, level of aggregation or other feasibility issues, and that the number of datasets requested is reasonable and well justified.

## *Confidentiality protection and disclosure avoidance review requirements*

Output from all research projects must undergo and pass disclosure review.

- The Census Bureau is required by law to protect the confidentiality of data collected under its authorizing legislation, Title 13, U.S. Code.
- Tabular and graphical output presents a higher risk of disclosure of confidential information than do coefficients from statistical models and is therefore strongly discouraged.
- Some data files are collected under the sponsorship of other agencies. In providing restricted access to these data, the Census Bureau must adhere to all applicable laws and regulations.
- Researchers may be required to sign non-disclosure documents of survey sponsors or other agencies that provide data for their research projects.

We look for several things in evaluating proposals for disclosure risk. To do this, the proposed methodology and output must be clearly stated in the proposal, especially the information requested above in Sections I.C.1.b and I.C.1.d. We realize that, because you are conducting research, you probably don't know all the details in advance. Nevertheless, we need enough detail to assess whether the proposed project can succeed without posing undue risk. In particular:

- We need to determine whether there may be possible thin "cells" in your output, where for this purpose, a "cell" is the group of observations (individuals or businesses) underlying any result you may release. To evaluate this, we need clear and accurate information on the types of output to be requested (e.g., models tabulations, graphs) and the units of analysis underlying the output (e.g., individuals and businesses, or groups of these); and the groups (levels) for which you will request output (e.g., industry, gender, geography, and possible crosses of these). For example, detailed industry-by-geography cells become thin very rapidly.
- We need to know whether variables included in models are discrete or continuous; both are allowable, but discrete variables (especially dummy (0, 1) variables) define "cells" that we must take into account.
- We need to know whether you will request tables, graphs, or maps, which can be especially problematic.
- We need to know whether you plan to produce sub-state estimates.
- We will also look for types of output that are relatively unfamiliar to us, such as from new statistical techniques. We must try to assess their disclosure risk, and we welcome any insights you may have on this.

Reviewed proposals receive one of two ratings:

- **Approved.** The proposal successfully addresses all of the review criteria mentioned above.
- **Not Approved.** The proposal fails to meet most or all review criteria. Subject to approval of the FSRDC administrator, not-approved proposals may be resubmitted as a new proposal, after addressing reviewers' concerns in a substantial revision.

The Project Review Coordinator (PRC) communicates the outcome of the review process to the researcher, which includes a review synopsis, an explanation for the decision, and copies of the expert reviews.

### ***B. Post Census Approval***

Following Census approval, the PRC will coordinate reviews from data sponsors and other federal agencies (including the IRS) as needed. Approval of research proposals by the Census Bureau (and the IRS if FTI is requested) is but one step in the process before research can commence. In many instances, the Census Bureau must obtain permissions to access certain data. This process can range from a few weeks to many months depending upon the data requested.

Once a project has all necessary approvals, all researchers who expect to access FSRDC data must undergo a background investigation including fingerprinting. After completion of the background check, the Census Bureau grants Special Sworn Status (SSS) to each researcher, which subjects them to incarceration of up to five years and/or fines of up to \$250,000 if they knowingly or inadvertently disclose confidential information on individuals, households or businesses. All SSS individuals must take annual trainings in proper Data Stewardship and the use and protection of Title 26 data.

All approved research projects are governed by a written agreement between the researcher(s) and the Census Bureau. The agreement stipulates the start and end dates for the project, and responsibilities of both parties with respect to procedures and practices. Each researcher on the project must sign an agreement with the Census Bureau. If researchers are added to the project after the project starts, these researchers will also each sign an agreement. Separately, an agreement with the FSRDC partner institution may be required, as well as arrangements for FSRDC fee payments.

### **Timing**

The review process is both lengthy and rigorous, requiring researchers to exhibit patience throughout. We continually strive to simplify and compress proposal review times, and are committed to process improvement for proposal review. However, experience has demonstrated that due to the number of steps, participants and data permissions involved, elapsed time between the final proposal submission and the actual commencement of research generally includes multiple months. This duration also varies substantially by individual proposal depending on the scope and complexity of the proposal, the number and variety of datasets, and other agencies involved.

Researchers can help speed up the process by:

- Adhering closely to all practices and procedures for proposal submission.
- Rationalizing and simplifying the scope and data requested in the proposal where appropriate.
- Working closely with their FSRDC administrator on proposal development and on any

- requested revisions or clarifications to proposals or predominant purpose statements.
- Providing the Census Bureau with the terms of use for any proprietary external datasets.
  - Processing their SSS paperwork quickly and completely.

We are proud of the many research projects conducted over the last three decades throughout the FSRDC network, and are grateful for your interest.

# Attachment 1: Proposal Registration Form

Title:

Principal Investigator:

Date:

### *Prerequisites Checklist*

1. I have read these research proposal guidelines.
2. I have corresponded with an FSRDC administrator about my prospective project.
3. I understand that the predominant purpose of this project must benefit the Census Bureau data programs.
4. I understand that there are usually lab fees for using an FSRDC.
5. I understand that work is performed in a restricted access environment and only cleared results are allowed to leave the FSRDC.
6. I understand that the approval process may take a considerable amount of time, for example, if review by other agencies or data owners is required or if data availability issues exist.

I have performed each checklist item above: Please initial \_\_\_\_\_

Duration:

Funding Source:

Description:

Census Datasets:

Years of Data:

Primary FSRDC Location:

Secondary FSRDC Locations:

Benefits:

### *Research Team*

Name	Affiliation	Email	Type	Involvement	Citizenship / Years in U.S.	FSRDC

## **Instructions for Proposal Registration Form**

These instructions provide some context for entering the information about your project in the Proposal Registration form. Note that it is not necessary to have all information exact to fill out this form, as the information can be updated any time up until the proposal is submitted for Census review. It is expected that you will update the information as you go along as part of the proposal development activity. The final version of this form will be submitted along with your proposal.

**Title:** The title of your proposal.

**Principal Investigator:** Researcher that will submit the proposal for Census review, and will be the central focal point of communication with Census on proposal and project affairs.

**Date:** Today's date

Please perform each checklist item and type your initials to indicate agreement.

**Duration (months):** Number of months of access to the data in the FSRDC starting from the first day of logon (after project approval) to FSRDC project completion.

**Funding Source:** List sources of funding for the FSRDC lab fees, as well as funding for the research project overall.

**Description:** The description is intended to give Census staff a short idea of what your research project is about, as well as the potential Census benefits.

**Census Datasets:** A list of the internal census datasets requested in the proposal. Please list the acronyms for the datasets from the links in Appendix A. The acronyms are in parenthesis after the dataset names in the tables.

**Years of Data:** The time span of the data you will use for your research project.

**Primary FSRDC Location:** The main FSRDC location from which the project will be administered.

**Secondary FSRDC Locations:** Any additional FSRDC locations researchers from the project will use regularly.

**Benefits:** The relevant benefit criteria the project will address. Please list the benefit numbers as taken from the benefit criteria listed in Attachment 3. Please note that only benefit numbers #5 to #13 are valid for projects requesting economic (Title 26) data.

### *Research Team*

List the full name, institutional affiliation, email address, and other requested information for all researchers affiliated with the project.

Name	Affiliation	Email	Type	Involvement	Citizenship/ Years in U.S.	FSRDC

Type: PI = Principal Investigator, CI = Co-investigator, RA = Research Assistant.

Involvement: Indicate the expected involvement of each researcher (the level of involvement determines whether the researcher will need to undergo the Special Sworn Status (SSS) background check and whether the researcher needs a computer account on the project):

A = Directly accessing data in the lab (SSS and Computer Account),

B = Viewing undisclosed output in the lab (SSS, no Computer Account),

C = Advising on or coauthoring papers based only on disclosed output outside the lab (no SSS, no Computer Account)

Citizenship: Indicate country of citizenship. If not a U.S. citizen, indicate number of years of residence in the U.S.

FSRDC: Indicate the FSRDC the researcher is planning to use.

## Example Proposal Registration Form

Title: R&D Investment and Plant Productivity

Principal Investigator: Researcher #1

Date: 03/13/2017

### *Prerequisites Checklist*

1. I have read these research proposal guidelines.
2. I have corresponded with an FSRDC administrator about my prospective project.
3. I understand that the predominant purpose of this project must benefit the Census Bureau data programs.
4. I understand that there are usually lab fees for using an FSRDC.
5. I understand that work is performed in a restricted access environment and only cleared results are allowed to leave the FSRDC.
6. I understand that the approval process may take a considerable amount of time, for example, if review by other agencies or data owners is required or if data availability issues exist.

I have performed each checklist item above: Please initial\_\_ RES\_\_\_\_\_

Duration (months): 36 months

Funding Source: National Science Foundation

Description: The project will investigate the effect of investment in research and innovation on the evolution of plant productivity over time. Benefits to Census include linking external patent data to explore the extent to which R&D activity is captured in the BRDIS and SIRD surveys.

Census Datasets: BRDIS, SIRD, ASM, CMF

Years of Data: 1992-2016, and if available 2017-2019

Primary FSRDC Location: Boston

Secondary FSRDC Locations: Chicago

Benefits: 5, 11

*Research Team*

Name	Affiliation	Email	Type	Involvement	Citizenship/ Years in U.S.	FSRDC
Researcher #1	University #1	<a href="mailto:researcher1@university.edu">researcher1@university.edu</a>	PI	A	USA	Boston
Researcher #2	University #2	<a href="mailto:researcher2@university.edu">researcher2@university.edu</a>	CI	A	Brazil / 7 years	Chicago

## **Attachment 2: Predominant Purpose Statement**

### **Predominant Purpose Statement (PPS)**

Draft, [insert date here]

#### **Project # - “Project Title”**

#### **Researchers**

*List all Census Bureau Employees and all SSS, with their affiliation. This should be in two separate sections. Below the list of SSS, include a brief description of their role on the project.*

[Additional project supporters may be added to this project after it commences as needs dictate. In accordance with the *Census Bureau - IRS Criteria Agreement* of March 2018, the Census Bureau will notify IRS Statistics of Income of any additions and/or subtractions of Special Sworn Status individuals on a quarterly basis.]

#### **Duration – (In months)**

#### **Locations – (HQ and/or all RDC Locations)**

**Funding** – *(If an internal project, this is ‘Internal Project – Census Bureau’; if there are additional funding sources, list them here.)*

#### **Background**

*This section should provide an understanding of the research questions, why these questions are of importance to the Census Bureau, the methodology that will be used, and how FTI will be used for the project. This section can be used as a form of proposal, explaining the research, how it will be conducted, and its importance.*

#### **Criteria**

*List all relevant criteria here and an explanation of how this project will further the mission of the Census Bureau.*

#### **Requested Datasets**

#### **FTI Datasets**

[Only the following Federal Tax return Information (FTI) items from the following files will be needed by the researcher(s) to conduct this project.]

*List all FTI datasets and variables contained in them here.*

## **Census Bureau Datasets**

*List all non-FTI datasets here. Variable listings are not needed.*

## **User Provided Datasets**

*List all public use datasets here, along with a brief description or link to it.*

## **Need and Use of FTI**

*In this section, list each FTI variable, the datasets it is contained in, and the description/label for the variable. Below this information, for each variable, provide a description of how the variable will be used in the proposed project and why it is needed.*

*Ex.:*

<i>EIN</i>	<i>LBD, BR</i>	<i>Employer Identification Number</i>
<i>Used to identify establishments and link to external data.</i>		

## **Project Completion**

[The researcher(s) will provide the Census Bureau with annual progress reports on this multi-year project. After this project is completed, the researcher(s) will produce a Center for Economic Studies (CES) Working Paper and related research papers for conference presentation and/or journal publication. They will also produce a CES technical memorandum summarizing the key findings pertinent to Census Bureau Title 13, Chapter 5 programs. In addition, the researcher(s) will produce a Post Project Certification (PPC) summarizing whether and how the criteria described in the project proposal's Predominant Purpose Statement (PPS), as well as any others, were or were not met.]

## **Disclosure**

Public releases of all statistics by the Census Bureau conform to policies and rules described in IRS Publication 1075 (September 2016) Tax Information Security Guidelines for Federal, State and Local Agencies adhering to cell thresholds established for Federal Tax Information including, but not limited to, business and geographic classifications. Furthermore, as agreed to in the May 2018 Memorandum of Agreement between the Census Bureau and the IRS regarding Confidentiality and Disclosure Protection of Federal Tax Information, the Census Bureau will develop additional disclosure avoidance methodologies for data releases when it is not practical to apply the framework expressed in Publication 1075. These methodologies will be certified by the Director, Statistics of Income Division, Internal Revenue Service.

## **Project Timeline**

The proposed duration of the project is [X] years. [Additional documentation if needed].

*Insert table with timeline. The timeline should include activities such as cleaning data, merging data, conducting specific analyses, and estimated timeframe for outputs, such as publications, presentations, or any public use products. Tie the activities in the timeline back to the benefits.*

## **Attachment 3 - Benefits to Census Bureau Programs**

Proposals must provide benefits to data programs the Census Bureau conducts under Title 13 U.S. Code. The Census Bureau developed the 13 criteria below for assessing this potential. Researchers must consult with an FSRDC administrator to determine whether the data they propose to use contains Federal Tax Information (FTI). All business data and some kinds of household data contain FTI. Projects using FTI must demonstrate that the project's predominant purpose is to benefit Census Bureau programs by meeting at least one of criteria 5 through 13 below, which are the same criteria as listed in the Criteria Document of September 15, 2000. Projects using data that have no FTI must show that the project will provide benefits under one or more of criteria 1 through 13.

1. Evaluating concepts and practices underlying Census Bureau statistical data collection and dissemination practices, including consideration of continued relevance and appropriateness of past Census Bureau procedures to changing economic and social circumstances;
2. Analyzing demographic and social or economic processes that affect Census Bureau programs, especially those that evaluate or hold promise of improving the quality of products issued by the Census Bureau;
3. Developing means of increasing the utility of Census Bureau data for analyzing public programs, public policy, and/or demographic, economic, or social conditions; and
4. Conducting or facilitating census and survey data collection, processing or dissemination, including through activities such as administrative support, information technology support, program oversight, or auditing under appropriate legal authority.
5. Understanding and/or improving the quality of data produced through a Title 13, Chapter 5 survey, census, or estimate;
6. Leading to new or improved methodology to collect, measure, or tabulate a Title 13, Chapter 5 survey, census, or estimate;
7. Enhancing the data collected in a Title 13, Chapter 5 survey or census. For example: improving imputations for non-response; developing links across time or entities for data gathered in censuses and surveys authorized by Title 13, Chapter 5;
8. Identifying the limitations of, or improving, the Business Register, Master Address File, and industrial and geographical classification schemes used to collect the data;
9. Identifying shortcomings of current data, collection programs and/or documenting new data collection needs;
10. Constructing, verifying, or improving the sampling frame for a census or survey authorized under Title 13, Chapter 5;
11. Preparing estimates of population and characteristics of population as authorized under Title 13, Chapter 5;
12. Developing a methodology for estimating non-response to a census or survey authorized under Title 13, Chapter 5;
13. Developing statistical weights for a survey authorized under Title 13, Chapter 5.

## **Appendix A: FSRDC Data**

### **Census Datasets**

Center for Economic Studies (CES) Data Products: <https://census.gov/programs-surveys/ces/data.html>

Economic Data: <https://census.gov/programs-surveys/ces/data/restricted-use-data/economic-data.html>

Demographic Data: <https://census.gov/programs-surveys/ces/data/restricted-use-data/demographic-data.html>

Longitudinal Employer-Household Dynamics (LEHD) Data: <https://census.gov/programs-surveys/ces/data/restricted-use-data/lehd-data.html>

UMETRICS Data: <https://census.gov/programs-surveys/ces/data/restricted-use-data/umetrics-data.html>

### **Other Agencies' Datasets**

Agency for Health Care Research (AHRQ) Data:  
[http://meps.ahrq.gov/mepsweb/data\\_stats/onsite\\_datacenter.jsp](http://meps.ahrq.gov/mepsweb/data_stats/onsite_datacenter.jsp)

National Center for Healthcare Statistics (NCHS) Data:  
<http://www.cdc.gov/rdc/>  
<http://www.cdc.gov/rdc/B2AccessMod/ACs220.htm>

Bureau of Labor Statistics Data:  
<https://www.bls.gov/rda/>  
<https://www.bls.gov/rda/apply.htm>

Bureau of Economic Analysis Data:  
<https://www.bea.gov/research/special-sworn-researcher-program>

## Appendix B: Useful Links

### General

FSRDC Website: <http://www.census.gov/fsrdc>

CES Website: <https://www.census.gov/programs-surveys/ces.html>

FSRDC Locations and Contact Information:

<http://www.census.gov/about/adrm/fsrdc/contact.html>

IRS Criteria Document: [https://census.gov/content/dam/Census/programs-surveys/center-for-economic-studies/IRS\\_Criteria\\_Document.pdf](https://census.gov/content/dam/Census/programs-surveys/center-for-economic-studies/IRS_Criteria_Document.pdf)

CES Working Papers: <http://ideas.repec.org/s/cen/wpaper.html>

and older CES Working Papers: <https://www2.census.gov/ces/wp/>

### Census Benefits

Writing Benefits Statements: [https://census.gov/content/dam/Census/programs-surveys/center-for-economic-studies/Writing\\_Benefit\\_Statements.pdf](https://census.gov/content/dam/Census/programs-surveys/center-for-economic-studies/Writing_Benefit_Statements.pdf)

Potential FSRDC Methodological Topics

[https://community.isr.umich.edu/public/Portals/9/Docs/Potential\\_RDC\\_Methodological\\_Topics\\_20060906.pdf](https://community.isr.umich.edu/public/Portals/9/Docs/Potential_RDC_Methodological_Topics_20060906.pdf)

Data Collection, Analysis and Dissemination Process Topics (NSF-Census Solicitation):

[http://www.nsf.gov/pubs/2010/nsf10621/nsf10621.htm#pgm\\_desc\\_txt](http://www.nsf.gov/pubs/2010/nsf10621/nsf10621.htm#pgm_desc_txt)

Abowd & Stinson: <http://www.census.gov/sipp/workpapr/abowd-stinson-ME-ss-20110707-sehsdworking.pdf>

Census Research Reports and Studies: <https://www.census.gov/srd/csrreports/byseries.html>

Research Opportunities at the Census Bureau:

<http://www.census.gov/srd/research.pdf>

<https://www.census.gov/srd/CSRMGeneralResearchTopics2015.pdf>

### Some Questionnaires

*Firm and Establishment Census and Surveys*

[https://bhs.econ.census.gov/ec12/pages/formarchive\\_default.html](https://bhs.econ.census.gov/ec12/pages/formarchive_default.html)

<https://www.census.gov/programs-surveys/economic-census/data/tables.html>

*Person and Household Census and Surveys:*

American Community Survey (ACS)

<http://www.census.gov/programs-surveys/acs/>

2010 Decennial Census

[https://www.census.gov/programs-surveys/decennial-census/technical-documentation/questionnaires.2010\\_Census.html](https://www.census.gov/programs-surveys/decennial-census/technical-documentation/questionnaires.2010_Census.html)

2000 Decennial Census

<http://www.census.gov/dmd/www/2000quest.html>

1990 Decennial Census

<http://www.census.gov/prod/1/90dec/cph4/appdxe.pdf>

Survey of Income and Program Participation (SIPP)

<http://www.census.gov/programs-surveys/sipp/tech-documentation/questionnaires.html>

Current Population Survey (CPS)

<http://www.census.gov/cps/>

American Housing Survey (AHS)

<http://www.huduser.org/portal/datasets/ahs.html>