

CJARS Title 13 Proposal Guide

-

For Data Access Through the Federal Statistical Research Data Center Network

The CJARS Team*

University of Michigan & U.S. Census Bureau

November 17, 2022

Contents

1	Purpose of proposal guide	2
2	Guidelines for proposal development	2
2.1	Preliminary proposal development	2
2.2	Final proposal development	3
2.2.1	Abstract	4
2.2.2	Project description	4
2.2.3	Benefits statement	10
3	Proposal review process	14
3.1	CJARS review of FSRDC proposals	15
3.2	Census Bureau review process	15
4	After Census Bureau approval	17

*Prospective data users who have questions about this proposal guide may contact cjars-data-users@umich.edu.

1 Purpose of proposal guide

This proposal guide is a resource for researchers interested in using data from the Criminal Justice Administrative Records System (CJARS) through the Federal Statistical Research Data Center (FSRDC) network. CJARS data infrastructure has been developed with the goal of improving research on the U.S. criminal justice system in order to expand knowledge and inform policy and public administration of the system. The CJARS team would like to facilitate research proposal development and lower barriers to access where possible.

This document is a complement to more [general guides to proposing research projects in the FSRDCs](#).¹ Its content aligns with those guides, but places specific attention toward assisting researchers who will include CJARS data in their FSRDC proposals.

Section 101 of Title 13 of the U.S. Code authorizes the Census Bureau to collect and publish statistics on crime, the criminal justice system, and the justice-involved population. Accordingly, this guide will provide researchers with additional information about how they can incorporate and make use of criminal justice data in their proposals to satisfy the requirements for proposals submitted to the Census Bureau. The CJARS data infrastructure can be used to produce Title 13 statistical material and improve Census Bureau operations. This proposal guide highlights several key areas (e.g., benefits statements) where researchers can leverage their request for CJARS data to bolster their proposal to the Census Bureau.

2 Guidelines for proposal development

Before preparing an FSRDC proposal, researchers should familiarize themselves with the CJARS project and its data holdings to assess whether the data infrastructure will be able to support their research interests. Extensive information about the CJARS project and its data holdings are available in the [CJARS data documentation](#) and [data benchmarking report](#).²

Once researchers have some idea of what they would like to propose, they should arrange a meeting with the FSRDC administrator at the location where they will be conducting the research. In this preliminary discussion, the administrator will outline the proposal process and can answer specific questions about data availability and the proposal development and review processes.

2.1 Preliminary proposal development

The first step in proposal development is to reach out to your local FSRDC administrator to receive consultation on the proposal development process. [A list of all FSRDC locations can](#)

¹The full link for this guide is: https://www.census.gov/content/dam/Census/programs-surveys/center-for-economic-studies/Research_Proposal_Guidelines.pdf

²The full link to the CJARS data documentation is: <https://cjars.org/data-documentation-download/>. The full link to the CJARS data benchmarking report is: <https://cjars.org/benchmarking-report-download/>.

be found on the Census Bureau website.³ There are several topics that should be discussed with the FSRDC administrator including whether the:

1. Research fits the Census Bureau’s authorities under Title 13.
2. Research is feasible.
3. Research will provide benefits to Census Bureau programs under Title 13.
4. Output is designed to produce model-based statistical estimates.
5. Output will not pose undue risk of disclosing information about survey respondents.

FSRDC administrators may request a two- to three-page proposal overview including a summary of the research questions, the internal Census Bureau datasets requested, and a general idea of the project’s methodology and output.⁴

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 names of the principal investigators (PIs) 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 United States for at least three of the previous five years. Researchers that wish to work in the FSRDC that do not currently reside in the United States or have not lived in the United States for three 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-PI. The Census Bureau recommends that the advisor also apply for Special Sworn Status (see below).

2.2 Final proposal development

FSRDC proposals submitted to the Census Bureau must include the following three documents (each of which will be described in more detail below):

1. A project abstract,
2. A project description, and
3. A statement of the benefits to the Census Bureau.

These documents are submitted to the FSRDC system through your FSRDC administrator. It is expected that the proposal development process will involve working with the FSRDC administrator through multiple iterations of the proposal and the benefit statement (called the Predominant Purpose Statement [PPS]) 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 Bureau benefits.

³The full link to the list of FSRDC locations is: <https://www.census.gov/content/census/en/about/adrm/fsrdc/locations.html>

⁴A list of available Census Bureau data can be found here: https://www.census.gov/about/adrm/fsrdc/about/available_data.html.

2.2.1 Abstract

An abstract should be submitted as a standalone document. It should be 150-250 words in length, written in first 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, list the project title and the names of all researchers.

2.2.2 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 Bureau and non-Census Bureau datasets requested, 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. Although the benefits to the Census Bureau are presented in a separate proposal document (discussed in the next section), your proposal should discuss how your research benefits the Census Bureau.

Keep in mind that the audience for an FSRDC 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:

1. Contain a title and the names of all researchers at the top of the first page.
2. Include appropriate headings and subheadings throughout the document to assist reviewers in following the proposal narrative.
3. Not include any appendices unless approved in advance by the FSRDC administrator.
4. Be limited to no more than 15 single-spaced pages or 30 double-spaced pages including references.
5. Use a twelve-point font and one-inch margins. Times New Roman or Calibri font are preferred.
6. Include page numbers.
7. Be written in revisable form text: .rtf, .doc, or .docx format.⁵

The full proposal should contain the sections described below. 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.

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

⁵Some data partners request revisable form documents in order to perform electronic comparison in case they receive multiple versions.

of your research questions and a brief discussion of the relevant literature. A full literature review is not necessary, but some references are fine. Include a list of all internal Census Bureau datasets that your project will request, such as CJARS. Discuss why you need non-public data and briefly describe what you will do with the data. For example, if your proposal makes use of CJARS incarceration data, you could explain why the National Corrections Reporting Program (NCRP) is not sufficient to answer your research questions (e.g., inability to link NCRP data with other individual-level data).

The introduction should also state that the primary purpose of this research is to benefit the Census Bureau. It is also helpful, where appropriate, to connect the proposal research to activities and initiatives performed by the Census Bureau, such as satisfying responsibilities of 13 U.S.C. §101 (i.e., producing statistics about crime and the justice system). More details and examples of how to generate Census Bureau benefits using CJARS will be provided in Section 2.2.3.

All research must have a statistical purpose. In particular, Title 13 specifically prohibits the use of census micro-data for legal or regulatory purposes.⁶ Please do not include analysis in your proposal that has or can be perceived to have a regulatory purpose. Proposals should not make specific policy recommendations.⁷

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 a restricted Census Bureau dataset, you should discuss in detail how you will evaluate these variables and what types of results you will report to the Census Bureau.

The requested Census Bureau 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

⁶Title 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."

⁷In addition, for proposals using establishment and firm data, we recommend avoidance of the word "policy" altogether. For example, substitute "decision maker" for "policy maker."

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 CJARS to a decennial census, the method of this linkage should be described. Also, linking some data sets may lead to relatively small sample sizes (and thus disclosure risk) if there is not a significant amount of overlap between the two data sets. An example of this could be linking CJARS data (which includes individuals involved in the justice system) to survey data which is based on a sub-sample of the population, such as the American Community Survey (ACS). For demographic data (e.g., linking CJARS 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 not longitudinal, or linkable across demographic surveys. However, many demographic datasets can be linked together using Protected Identification Keys (PIKs). CJARS also contains this linkage variable. Ask your FSRDC administrator for details.

Record level linkage between researcher-provided data and demographic data must be performed by internal Census Bureau staff.⁸ 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 Bureau data, which is a violation of Title 13 protections.⁹ Data that fits into the CJARS data schema can be incorporated into CJARS through data donations made by researchers directly to the project.¹⁰ These records will first be processed at UM to prepare the records for transfer to the Census Bureau where they will ultimately be integrated into the CJARS data infrastructure that is available in the FSRDC system.

It is important to convey to proposal 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. Documentation on the CJARS data can be downloaded from the [CJARS website](#).¹¹ Other restricted-use data commonly used by approved project in the FSRDCs can be found on the [Census Bureau website](#).¹²

Often three sources of information about datasets are helpful: survey questionnaires, Census Bureau web pages detailing survey and sample design, and working papers that

⁸Within the Census Bureau, availability 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.

⁹Furthermore, for business data with FTI, a complete match results in a violation of the IRS "fact-of-filing" information protection.

¹⁰For more information about donating data to CJARS, please contact the CJARS staff at cjars-staff@umich.edu.

¹¹The full link to the CJARS website is <https://cjars.org>.

¹²The full link to the list of restricted-use Census Bureau data is <https://www.census.gov/programs-surveys/ces/data/restricted-use-data.html>.

researchers have written using the data. Making yourself familiar with all available documentation for a dataset you are requesting will help to ensure that the way you propose using the data in your research is in fact feasible, otherwise you could hinder the chances of receiving project approval. 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 individual, arrest, court case filing? Alternatively, will your unit of observation be at a more aggregated level (e.g., Department of Corrections)? 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) state or sub-state levels; if so, what levels? In addition, will you run separate models by gender, race, etc.? If you have external or public-use data, how large 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 FSRDC reviewers assess both the feasibility of your project and the risk of disclosing confidential information. This information will also be important for CJARS staff to assess whether your proposal aligns with the data use agreements that it has in place with agencies (see section 3.1 for more information on requirements for CJARS review requirements).

Write out the equations you plan to estimate. You do not need to give the exact functional form you plan to use, but provide as much detail as possible about 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.

Data sources

Data sources should be divided into two sections: "Census Provided Data" and "Researcher Provided Data." In the "Census Provided Data" section, specify the exact data you are requesting. For example, this would be the most recent vintage of CJARS data. More information about vintages can be found in the [CJARS data documentation](#).¹³

If a dataset is listed as a separate entity in the table, it needs to be listed separately in your proposal. As an example, "LEHD Data" is insufficient information, instead list the requested datasets individually (e.g., EHF, ICF, etc.). In regard to requesting CJARS, the CJARS data infrastructure should be thought of as a single dataset even though it is comprised of several relational tables. CJARS should be requested as a single entity.

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. An example is provided at the end of this section. If you are requesting years of

¹³The data documentation can be downloaded on the CJARS website: <https://cjars.org/data-documentation-download/>.

data not yet released, detail these with an "if available" caveat (see table at the end of this section for an example). The years of requested data need to be clearly justified, particularly for proposals requesting a long time series.

It is straightforward for the Census Bureau to provide the standard FSRDC datasets for an approved project. 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 (e.g., UCR or NIBRS), 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.¹⁴ Future requests for researcher-provided data not listed in your proposal may require additional justification and approval. Note: public-use Census Bureau 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).

Table 1: Example data request table with justification

Dataset	Years	Reason/Use
Criminal Justice Administrative Records System (CJARS)	2020q4; updated vintages as they become available	Determine individual terms of incarceration
LEHD Employment History File (EHF)	1985 to 2014; 2015 to 2024 as it becomes available	Identify employment histories

Project output and disclosure risk

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 rather than

¹⁴Outside data linkable at the individual-level would need to go through the Census Bureau's internal data linkage procedure. For example, data from a criminal justice agency that is outside of the scope of CJARS, but contains information needed to answer research questions (e.g., inmate intake risk assessment scores) would need to go through the Census Bureau's data linkage process.

requesting access to micro data at the FSRDC.¹⁵

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 interested in terms of incarceration for individuals incarcerated in a specific state? 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.¹⁶ 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.

There are requirements for proposals that include CJARS that researchers should be aware of. CJARS collects data through a number of different methods, such as: data use agreements, web scraping, public records requests, and data donations. Some of the data use agreements that CJARS has with agencies put stipulations on how the data can be used. For instance, some agencies request that they are given the opportunity to review research proposals that isolate the agencies data in analyses or output. For this reason, it is important for CJARS to know the specific geographies and type(s) of records (e.g., court case filings, terms of incarceration) which output will be released so that it can abide by the terms of all data use agreements that it has with agencies.

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. Please reach out to the FSRDC executive director of the location where you will be conducting the research to inquire about lab fees. Also inquire about how these fees can be covered, sponsoring institutions might cover these fees.

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 and up to the maximum of 60 months is recommended. All projects are required to submit a yearly progress report/interim benefits statement.

Project description checklist

Below is a checklist of items that should be covered in the project description of your research proposal. Please use this checklist to make sure that you cover all required aspects of the project description.

1. Introduction (1-3 pages)

¹⁵Demographic 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.

¹⁶If 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."

- (a) Overview of benefits to Census Bureau.
 - (b) Description of research questions, brief discussion of relevant literature and research.
 - (c) Basic overview of research plan, including a brief summary of all internal Census Bureau datasets requested.
2. Methodology (8-9 pages)
- (a) Overview of empirical methodology.
 - (b) Subsections describing detailed empirical specifications, estimation methodology, key variables, and the specifics of how the data will be used in estimation.
 - (c) Subsection describing methodology used in providing benefits to Census Bureau.
3. Data (1-3 pages)
- (a) Census provided data
 - i. List of all dataset, including years.
 - ii. Description of planned usage.
 - iii. Summary of justification and need for restricted use data.
 - (b) Researcher Provided data
 - i. List of all researcher-provided datasets, including years.
 - ii. Description of planned usage and need for researcher provided data.
 - (c) Data Linkages and Justification
 - i. Description of how data will be linked together.
4. Project Output and Disclosure Risk (1-3 pages)
- (a) Description of expected model-based output from the methodology.
 - (b) Description of expected supporting tabular output, if any.
 - (c) Description of products expected from the project including working papers and technical memos
 - (d) Discussion of any possible disclosure risks and mitigation strategy (disclosure risk and mitigation is explained further in section 3.2).
5. Project Duration and Funding Sources (less than 1 page)
- (a) Discussion of needed timeframe, including any travel requirements for accessing FSRDC and how those requirements may affect work on project.
 - (b) Statement of funding source.
6. References

2.2.3 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 their 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 censuses, 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 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 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.

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 Bureau 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 Bureau-produced publications, and further informs Census Bureau 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.¹⁷

Your access to the internal microdata is covered under legal statute because of your role as a consultant to assist the Census Bureau with its work.¹⁸ 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. With the CJARS data infrastructure, researchers can generate benefits for the Census Bureau that cover the justice-involved and their outcomes.

A good Census benefit accomplishes three elements: (i) a connection to Census Bureau program activity, (ii) articulates a specific analysis exercise, its purpose and approach, that can be clearly understood by reviewers and Census Bureau 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 the Census Bureau is working on or potentially thinking about strengthens its relevance. Census Bureau program staff describe much about their activities on the Census Bureau web pages, easily accessible through searches.¹⁹ Note that contributions that benefit other researchers, including Census Bureau 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 Bureau activities.

¹⁷For some federal agencies, including the IRS, the PPS is the document and project scope they review and approve.

¹⁸Title 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."

¹⁹Proposals requesting LEHD data are expected to include benefit to the LEHD program.

The PPS document is formatted in a specific way, and an [example draft](#) is available (remember that you should not begin preparing documents until you have met with your FSRDC administrator).²⁰ In the section entitled "Criteria", specify which of the 13 benefit criteria listed above are relevant to your project. Then provide a description under each relevant criterion of 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. For CJARS, as an example, you would state "Criminal Justice Administrative Records System (CJARS) 2020q4; updated vintages as they become available."

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.

To help you write benefits statements, the table below includes examples of how CJARS data could be leveraged to provide Criteria 5, 9, and 11 benefits to Census. Please note that this is a non-exclusive list and there may be other criteria/census benefits that are better suited for your project. Also, we are providing the examples below to aid with your application and do not make any guarantees that they will be sufficient as written for your specific project. Benefit statements should be tailored to the specific project you are proposing.

²⁰An example draft of a Predominant Purpose Statement document can be found on the Census Bureau website: https://www.census.gov/content/dam/Census/programs-surveys/center-for-economic-studies/PPS_Template.rtf.

Table 2: Example PPS benefits statements

Criterion	Census benefit example
Criterion 5	<p>There is a substantial amount of research effort being put into understanding the impact of contact with the justice system on outcomes that extend beyond those related to criminal justice involvement (e.g., enrollment in social safety net programs or employment status). Researchers interested in pursuing this line of inquiry could satisfy Criterion 5 while conducting their research by performing logical checks to understand the extent to which enrollment in social safety net programs overlaps with supervision by the justice system that would make being enrolled impossible. For example, it would be logically inconsistent if an individual was incarcerated between 2010 and 2017 according to CJARS data, but received Temporary Assistance for Needy Families (TANF) benefits in 2013. This could help to elucidate data quality issues that could improve a Title 13 program. And since CJARS itself is a Title 13 program, proposals can satisfy Criterion 5 by generating analysis that improves the quality of current and future vintages of CJARS itself.</p>
Criterion 9	<p>Recent federal legislation has been passed (i.e., First Step Act) that is intended to reduce the scope of the justice system in the United States, and particularly the population of incarcerated individuals. Increased numbers of prisons releasees will pose a challenge to sample-based Census Bureau programs such as the American Community Survey (ACS). Operations of sample-based programs would be improved by research that identifies specific subgroups and geographic areas for which measurement challenges are likely to be greatest.</p>
Criterion 11	<p>Inequities in the justice system is a topic at the forefront of research and policy discussions on the justice system. Researchers could satisfy Criterion 11 by generating estimates of justice-involved populations by leveraging linkages between CJARS data and other Census Bureau data could generate novel estimates. For example, CJARS data could be linked to LEHD data to understand the nature of employment gaps that occur after arrests and how this may vary across individual characteristics (e.g., race/ethnicity, gender, geography).</p>

3 Proposal review process

All research proposals that are submitted are evaluated by the Census Bureau on several criteria. It is important to note that when proposals request CJARS data, the proposals must also be reviewed by the CJARS team to ensure compliance with all data use agreements

that CJARS has with agencies from which it receives data. More about each of these review processes is described in the following sections.

3.1 CJARS review of FSRDC proposals

The CJARS project acquires data through many different methods, such as data-use agreements, public information requests, web scraping, and data donations. As discussed above, there are some data use agreements that require CJARS to give an agency the opportunity to review research proposals that will isolate the agency's data in analyses or output. The Census Bureau forwards requests to use CJARS data to the University of Michigan. The CJARS team will examine whether the proposals must be forwarded to CJARS data providers for further review. The questions that must be addressed in a proposal in order for the CJARS team to assess whether or not proposals must be forwarded to agencies for approval include the following:

1. What is the minimum level of geography for tabulations?
2. What is the minimum level of geography for modeled parameters?
3. Will agencies be identified in any output?
4. Will any steps be taken to minimize identification of agencies beyond those required by the Census Bureau Disclosure Review Board?

Each of these questions must be addressed in proposals that request CJARS data.

3.2 Census Bureau 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: (1) benefits to Census, (2) scientific merit, (3) need for non-public data, (4) feasibility, and (5) confidentiality protection and disclosure avoidance.

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 Bureau 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.

1. The Census Bureau is required by law to protect the confidentiality of data collected under its authorizing legislation, Title 13, U.S. Code.
2. Tabular and graphical output presents a higher risk of disclosure of confidential information than do coefficients from statistical models and is therefore strongly discouraged.
3. 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.
4. Researchers may be required to sign non-disclosure documents of survey sponsors or other agencies that provide data for their research projects.

There are several ways proposals are evaluated for disclosure risk. To evaluate disclosure risk, the proposed methodology and output must be clearly stated in the proposal. It is

understood that, because you are conducting research, you probably do not know all the details in advance. Nevertheless, there needs to be enough detail to assess whether the proposed project can succeed without posing undue risk. In particular:

1. It has to be possible to determine whether there may be 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, there needs to be 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 agencies, or groups of these); and the groups (levels) for which you will request output (e.g., criminal justice event type, gender, geography, and possible crosses of these). For example, detailed arrest-by-geography cells can become thin at a small enough geographical unit of analyses.
2. It has to be possible to determine whether variables included in models are discrete or continuous; both are allowable, but discrete variables (especially dummy (0, 1) variables) define "cells" that must be taken into account.
3. It has to be possible to determine whether you will request tables, graphs, or maps, which can be especially problematic.
4. It has to be possible to determine whether you plan to produce sub-state estimates.
5. Proposals will be evaluated for types of output that are relatively unfamiliar to some, such as from new statistical techniques. It must be possible to assess their disclosure risk, and insights you may have on this is welcomed.

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.

4 After Census Bureau approval

Following Census Bureau 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 been given approval by the Census Bureau, 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.

The review process is both lengthy and rigorous, requiring researchers to exhibit patience throughout. 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 data product that the CJARS project has developed and built and are grateful for your interest.