California Military Economic Impact Study Process Guide

Britnee Pannell & Sumeet Bedi

2022-01-20

Contents

Chapter 1

Introduction

The California Military Economic Impact Study report series began in 2018, when the California Governor's Office of Planning & Research and the Governor's Military Council requested the California Research Bureau (CRB, a unit of the California State Library) to conduct this study. This report provides detailed statewide and localized economic impacts of federal national security activity in the state of California. The federal agencies identified as relevant to national security includes the Departments of Defense (DOD), Homeland Security (DHS), Veterans Affairs (VA), and specified sub-agencies of The Department of Energy (DOE). The type of economic activity detailed includes spending (contracts, grants, veterans' benefits, and SmartPay charge card) and employment data (civilian and military). The first report, published in August 2018, utilized federal fiscal year 2016 data, while the second report, published in December 2019, used fiscal year 2018 data.

Following these 2 reports, the Office of Planning & Research and the CRB secured grant funding from the DOD to support two full-time equivalent (FTE) positions to develop supplemental reports that help localize the economic impacts detailed in the statewide report. In December 2020, the CRB produced the 2020 California Military Economic Impact Study, and followed this third edition of the statewide report with 2 first-time supplements that discuss the economic impacts in every county and congressional district in California. In December 2021, the Research Bureau completed the fourth version of the statewide report as well as the second edition of the county and congressional district supplements. After completion of the second round of supplements, the CRB was tasked with producing this process guide document.

This process guide and supporting documentation were developed in order to allow other states to replicate the methodology of this study for their respective geography of interest. Additionally, this documentation serves to provide the rationale behind how the data was gathered, wrangled, and analyzed in order to justify the conclusions in our main reports.

Chapter 2

Software Requirements for Project

To recreate California's study or perform studies in additional areas, the software used to obtain and process that data is needed. Fortunately all software is free and available online.(Where possible, make sure to use the most recent and fully updated versions of the software).

1. The R coding language from cloud.r-project.org. This language is used to obtain and process data.
2. RStudio Desktop from rstudio.com. RStudio is the integrated development environment (IDE) used to run R scripts and develop code.
3. Git from git-scm.com. A comprehensive guide to installing Git is available at happygitwithr.com. Git allows version control of edits across a multi-person team of researchers.
4. If a Github account is needed, one can sign up and register at https://github.com/. Individual free plans are available, as well as free upgrades for qualifying academic purposes.Github is used to develop and host this project.

The next section details data necessary to complete this project as well as how to obtain it.

Chapter 3

How to Obtain Necessary Data

This section details how to obtain each data type, whose categories are defined in the "requirements" section

3.0.1 Obtainable With Code

Unfortunately, there is no simple way to obtain employment data with code.

3.0.2 Manual Data Retrieval

Employment data was obtained from 2 websites - 1 for civilian employment figures, and 1 for military employment. There is no guarantee that these websites will exist in this form indefinitely, and care will be taken to keep this document as up to date as possible.

- Civilian employment: Obtained from the federal OFfice of Personnel Management's (OPM) FedScope website, which provides federal workforce data. Given that the report works based on federal fiscal years, the employment data of interest is of Quarter 3 of a given year (i.e. September). From this site, one can grab the civilian employment numbers for the Departments of Defense (which is an aggregate of the Air Force, Army, Defense, and Navy cabinet agencies), Homeland Security, Veterans Affairs, and Energy.
- Military employment: Obtained from the Defense Manpower Data Center (DMDC) website, which serves under the Office of the Secretary of

Defense. Similar to civilian employment, the data of interest is from September of a given year. The data referenced is the "Military and Civilian Personnel by Service/Agency by State/Country" Excel file, available under DoD Personnel, Workforce Reports & Publications on the DMDC website.

3.1 Spending Data

3.1.1 Code to Obtain Spending Data

3.1.2 Spending Data Obtained via FOIA

3.2 Raw Data Provided by

Include details on Data provided in the "data/raw" folder in the code repo- and the justifications on why it was included and not others

Here is where things get a little annoying - Each file used, with an explanation of where to get it and how to navigate the sites used for obtaining data - What information each file provides - Detailed information on how to make a file custom to specific data needs (where applicable) - Any notes on how files may differ according to region and individual project goals

Chapter 4

Methods

The following section details how to use the data and R code provided as well as an explanation of how the code works.

4.1 Process Outline

The over all process for this project is as follows:

- Data was Obtained
- Data was Filtered for relevance.
- Errors in data were for checked and repaired where found.
- Data was formatted for use in IMPLAN.
- Data was run through IMPLAN.
- IMPLAN outputs were graphically displayed and distributed via report.

4.2 Process in Detail

- Obtain data
  - Spending data
    - Grants
    - Contracts
    - SmartPay (FOIA Required)
  - Employment Data
    - Military Personnel
    - Civilian Employment

DMDC- download and parse csv FedScope- Initially have users go to website and save values of interest to a separate csv file to use in IMPLAN later Eventually set up a code to generate these values acording to state and national level based on parsing out the download from the site.

  - NAICS to IMPLAN crosswalks
  - Spreadsheets provided to aid processing and format outputs

- Process data

  - Clean contracts and grant data-
  - Clean spending data
  - Error check contract spending data

  Will need to go into detail about changes in the code between this year (2021) and subsequent years

  More detailed mention of how the error checking of the USASpending.gov contract data is needed A detailed walk through of how to manually check data and use the multiple NAICS to IMPLAN crosswalks to catch data Mention how IMPLAN automatically removes any codes having to do with construction so those have to be manually coded

  Some errors occur due to the transaction not being given a NAICS code, those need to be manually fixed

  Issues occur with NAICS codes that apply to multiple IMPLAN codes- give detailed explanation of how this was worked around and data was processed and added back to the main cleaned data.

- Run Data Through IMPLAN At this point we stop giving details to users about subsequent processes- we are not responsible for teaching users how to use IMPLAN. We should go over the general steps in what we did next to process data from IMPLAN, and how it was displayed graphically to educate the customers and summarize results for easier understanding

Chapter 5

Using IMPLAN

Place holder for section on how to enter the output files into implan and what IMPLAN analysis was run so that our study can be repeated.

Should not go into too much detail, as full instructions on how IMPLAN works is outside of the scope of this process guide.

Once the IMPLAN activity sheets for your desired geographies have been produced, it is time to upload them into IMPLAN and allow the tool to run its analysis. When you log into IMPLAN, the homepage should display 3 boxes to navigate through: Regions, Impacts, and Projects.

- **Regions:** Click on Regions to begin your analysis. This should open up a map of the United States. Select the region of interest for the analysis and give this project a name. After saving the project, move on to the Impacts tab.
- **Impacts:** In the Impacts tab, look for the button to the right of the "Save" button. Click on that, and select "upload event template". From there, find the activity sheet for the geography of interest and select that for upload. Ensure that "Industry Change Event" and "Household Spending" can be selected, and choose those for uploading into IMPLAN. The data points for the respective IMPLAN sectors should then auto-populate on IMPLAN. If your analysis is including SmartPay spending data, be sure to manually add that in as type "Institutional Spending", specification 11001, and the value amount. Choose to select all events (to the left of the "Save" button) and drag them over to the geography's model on the right side of the IMPLAN screen. After that, select to run the model at the bottom right. Depending on how big the model, IMPLAN will take time to generate the results. Once the model is done with its analysis, click on the "View Results" button.

Chapter 6

Conclusion/ Discussion

- Importance of modern techniques to get more efficient data analysis in a timely fashion
- Other closing remarks

Pitfalls, how re factoring code caught some of them. How this process will result in a more robust study over subsequent years

Government spending data is very difficult to obtain and there is not a lot of good documentation to help lay people use this data Hopefully we provide some guidelines and aid in discovering and processing this data so that quality studies can come about and Government spending can become more transparent.

Feel good hopeful stuff next.

Chapter 7

What's Next?

A section on where we hope to add and develop this process. Potentially the section to outline changes to the code we have already made for the upcoming 2021 report.

Chapter 8

License your GitBook

In the spirit of Open Science, it is good to think about making your course materials Open Source. That means that other people can use them. In principle, if you publish materials online without license information, you hold the copyright to those materials. If you want them to be Open Source, you must include a license. It is not always obvious what license to choose.

The Creative Commons licenses are typically suitable for course materials. This GitBook, for example, is licensed under CC-BY 4.0. That means you can use and remix it as you like, but you must credit the original source.

If your project is more focused on software or source code, consider using the GNU GPL v3 license instead.

You can find more information about the Creative Commons Licenses here. Specific licenses that might be useful are:

- CC0 ("No Rights Reserved"), everybody can do what they want with your work.
- CC-BY 4.0 ("Attribution"), everybody can do what they want with your work, but they must credit you. Note that this license may not be suitable for software or source code!

For compatibility between CC and GNU licenses, see this FAQ.