Connecting Federal and State Financial Data
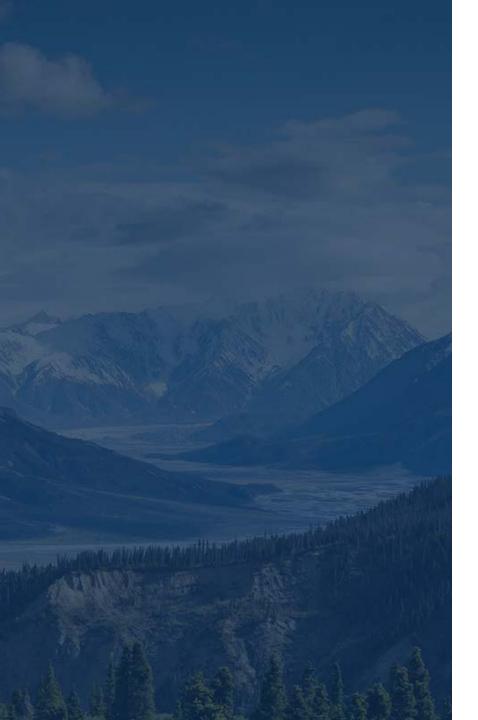
USASPENDING.GOV PILOT WITH THE STATE OF CONNECTICUT

AUGUST 2017

OVERVIEW

1. OBJECTIVE
2. METHOD
3. ROADBLOCKS
4. SOLUTION
5. CONCLUSION

OBJECTIVE AND APPROACH

Objective

To test the accuracy of connecting USAspending data with state-level financial data from Connecticut.

Problem

USAspending and Connecticut financial data do not contain a unique identifier that would allow joining at the award level.

Approach

As a result, we identified the common data between the two datasets that would produce the lowest level of granularity.

▸ USAspending contains a CFDA number, which is a unique number assigned to each assistance program administered by a department within the federal government. Connecticut data contains a SID number, which is a Special Identification Code assigned by Connecticut's Central Accounting for each specific appropriation.

▸ There are multiple SIDs per CFDA, so a one-to-one match was not possible. However, we were able to aggregate SIDs within the same CFDA family to compare Connecticut's equivalent of a CFDA to USG's CFDA using total dollars spent as a comparison.

SID	SID Description	CFDA	CFDA Description
20423	Control Of West Nile Virus	10001	AGRICULTURAL RESEARCH - BASIC & APPLIED RESEARCH
20442	Evaluation Arugula Accessions	10001	AGRICULTURAL RESEARCH - BASIC & APPLIED RESEARCH
20463	Aquatic Weed Management	10001	AGRICULTURAL RESEARCH - BASIC & APPLIED RESEARCH
22697	Chemical Analysis - Honey Bees	10001	AGRICULTURAL RESEARCH - BASIC & APPLIED RESEARCH

MATCHING PROCESS

USAspending data

Grants, loans, and other financial assistance from July 1, 2015 – June 30, 2016 with Connecticut as the recipient state

▸ We start with 22,768 records

▸ filtering by (1) award id's that contain a new award in that time period results in 11,476 records

▸ dropping records where there is no federal award id results in 10,068 records

▸ bringing in data from July 1, 2016 – present for any modifications to new awards in FY 2016 results in 2,287 additional records for a new total of 12,355 records

▸ filtering by (2) CFDA programs that appear in Connecticut data results in 8,418 records

▸ filtering by (3) recipient type results in 1,096 records

▸ The total amount of obligations is $6,139,429,178.25

Connecticut data

▸ We start with 242,535 records

▸ filtering by (1) records where the Budget Reference column is 2016 results in 121,787 records

▸ filtering by (2) CFDA program that appears in USAspending data results in 64,601 records

▸ The total amount paid out is $5,423,119,397.14

MATCHING PROCESS

	USG	CT	Pct diff
Medical Assistance Program (Medicaid)	$4,614,183,523	$4,079,254,082	-11.6%
Highway Planning and Construction	$320,311,370	$489,574,311	52.8%
Foster Care – Title IV-E	$64,658,796	$558,495	-99.1%
Unemployment Insurance	$55,849,229	$8,575	-99.9%
Children's Health Insurance Program	$74,017,713	$28,260,542	-61.8%
Child Support Enforcement Educational Agencies	$42,852,898	$319,581	-99.2%
SNAP State Administrative Match	$48,656,363	$7,027,001	-85.5%

ROADBLOCK: DATA QUALITY

Awards to the state are identified through USAspending data by the recipient type variable, but not all state agencies are correctly labeled as a state government recipient type.

Solution: Our initial filter to include only state government as a recipient type led to 1,022 records. However, a manual review of the remaining records not classified as state government led to 74 instances where a state agency was not correctly labeled as a state government recipient type, based on a name indicating the recipient was actually part of the state government (i.e. "Department of"). This workaround involved record-by-record review, and is not necessarily replicable in an automated analysis.

A sample of the recipient names not classified as state government recipients:

```
"STATE OF CT DEPT EMER MGMT HOME SEC"
"Public Health, Connecticut Department of"
"STATE OF CONNECTICUT"
"Connecticut State Library"
"DEPARTMENT OF EMERGENCY SERVICES"
"PUBLIC HEALTH, CONNECTICUT DEPARTMENT OF"
"CONNECTICUT OFFICE OF EARLY CHILDHOOD"
"CONNECTICUT"
"STATE OF CONNECTICUT, DEPARTMENT OF HOUSING"
"MENTAL HEALTH AND ADDICTION SERVICES, CONNECTICUT DEPARTMENT OF"
"ECONOMIC AND COMMUNITY DEVELOPMENT, CONNECTICUT DEPARTMENT OF"
"DEPARTMENT OF EMERGENCY SERVICES AND PUBLIC P"
"Transportation Connecticut Department of"
"TRANSPORTATION, CONNECTICUT DEPARTMENT OF"
"CONNECTICUT DEPARTMENT OF TRANSPORTATION"
"SOCIAL SERVICES, CONNECTICUT DEPARTMENT OF"
"AGING, CONNECTICUT DEPARTMENT ON"
```

ROADBLOCK: TIME DISCREPANCIES

Varying Program Start Dates: Many CFDAs distribute awards over multiple years, and it is difficult to track whether a payment captured by Connecticut reflects this year's award or the prior year's award.

Solution: We addressed this by using USAspending data that matched Connecticut's FY 2016 based on their budget reference column and isolating new federal awards that began in FY 2016 and any modifications to those awards in subsequent years. While some records did not have the budget reference field completed, those records were all in FY 2017 or FY 2018, so this was a reliable filter for this process but would not be usable in later analyses.

Example: This program awards funds on July 1 and October 1 of the following fiscal year and funds remain available for an additional fiscal year for obligation and expenditure.

	USG	CT	Pct diff
Title I Grants to Local Educational Agencies	$ 116,021,685	$ 112,418,181	3.1%

Time lag: A second issue with spending over multiple years is that aggregating dollar amounts from Connecticut data will often produce a lower dollar amount, reflecting the portion of the total award that has been disbursed. Therefore, comparing dollar amounts will not be a reliable mechanism for comparing state and federal data until the obligation period has ended.

CONCLUSION

▸ In order to properly match state and federal data, both systems require a unique identifier to reliably link data about awards to payments about those awards.

  ▸ The most efficient solution to this challenge would be for states to capture a unique award ID in their financial management systems.

  ▸ This would allow matching to show final recipients of federal funds, and to allow the public to calculate how quickly states are spending against their federal awards.