CPI for Diesel

Measuring Price Change for Diesel Fuel Using Alternative Data

Sarah Niedergall and Alec Harkins
Economists
DCPPI/OPLC

2024 FCSM Research and Policy Conference

October 23, 2024

This presentation reports on the results of ongoing research and analysis undertaken by Bureau of Labor Statistics staff. It has undergone more limited review than official publications.

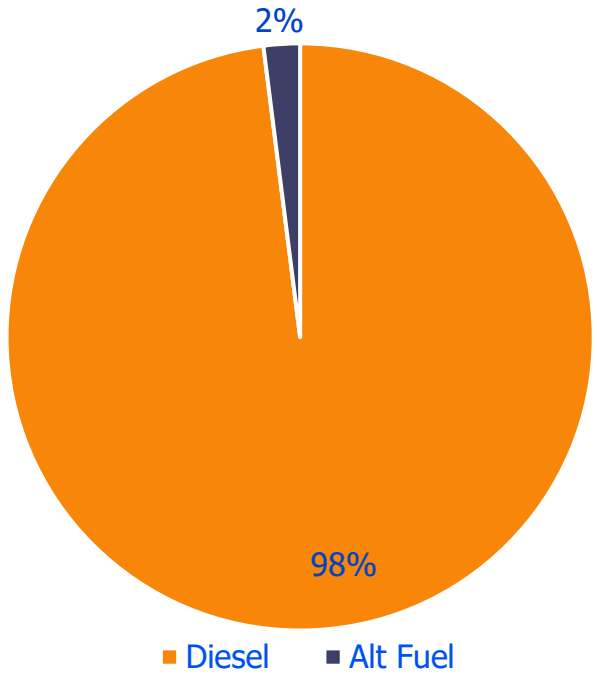
Crowd-source (Corp5) data source

- CORP5 is an online crowdsourced aggregator of motor fuel price data, with a database of over 100,000 stations across the U.S.
  - Average of 205,000 reported price observations every day
  - Roughly 6.23 million prices every month
  - Prices are updated in real time by app-users and station owners
- CORP5 data includes prices for gasoline, diesel, and E85
  - In June 2021, the CPI replaced the data collected by the BLS in the gasoline index with data from Corp5.
- BLS works with CORP5 to continue to obtain daily deliveries of their data

Importance of motor fuels in the CPI

Item	Relative importance*
Transportation	15.898
Motor Fuel	3.372
Gasoline (all types)	3.261
Other motor fuels	0.111

* December 2023 (2022 expenditure weights)



Data comparison of diesel sources

	Survey data	Crowd-sourced data
Frequency	Monthly	Daily
Number of retail stations	715 stations/month	32,385 stations/month
Number of price observations	760 price quotes/month	1.3 million observations/month

Benefits of CORP5

- Permits the use of substantially more observations with broader coverage of the CPI geography.
- Improved representativeness of reference month
  - Weekends and holidays now included
  - Price data are available daily, leading to improved estimates of average price levels for each calendar month.
- Reduces BLS data collectors' and survey respondents' burden
- Opportunity to improve review and analysis of input data

Corp5 hypothetical raw data

Station ID	Average price	Fuel type	Zip code	State
1	4.24	Diesel	54566	Wisconsin
2	3.55	Diesel	48002	Michigan
3	3.48	Diesel	35005	Alabama
4	4.59	Diesel	46011	Indiana
5	6.55	Diesel	53227	Wisconsin
6	3.49	Diesel	84101	Utah
7	3.43	Diesel	98373	Washington
8	2.48	Diesel	04403	Maine
9	3.09	Diesel	06001	Connecticut
10	3.29	Diesel	02801	Rhode Island
11	3.19	Diesel	43106	Ohio
12	4.55	Diesel	20602	Maryland
13	4.19	Diesel	20124	Virginia

Calculation of station relatives

- Monthly Station Price Relatives
  - Month-to-month change of same station ID
  - No item replacement
  - No imputation when a station ID is unavailable
    - Reasonable strategy since there is a large enough sample
    - Further research will look into imputation

Calculation of county and CPI area relatives

1. Calculate geometric mean of the station relatives in each county (Jevons Relative)
2. Allocate CPI area diesel expenditure weights to counties based on their Census Urban population estimate
3. Calculate CPI area relative using weighted county relatives:
   i. weight ratio = county weight / sum of county weights
   ii. Geomeans price relative component of each county = Jevons relative ^ weight ratio

Hypothetical Allocated county-level expenditure weights

Example of CPI area with three counties

County codes	Urban population (in millions)	Allocated expenditure weight ($, in millions)	County population share	Jevons relative	Geomeans relative price component
10000	1.86	62	0.31	1.30	1.09
20000	2.7	90	0.45	1.16	1.07
30000	1.44	48	0.24	0.98	1.00
Total	**6**	**200**	**1.00**	--	--

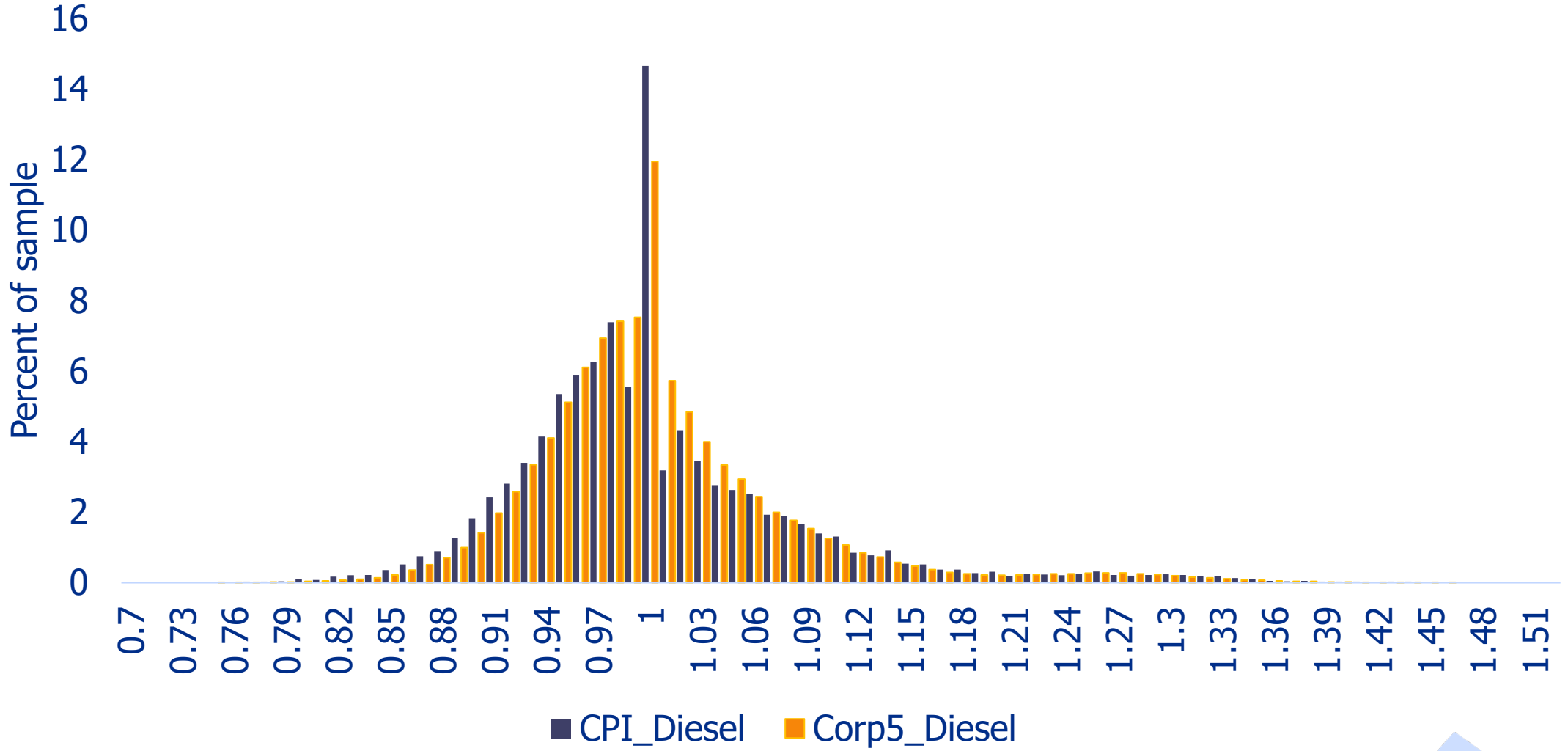
Calculation of item/area relatives

$$CPI\ Area\ Relative_t = \prod_{n=1}^{n} County\ Jevons\ Relative_t^{Weight\ ratio}$$

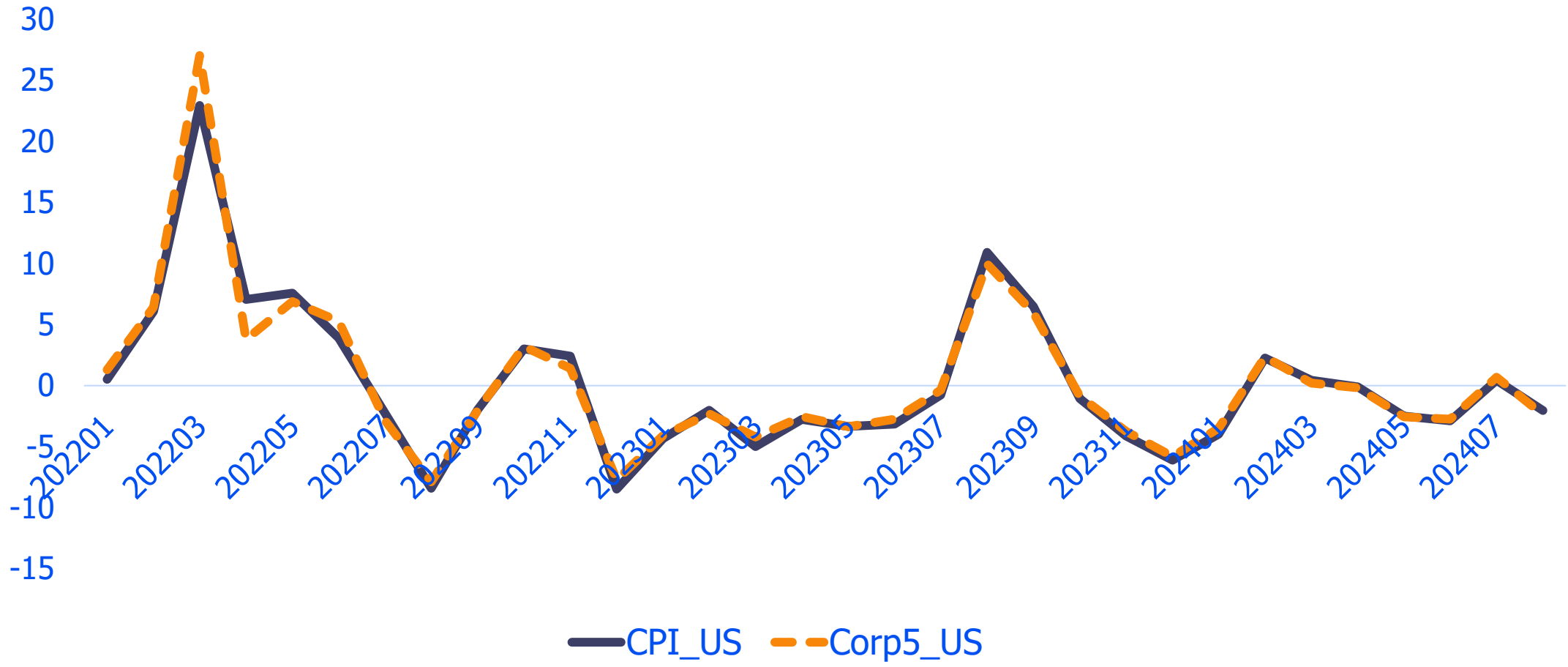
Imputation

- Missing **STATION** prices are IMPLICITLY imputed based on the COUNTY relative (excluded from COUNTY relative calculation)
- Missing **COUNTY** relatives are IMPLICITLY imputed (excluded from AREA relative calculation)
- Missing **AREA** relatives are EXPLICITLY imputed using CPI area imputation hierarchy

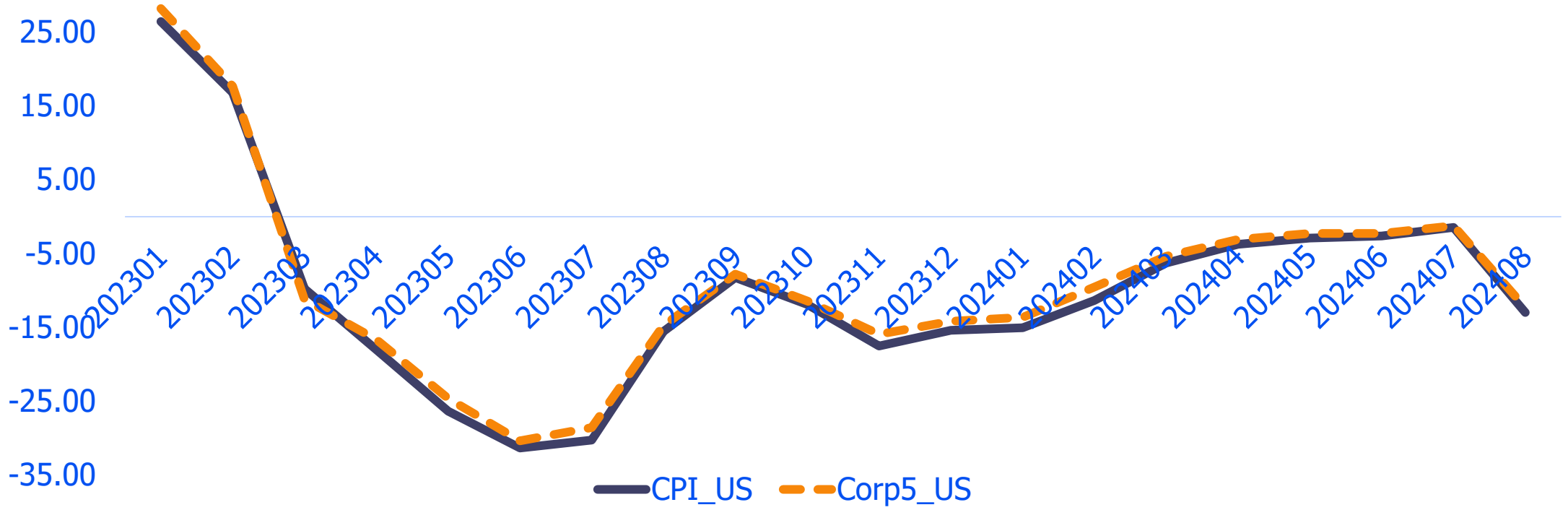
Distribution of 1-month price change
January 2022 – August 2024



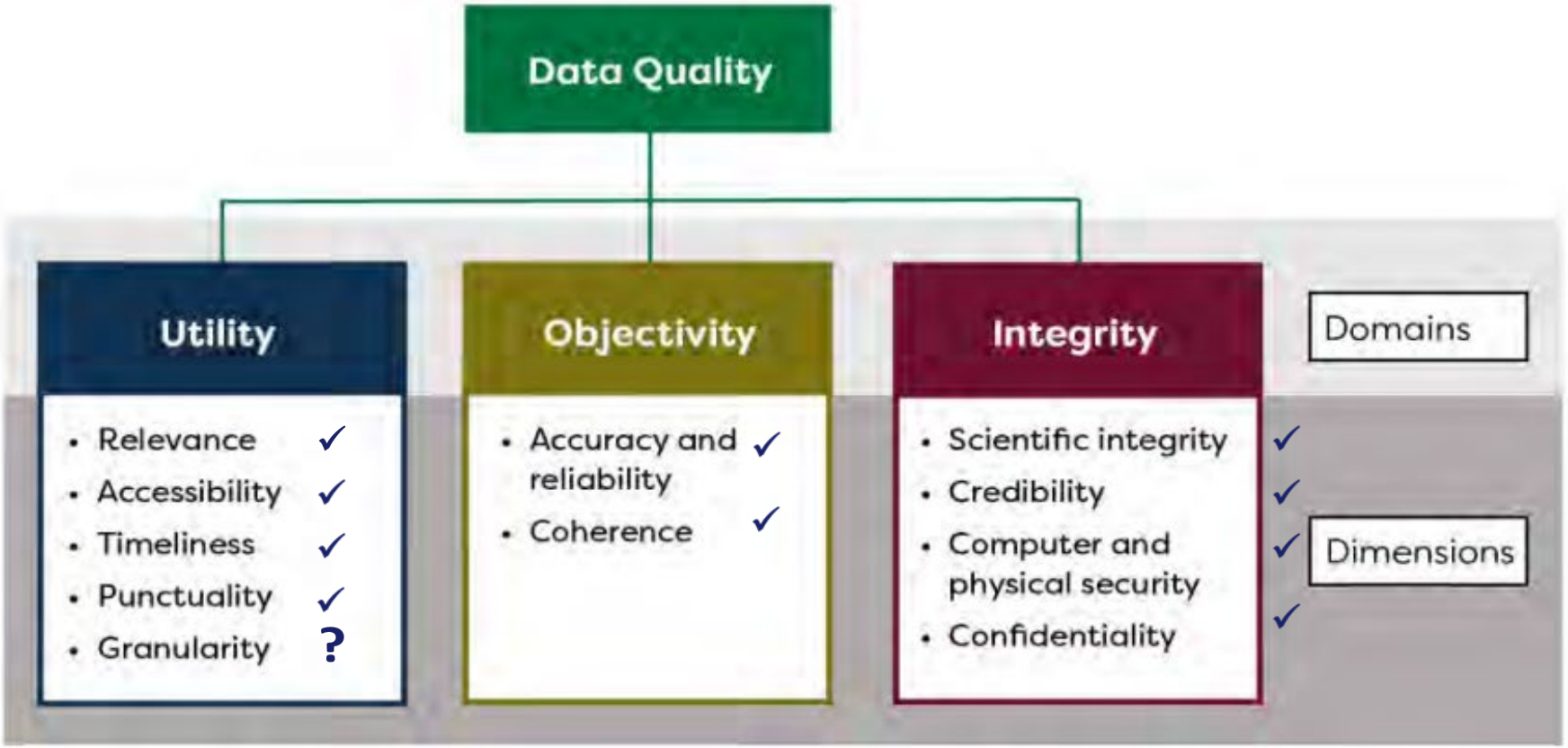
Experimental diesel 1-month percent change using CPI and Corp5 data



Experimental diesel 12-month percent change using CPI and Corp5 data



FCSM framework for data quality

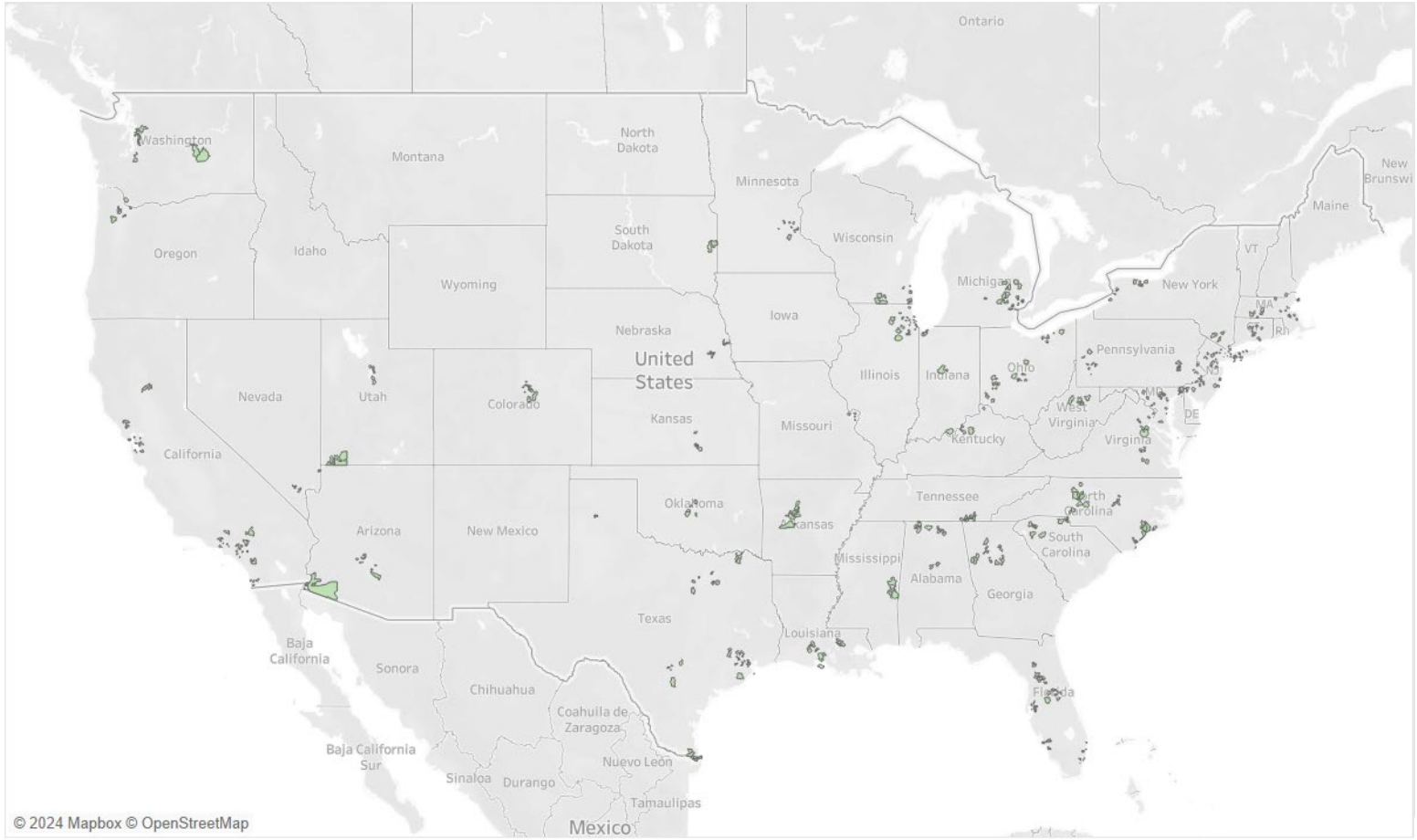


Federal Committee on Statistical Methodology. 2020. *A Framework for Data Quality*. FCSM 20-04, September 2020.

CPI and Corp5 granularity

Data element	CPI	Corp5
Unique station identifier	✓	✓
Location	✓	✓
Price	✓	✓

Geographic Coverage



Map based on Longitude (generated) and Latitude (generated). Color shows count of Outlet Cd. Details are shown for Outlet Zip.

Next steps

- Calculation of expenditure weights for diesel and alternative fuel (E15/E85, biodiesel, EV charging stations)
- Imputation of missing station relatives
- Average prices
- Variance estimates

Item	Relative importance*
Transportation	15.898
Motor Fuel	3.372
Gasoline (all types)	3.261
Other motor fuels	0.111

* December 2023 (2022 expenditure weights)

Contact Information

Sarah Niedergall and Alec Harkins
DCPPI/OPLC
www.bls.gov/cpi
niedergall.sarah@bls.gov
harkins.alexander@bls.gov