Appendix D:
Statistical Methodology of Estimating Petroleum Exports Using Data from U.S. Customs and Border

Updated June 1, 2023

Independent Statistics and Analysis
U.S. Energy Information Administration

www.eia.gov
U.S. Department of Energy
Washington, DC 20585

The U.S. Energy Information Administration (EIA), the statistical and analytical agency within the U.S. Department of Energy (DOE), prepared this report. By law, our data, analyses, and forecasts are independent of approval by any other officer or employee of the U.S. Government. The views in this report do not represent those of DOE or any other federal agencies.

Table of Contents

Introduction

On August 31, 2016, we started incorporating export data collected by U.S. Customs and Border Protection (CBP) through the Automated Commercial Environment (ACE) system to help estimate U.S. petroleum exports reported in the *Weekly Petroleum Status Report* (WPSR). The updated WPSR methodology provides a more accurate estimate of U.S. petroleum exports to better inform our weekly estimates of domestic quantity demanded for petroleum products and biofuels and the quantity of the crude oil supply adjustment (that is, the balance quantity for crude oil).

Exports are an important component for calculating our weekly estimates of quantity demanded for petroleum products. We calculate domestic quantity demanded (measured as product supplied) for petroleum products as:

Domestic quantity demanded = Production + Imports - Stock change - Exports

We conduct weekly surveys to collect the data that we use to estimate production, imports, and stocks but not exports. The U.S. Census Bureau (USCB) publishes export data once a month. To estimate weekly exports, we previously used forecasts based on an autoregressive integrated moving average (ARIMA) procedure to model weekly export statistics using the latest USCB monthly data. We obtained official USCB export data for crude oil, petroleum products, and biofuels each month, approximately six weeks after the close of the reporting month, and we published them in our *Petroleum Supply Monthly* (PSM). These data provided the inputs to the weekly export ARIMA model. However, as the volume and variability of U.S. exports increased, we sometimes found significant discrepancies between modeled weekly estimates and final monthly exports data provided by the USCB. Because of the growing impact of exports on our weekly estimates of U.S. petroleum quantity demanded, we needed to improve the weekly export estimates we were using in the WPSR.

Beginning in 2013, we investigated data sources and methodologies to improve our weekly estimates of petroleum and biofuels exports. We found that we could use the administrative data collected by CBP through the ACE system, which cover U.S. exports (except for most exports to Canada), to improve our weekly export estimates. To obtain these data, we established a Memorandum of Understanding with CBP in September 2014 through the International Trade Data System (ITDS) process. In August 2016, the Office of Management and Budget (OMB) granted us an exception to the provisions of OMB's *Statistical Policy Directive No. 3 on the Compilation, Release, and Evaluation of Principal Federal Economic Indicators.* This exception permits us to publish weekly estimates of export volumes of crude oil and six refined petroleum products and biofuel categories before USCB releases the official U.S. trade data.

Since August 31, 2016, we have continuously monitored and analyzed the estimates derived from the recently acquired near real-time CBP export data. The remainder of this document describes the method we used for the WPSR release on June 1, 2023, and onward. You can find a description of changes to our estimation methodology over time in Attachment B.

Trade Data Collection and Dissemination

To derive weekly estimates of domestic petroleum product supplies, WPSR relies on estimates of weekly U.S. exports of:

- Crude oil
- Finished motor gasoline
- Motor gasoline blending components
- Fuel ethanol
- Kerosene-type jet fuel
- Distillate fuel oil
- Residual fuel oil
- Propane
- Other oils
  - Biofuels
  - Asphalt
  - Aviation gasoline
  - Butanes
  - Ethane
  - Lubricants
  - Miscellaneous products
  - Non-biofuel oxygenates
  - Petroleum coke
  - Unfinished oils
  - Wax

CBP collects the export data that we use to develop these weekly estimates.

In general, entities exporting products from the United States must file export transaction data with CBP. This information is included in CBP's ACE system. As a result of an agreement with CBP, we now receive these unedited transactional ACE data files from CBP and use them to develop aggregate estimates of weekly exports for the WPSR. The data from CBP include an export date that makes it possible to group transactions into WPSR weekly reporting periods. We must edit and impute the data received from CBP to avoid using questionable data values for analysis, modeling, and aggregation. This document later describes the edit and imputation methods we use.

One drawback to using these data is that we cannot follow up with exporters to resolve questions about the CBP data. As a part of their standard procedures, USCB can follow up with exporters to address questionable data and, if necessary, obtain corrections while it prepares official U.S. trade statistics. However, we do not receive a record of these corrections.

Under agreements between the United States and Canada, entities exporting to Canada are not required to report to CBP. As a result, CBP data generally exclude U.S. exports to Canada. Products that require an export license are an exception, and export transactions involving licensed products must be reported to CBP regardless of the destination country. Crude oil was a licensed export product until December 2015. USCB has an agreement with Statistics Canada to obtain Canada's monthly imports data from the United States, which USCB uses to produce monthly estimates of U.S. exports in official U.S. trade statistics. However, we do not have early access to Canada's import data. To address these issues, we developed methodologies to edit the raw CBP data and to estimate weekly exports to Canada.

The CBP data we use to estimate weekly exports begin with the same data that the USCB uses as its initial input to prepare its monthly *International Trade in Goods and Services* report (FT900), a principal economic indicator. The FT900 has a six-week publication lag, but we release WPSR data each Wednesday for the weekly report period ending the previous Friday. As a result of the lag, we incorporate historical, but not current, FT900 data in the WPSR exports estimation methodology, described below. Because USCB publishes FT900 data monthly and the monthly publication cycles are more closely aligned, we use the FT900 data to report monthly export volumes we include in our *Petroleum Supply Monthly* (PSM) and *Petroleum Supply Annual*.

Editing Methodology

Editing methods

Before producing the published estimates for the WPSR, we review several aspects of the CBP data after we receive it. For all products, we:

- Remove duplicate records
- Exclude records with export dates beyond the reporting period
- Exclude products that are out of the WPSR's scope
- Convert all commodity quantities to barrels exported
- Calculate the dollar value per barrel and compare against daily spot prices from Refinitiv (formerly Thomson-Reuters) to identify extreme values
- Use a procedure based on shipment weight, commodity weight, or commodity value to replace the barrels exported based on reported commodity quantities, if necessary (shipment weight is the combined weight of the commodity being shipped and its packaging, but not the weight of the vessel)
- Use estimate based on historical CBP data if replacements based on commodity value or weight are unacceptable

See Attachment A for a more complete list of editing rules that we apply to the CBP data.

Estimation methods

Once editing is complete, we apply a statistical model to correct for coverage differences in the CBP data based on our comparisons to the USCB monthly volumes that are published in the PSM.

Model 3

We refer to the post-editing statistical model as Model 3 to distinguish it from earlier models. We use a two-component approach to generate estimates. The first component is a linear regression model to estimate the exports to destinations other than Canada. The formula for the regression component is:

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$$

where

t denotes a month and

y_t = exports of crude oil, finished motor gasoline, motor gasoline blending components, fuel ethanol, kerosene-type jet fuel, distillate, residual fuel oil, propane, or other oils to destinations other than Canada in barrels per day as published in the PSM for month t;

x_t = edited CBP export data to destinations other than Canada in barrels per day aggregated to match the monthly frequency and reference period of y_t;

β_0 = the intercept;

β_1 = the regression coefficient for the regressor x_t; and

ε_t = the error term.

After we edit the CBP export data (which excludes Canada), we apply the estimated regression coefficients to the edited data each week to produce the published weekly exports data. This method assumes a similar linear relationship between the PSM and CBP data (excluding Canada) on a weekly basis for the weeks in each month. We will continue to assess the results of this approach and identify alternative methods to develop weekly export volume estimates, if needed.

The second component is a three-month moving average of U.S. exports to Canada in thousand barrels per day as reported in the latest series of the PSM.

The WPSR Review Team may adjust the estimates calculated using the above procedures, if necessary, as part of the macro-editing process they conduct before we release the published estimates. As part of this process, the team may adjust the exports data to resolve inconsistencies in the product supplied calculation and the balance quantity for crude oil. Adjustments may include replacing export quantities estimated from CBP data with averages based on historical aggregates published in WPSR, possibly with the historical aggregates published in the PSM. For more information on the macro-editing process, see the Macro Editing section of WPSR's *Appendix B: Explanatory Notes and Detailed Methods Report.*

Summary

The accuracy of our WPSR estimates of U.S. weekly exports has improved because we are now using near real-time petroleum export data from CBP, published monthly trade data from the USCB, and revised estimation methodologies. We will continue to monitor the performance of these methods and will modify them as needed to further improve their accuracy.

Attachment A. Edit Rules

The following are the calculations and edit rules that we use in preparing the CBP data to produce published aggregate statistics of U.S. exports:

- Identify and remove records that are outside the reporting period.
- Identify and include records from the previous week's reporting period that were received too late for WPSR processing to be included in the previous week's estimates.
- Identify and remove duplicate records.
- Exclude exports originating from U.S. territories (for example, Puerto Rico and U.S. Virgin Islands).
- Map Schedule B and Harmonized Tariff Schedule (HTS) codes to WPSR product codes.
- Keep only products that are within the scope of WPSR products.
- Convert units of quantity not already in barrels (for example, kilograms, liters, and tons) to barrels.
- Calculate a dollar value per barrel using spot prices from Refinitiv (formerly Thomson-Reuters).
- Calculate a unit price (dollars per barrel), based on value (in dollars) divided by quantity of commodity exported.
- Compare the unit price with spot prices to determine outliers based on an acceptable range around spot prices.
- Flag extreme values.
  - Recalculate weight of the commodity exported based on quantity (in barrels) and a conversion factor.
  - Compare the recalculated weight against a proxy quantity such as shipment or commodity weight based on an acceptable range around shipment or commodity weight.
  - Use the original reported quantity in barrels if the recalculated weight is within an acceptable range around shipment or commodity weight.
  - Use the gross weight to impute reported quantity if the recalculated weight is not within an acceptable range around the shipment or commodity weight.
    - Recalculate a unit price (dollars per barrel), compare with spot prices, and flag those values outside of an acceptable range.
  - Calculate an estimate based on commodity value (in dollars) divided by spot price per barrel for values outside of the acceptable range.
  - Compare the original recalculated weights and estimate based on commodity dollar value with maximum capacity.
    - If estimates based on both weight values and commodity dollar values exceed an acceptable range around maximum capacity, use a specified upper percentile based on an analysis of the upper tail of the distribution of quantity from historical CBP data relative to maximum capacity.

Attachment B. Updates to WPSR's Methodology for Estimating Exports

Phase 1: Implemented starting with the WPSR published on March 8, 2017

- We updated the regression model (Model 1), which we previously used to estimate total exports from edited CBP data. The regression model used CBP data as an input and accounted for U.S. exports to Canada that were missing from CBP data. The updated model (Model 2) included an intercept term and used daily CBP data that we edited to identify and impute questionable quantities and then we aggregated to monthly totals as input. Model 1 used monthly USCB data that we interpolated to a weekly series.
- We began using Model 2 to estimate finished motor gasoline exports. We had used the previous method to report finished motor gasoline exports equal to the aggregation of CBP data after editing and imputation.

Phase 2: Implemented starting with the WPSR published on June 7, 2017

- We used Model 2 to estimate crude oil exports, replacing the Unobserved Components Model (UCM).

Phase 3: Implemented starting with the WPSR published on March 14, 2018

- We continued using edited CBP data to estimate distillate fuel oil and propane (unchanged since August 2016).
- We updated the methodology to estimate exports of crude oil, kerosene-type jet fuel, residual fuel oil, finished motor gasoline, and other oils.
- For residual fuel oil and finished motor gasoline, we reported edited CBP data without further modeling.
- For crude oil, kerosene-type jet fuel, and other oils, we used an updated methodology (called Model 3) as described in the Estimation Methods section.

Phase 4: Implemented starting with the WPSR published on July 8, 2020

- We updated the methodology to estimate exports of residual fuel oil using Model 3 as described in the Estimation Methodology section.
- We continued using the same methodology for all other products (unchanged since March 2018).

Phase 5: Implemented starting with the WPSR published on July 8, 2021

- We updated the methodology to estimate exports of propane using commodity weight instead of commodity quantity.
- We continued using the same methodology for all other products.

Phase 6: Implemented with the WPSR published on July 21, 2021

- We updated the methodology to allow for adjustments by the WPSR Review Team, if necessary, to resolve inconsistencies in the calculation of product supplied during the macro-editing process.

Phase 7: Implemented with the WPSR published on August 11, 2021

- We reverted to the original methodology to estimate exports of propane (commodity quantity instead of commodity weight).
- During July 2021, we updated the methodology to estimate exports of propane using commodity weight, instead of commodity quantity, as described in Phase 5. However, the WPSR Review Team repeatedly observed data inconsistencies when using the estimates based on the new methodology in the calculation of weekly product supplied in July. As a result, we returned to the original methodology and will estimate exports of propane by converting the first unit of quantity (cubic meters for propane) to barrels, starting with the WPSR published on August 11, 2021. We will continue to research possible improvements to the original methodology for estimating exports of propane.

Phase 8: Implemented with the WPSR published on September 15, 2021

- We updated the methodology to estimate exports of propane using Model 3 as described in the Estimation Methodology section.

Phase 9: Implemented with the WPSR published on March 16, 2022

- We updated the methodology to estimate exports of finished motor gasoline and distillate using Model 3 as described in the Estimation Methodology section.
- We updated our edit rules for the current week to include additional records from the previous week's reporting period that were received too late to be included in previous week's estimates. We include any of these records that pass edit rules in the current week's estimates.
- We updated our deduplication procedure, which will result in more records input to the editing process.

Phase 10: Implemented with the WPSR published on June 1, 2023

- We began showing fuel ethanol and total motor gasoline (sum of finished motor gasoline and motor gasoline blending components) export estimates separate from other oils export estimates.
- We included motor gasoline blending component export estimates with the finished gasoline export estimates due to the low motor gasoline blending components export participation.