Avista Corp.
1411 East Mission P.O. Box 3727
Spokane. Washington 99220-0500
Telephone 509-489-0500
Toll Free 800-727-9170

January 14, 2025

Public Utility Commission of Oregon
Attn: Filing Center
201 High St. SE, Suite 100
Salem, OR 97301

RE: UG 519 – Filing - Substitute Exhibit 701

Attached for electronic filing is the filing of Substitute Exhibit 701 in Docket UG-519. After discussions with Staff, it was discovered that the originally filed Exhibit 701 incorrectly contained data using the Spring 2024 forecast as opposed to the Fall 2024 forecast. Substitute Exhibit 701 corrects for that, and should replace the originally filed Exhibit 701, in its entirety.

Please direct any questions regarding this filing to Kaylene Schultz at (509) 495-2482.

Sincerely,

/s/ David J. Meyer

David J. Meyer
Vice President and Chief Counsel for Regulatory
and Governmental Affairs

Enclosure

BEFORE THE
PUBLIC UTILITY COMMISSION OF OREGON

DOCKET NO. UG-519

DR. GRANT D. FORSYTH
Substitute Exhibit No. 701

Load Forecast Modeling Overview

1. Introduction

This exhibit covers four main areas. Section 2 provides information on weather forecast drivers, and how they are adjusted for the Company's billing period. Section 3 provides information on non-weather forecast drivers used in conjunction with autoregressive-integrated-moving average (ARIMA) models. Section 4 presents the use per customer and customer forecasting models using standard econometric notation. That section is organized around the four main regions in the Company's service territory: Medford, Roseburg, Klamath Falls, and La Grande. Section 4 also provides a broad overview of a recent change in how calendar core natural gas load is derived for the unbilled calculations in the revenue model.

2. Weather Forecast Drivers

Degree days are based on temperature data and are divided into heating degree days (HDD), quality heating degree days (QHDD), and cooling degree days (CDD). HDD reflect usage in the colder months; CDD reflects usage in the summer months; and QHDD reflect usage in the coldest winter months of December, January, February, and March. The baseline for calculating HDD and QHDD is 65 degree Fahrenheit.

Because of Avista's (AVA) billing lags, calendar (CAL) degree day data has to be adjusted as follows:

[2.1] $HDD_t^{AVA} = 0.5 \cdot (HDD_t^{CAL}) + 0.5 \cdot (HDD_{t-1}^{CAL})\ for\ month\ t = Jan, \ldots, Dec$

QHDD are calculated as:

[2.3] $QHDD_t^{AVA} = 0.5 \cdot (HDD_t^{CAL})\ for\ month\ t = Dec$

[2.4] $QHDD_t^{AVA} = 0.5 \cdot (HDD_t^{CAL}) + 0.5 \cdot (HDD_{t-1}^{CAL})\ for\ month\ t = Jan\ and\ Feb$

[2.5] $QHDD_t^{AVA} = 0.5 \cdot (HDD_{t-1}^{CAL})\ for\ month\ t = Mar$

[2.6] $QHDD_t^{AVA} = 0\ for\ t = Apr, \ldots, Nov$

Below, HDD_t^{AVA}, CDD_t^{AVA}, and $QHDD_t^{AVA}$, is referred to as Avista adjusted (AVA) data. Normal weather is defined as a 20-year moving average. All forecasts use the most recent 20-year moving average as normal weather going forward. This calculation is conducted for each of Avista's four Oregon regions: Medford, Roseburg, Klamath Falls, and La Grande. As can be seen in Section 4, degree days are often squared to take into account non-linear relationships between customer usage and weather.

3. Non-Weather Forecast Drivers

Non-weather drivers are energy price (RAP); population (POP); U.S. Federal Reserve industrial production index (IP), non-weather seasonal dummies (SD); trend functions (T, the natural log, lnT, and/or RAMP); Western housing starts (WHS); employment (EMP); dummies for outliers

(OL); and periods of possible structural change (SC). The SC dummies control for periods where there are deviations from long-run behavior trends. This could be due to unique economic shocks and/or the sudden in- or out-migration of customers that changes the series behavior. Household Income does not appear as an explanatory variable in any of the residential models because it was found not to be statistically significant. In the case of Oregon, RAP appears in the residential schedule 410 for Roseburg and is lagged one year. This means the model indicates that it takes one year for a price change to impact behavior. In some of the schedules there is a special dummy control variable called COVIDD; COVIDD controls for the influence of the COVID pandemic on load behavior in some, but not all, schedules.

Pure ARIMA and ARIMA "transfer function" models are frequently used. In these cases, the error structure is expressed as $\epsilon_{t,y}$ = ARIMA$\epsilon_{t,y}(p,d,q)(p_k,d_k,q_k)_k$. The term p is the autoregressive (AR) order, d is the differencing order, and q is the moving average (MA) order. The term p_k is the order of seasonal AR terms, d_k is the order of seasonal differencing, and q_k is the seasonal order of MA terms. The seasonal values are related to "k," which is the frequency of the data. With the current data set, k = 12 for both use per customer (THM/C, THM = therms) and customers (C) for each schedule.

For the main residential and commercial schedules, the modeling approach takes in account that historical customer growth between the main schedules in Medford and Klamath Falls is highly, positively correlated. To ensure this relationship is reflected in the customer and load forecasts, the customer models for the 420 commercial schedules in Medford and Klamath Falls use 410 residential customers as a forecast driver. In the case of Roseburg and La Grande, the correlation between residential and commercial growth is weak, so the residential customer forecast is not used as a forecast driver. This means for Medford and Klamath the final customer forecast for residential schedule 410 are used as a variable to forecast commercial customers. In all city areas, the 410 customer forecasts are driven by population forecasts.

Note that dates on the some of the dummy variables are followed by "↑," which means "going forward in time." For example, "Jan 2009↑=1" means, "From January 2009 forward the dummy variable equals 1." Also note that t = month and y = year. For example $THM/C_{t,y,MED410.r}$ should be read as, "Therms per customer in month t, of year y, for Medford residential (r) schedule 410. For industrial (i) and commercial (c) similar notation is used.

Not all schedules require an ARIMA based model. In some schedules, simple regression and smoothing methods are used because they offer the best fit for usage that is periodic and/or irregular; is in a long-run, but steady, decline; and/or is seasonal but not weather related.

Total THM for each schedule is arrived at by multiplying customer forecasts by use per customer forecasts. In some cases, these forecasts are adjusted to reflect information that cannot be accounted for a model based on historical data.

<u>4. Use Per Customer and Customer Forecast Models by Region</u>

This section presents the use per customer (UPC) and customer forecast models. The total load for a given schedule is derived by multiplying the UPC forecast by the customer forecast. The system load is then generated by summing across all the forecasts by schedule.

4a. ***Medford, OR Forecasting Models***

The forecasting models for the Medford region (Jackson and Josephine counties) are given below for the residential, commercial, and industrial sectors:

Residential Sector, THM:

[7.50] $THM/C_{t,y,MED410.r} = \alpha_0 + \alpha_1 HDD_{t,y}^{AVA} + \alpha_2 (HDD_{t,y}^{AVA})^2 + \boldsymbol{\omega_{SD} D_{t,y}} + \omega_{OL} D_{Mar\ 2008=1} +$
$\omega_{OL} D_{Feb\ 2011=1} + \omega_{OL} D_{Dec\ 2015=1} + \omega_{OL} D_{Feb\ 2018=1} + \omega_{OL} D_{Dec\ 2019=1} + \omega_{OL} D_{Mar\ 2020=1} +$
$+ \omega_{OL} D_{Dec\ 2020=1} + \omega_{OL} D_{Feb\ 2021=1} ARIMA \epsilon_{t,y}\ (11,0,0)(1,0,0)_{12}\ for\ y = 2007\uparrow$

[7.50] Model notes:
1. AHS is not included because its impact is unstable and not statistically significant. RAP was removed in the spring 2024 forecast because it was no longer statistically significant and its sign was unstable.
2. Model is restricted to 2007 ↑ because including earlier periods produces unacceptable model diagnostics.
3. In the fall 2024 forecast a new residential schedule was introduced—schedule 411—for multifamily gas users. Previously, the customers in schedule 411 where part of schedule 410. Because there is very little historical usage data available, the fall 2024 forecast had to estimate the load profile and annual usage to derive monthly UPC forecasts. This method is described in the fall 2024 forecast sheet for Oregon. See also the notes to [7.51]. The removal of schedule 411 customers from schedule 410 did not appear to significantly change the existing schedule UPC.

Residential Sector, Customers:

[7.51] $C_{t,y,MED410+411.r} = \alpha_0 + \alpha_1 POP_{t,y,JACK+JOS} + \omega_{SC} D_{Nov\ 2004\uparrow=1} + \omega_{SC} D_{Nov\ 2019\uparrow=1} + \omega_{SC} D_{Oct\ 2020\uparrow=1} +$
$\omega_{OL} D_{Dec\ 2005=1} + \omega_{OL} D_{Sep\ 2020=1} + \omega_{OL} D_{Oct\ 2018=1} + ARIMA \epsilon_{t,y}\ (11,1,0)(2,1,0)_{12}$

[7.51] Model notes:
1. SC dummy for 2004↑ controls for a step-up in customers; SC dummy for November 20219 controls for a step-down in customers; and SC dummy for October 2020↑ and OL dummy for September 2020 control for the impact of the 2020 wildfires which destroyed around 1,000 customers (both residential and commercial) in the Medford region.
2. POP is Jackson plus Josephine counties. The coefficient on POP is positive, but no longer statistically significant.
3. The creation for schedule 411 in fall 2024 caused a step down in customers in schedule 410 because historically the schedule 411 customers where part of schedule 410. To control for this, the fall 2024 forecast continued to forecast schedules 410 and 411 together and then subtract off the forecast of the schedule 411 customers. In the fall 2024 forecast, the schedule 411 customer forecast was the average number of customers for June, July, and August 2024. Even though schedule 411 started in April 2024, the number customers estimated for April 2024 was too high and dropped in May 2024. It appears to have stabilized over the Jun-Aug period.

Commercial Sector, THM:

[7.52] $THM/C_{t,y,MED420.c} = \alpha_0 + \alpha_1 HDD_{t,y}^{AVA} + \alpha_2 (HDD_{t,y}^{AVA})^2 + \omega_{SC} D_{Aug\ 2007\uparrow=1} + \omega_{SC} D_{Aug\ 2020\uparrow=1} + \omega_{OL} D_{Dec\ 2005=1} +$
$\omega_{OL} D_{Mar\ 2010=1} + \omega_{OL} D_{April\ 2010=1} + \omega_{OL} D_{Feb\ 2018=1} + ARIMA \epsilon_{t,y}\ (11,0,0)(3,1,0)_{12}$

[7.52] Model notes:
1. SC dummies controls for a step-down in UPC starting in August 2007 and a step-up in August 2020.
2. Work is ongoing to determine if RAP can be integrated into the model.
3. Error structure is white noise, but not quite normally distributed.

[7.53] $THM/C_{t,y,MED424.c} = \alpha_0 + \alpha_1 HDD_{t,y}^{AVA} + \omega_{SC} D_{Jan\ 2006\uparrow=1} + \omega_{OL} D_{Dec\ 2004=1} + \omega_{OL} D_{May\ 2005=1} +$
$\omega_{OL} D_{Apr\ 2006=1} + \omega_{OL} D_{Feb\ 2011=1} + \omega_{OL} D_{Dec\ 2012=1} + \omega_{OL} D_{Jan\ 2013=1} + \omega_{OL} D_{Feb\ 2014=1} +$
$\psi COVIDD_{Apr-Mar\ 2021=1} + ARIMA \epsilon_{t,y}\ (10,0,0)(3,1,0)_{12}$

[7.53] Model notes:
1. SC Dummy controls for a step-down in UPC starting in January 2006.

2. COVIDD dummy controls for the impact of the shut-down shock.

[7.54] $THM_{t,y,MED444.c} = \frac{1}{N}\sum_{j=1}^{N} THM_{t,y-j} \; for \; y - j = 2004 \uparrow$
$up \; to \; the \; most \; recent \; month, then \; repeat \; forecast \; values$

[7.54] Model notes:
1. A regression model produces poor diagnostics because of the many sequential months (mostly in the winter and spring) with zero customers.
2. This schedule appears to be dead. There has not been any significant load since 2019. The current forecast assumes not load over the forecast horizon.

[7.55] $THM/C_{t,y,MED440.c} = \alpha_0 + \alpha_1 HDD_{t,y}^{AVA} + \omega_{SC} D_{Dec\,2019\uparrow=1} + \omega_{SC} D_{Feb\,2023-Sept\,2023=1} + \omega_{OL} D_{Sep\,2008=1} +$
$\omega_{OL} D_{Oct\,2008=1} + \omega_{OL} D_{Nov\,2008=1} + \omega_{OL} D_{May\,2009=1} + \omega_{OL} D_{Jan\,2014=1} + \omega_{OL} D_{Aug\,2016=1} + \omega_{OL} D_{Sep\,2023=1} +$
$\omega_{OL} D_{Apr\,2024=1} + ARIMA\epsilon_{t,y}\,(1{,}0{,}0)(0{,}0{,}0)_{12} \; for \; t, y = May\; 2007 \uparrow$

[7.55] Model notes:
1. Model is restricted to May 2007 ↑ because schedule does not start until February 2007. Estimation period starts in May 2007 to account for the lag in the error term. In the underlying regression data base (ORALLOTHERSCHEDULES), data prior to February 2007 is recorded as -9999 going back to January 2004. Failure to move up the estimation period will cause the -9999 values to be incorporated into the estimation.
2. Starting in January 2020, there was a movement of meters for the Medford school district; the number of meters was reduced by nine so the customer count for this schedule fell roughly in half. Those nine meters, representing a small amount of load, were moved to schedule 420. Therefore, UPC is recalculated by reducing the number of historical meters by nine. This is an approximation of what UPC should have been because the load of nine meters is still embedded in the historical data—but as noted, the load associated with the moved meters is small. Because the nine meters are small relative to 420 customers and load, no adjustment is made to schedule 420. See also discussion for customer equation [7.60].
3. The first SC dummy controls for a step-down in UPC starting in December 2019. The second SC dummy for February-September 2023 controls for a temporary step-up in UPC starting February 2023 to September 2023; this reflects the movement of Murphy Plywood from schedule 456 industrial to this schedule. However, it should have been moved to schedule 440 industrial. This mistake was not discovered until after the fall 2023 forecast was complete. For rate making and revenue purposes, the accidental inclusion in this schedule has no material impact. For the spring 2024 forecast, Murphy Plywood will be correctly reflected in 440 industrial.
4. Although the Shapiro-Wilk test indicates error term normality, it is still close to failing the test. New specifications may be required in future forecasts.

[7.56] $THM/C_{t,y,MED456.c} = \alpha_0 + \alpha_1 HDD_{t,y}^{AVA} + \boldsymbol{\omega_{SD} D_{t,y}} + \omega_{SC} D_{Feb\,2024\uparrow=1} + \omega_{OL} D_{Apr\,2014=1} +$
$\omega_{OL} D_{Apr\,2022=1} + ARIMA\epsilon_{t,y}\,(2{,}0{,}0)(0{,}0{,}0)_{12} \; for \; y = 2014 \uparrow$

[7.56] Model notes:
1. Model restricted to January 2014↑ because including earlier periods, produce fitting problems as well as forecasts that do match current behavior. There appears to have been a significant change in seasonality.
2. SC dummy controls for a significant step-up in UPC because of an expansion of the Rogue Valley Medical center starting in February 2024.
3. In late 2023 Tree Top Inc. closed. The current forecast was run assuming that Tree Top Inc. was never in schedule 456. This data series is ORMEDSCH456ADJupc.c.

Commercial Sector, Customers:

[7.57] $C_{t,y,MED420.c} = \alpha_0 + \alpha_1 C_{t,y,MED410+411.r} + \boldsymbol{\omega_{SD} D_{t,y}} + \omega_{SC} D_{Feb\,2016\uparrow=1} + \omega_{OL} D_{Jan\,2016=1} + \omega_{OL} D_{May\,2020=1} +$
$\omega_{OL} D_{Jun\,2020=1} + \omega_{OL} D_{May\,2022=1} + ARIMA\epsilon_{t,y}\,(5{,}1{,}0)(0{,}0{,}0)_{12}$

[7.57] Model notes:
1. $C_{t,y,MED410+411,r}$ are residential customers from residential schedule 410. They are being used as a forecast driver because of the historical positive correlation between residential and commercial customer growth. See Tables 5.1 and 5.2. However, in the future, POP may become a better driver. Model results with POP are fairly close to model shown above.
2. OL dummies for May and June may reflect short-term impacts of the COVID shock. SC dummy controls for a step-up in UPC starting in January 2016.
3. Because the impact of the wildfires is reflected in $C_{t,y,MED410+411}$, they are controlled for through that variable and not an SC dummy.

[7.58] $C_{y,MED424.c} = C_{y-1} + (\hat{\alpha}_0 + \hat{\alpha}_1 \Delta EMP_{y-1,4County})$

[7.58] Model notes:
1. This model reflects a recommendation by Oregon staff in the 2016 rate case to include employment as an economic driver for schedule

424 commercial customers. The estimated equation in parenthesis reflects the regression estimated of $\Delta C_{y,MED424.c} = \alpha_0 + \alpha_1 \Delta EMP_{y-1,4County} + \varepsilon_t$ using annual customer data since 2004. Annual data is used to smooth over the sometimes volatile changes in the monthly customer number. In addition, customer increases and decreases around the long-run trend tend to occur in steps. The combination of steps and month-to-month volatility creates significant economic problems when trying to model around the monthly data. For example, even with intervention variables, tests for error normality always indicated non-normal error terms with the use of monthly data.
2. $\Delta C_{y,MED424.c}$ is the change in customers in year y (customer change between year y and y-1) and $\Delta EMP_{y-1,4County}$ is the change in total non-farm employment in Jackson+Josephine, Klamath, and Douglas counties in year y-1 (employment change between year y-1 and y-2). Staff originally suggested lagged total employment for Oregon, but the correlation between schedule 424 customers and employment for the four-county area is higher. The forecasted employment values for Jackson+Josephine County are derived from the employment growth forecasts used in the Jackson+Josephine County population forecast. The forecasts for Douglas and Klamath counties come from IHS. In IRP years, IHS forecasts all counties will be used for the out years.
3. The annual forecast value for each year, F(·), is assumed to hold for each month of that year. That is: $F(C_{y,MED424.c}) = F(C_{t,y,MED424.c})$. Given the step-like behavior of the monthly series, this is a reasonable assumption.
4. The forecast and regressions for this schedule can be found in the Excel file folder “OR 4County Sch 424c Cus.”

[7.59] $C_{t,y,MED444.c} = 1 \; if \; (THM/C_{t,y})_{MED,444.c} > 0$

[7.59] Model notes:
1. There is typically only one customer served by this schedule. Therefore, the customer forecast is automatically set to one whenever the load forecast is greater than zero. Schedule appears to have died. See notes for [7.54].

[7.60] $C_{t,y,MED440.c} = C_{t-1}$

[7.60] Model notes:
1. Starting in January 2020, there was a movement of meters for the Medford school district; the number of meters was reduced by nine so the customer count for this schedule fell roughly in half. This resulted in a significant change in the historical series. These nine meters, representing a small amount of load, were moved to schedule 420. Because the nine meters are small relative to 420 customers and load, no adjustment is made to schedule 420. See also discussion for customer equation [7.55].
2. The forecast was moved to a simple model in the fall 2023 forecast because periodic in-migrations of large customers from schedule 456.

[7.61] $C_{t,y,MED456.c} = \frac{1}{12}\sum_{j=1}^{12} C_{t-j}$

[7.61] Model notes:
1. Data starts January 2004. Excluding a March 2013 outlier, the customer count fluctuates between 5 and 8 with changes occurring in steps between prolonged periods of stability. The customer count shows no trend, limited volatility, and no seasonality. In the fall 2024 forecast the moving average was adjusted for the closure of Tree Top Inc.

Industrial Sector, THM:

[7.62] $THM/C_{t,y,MED420.i} = \alpha_0 + \alpha_1 HDD_{t,y}^{AVA} + \delta_1 IP_{t,y} + \omega_{SC} D_{May\ 2019\uparrow=1} + \omega_{SC} D_{Jun\ 2020\uparrow=1} + \omega_{OL} D_{Mar\ 2011=1} + \omega_{OL} D_{Jan\ 2018=1} + \omega_{OL} D_{Nov\ 2022=1} + ARIMA\epsilon_{t,y}\,(1,0,0)(2,1,0)_{12} \; for \; t,y = May\ 2010$

[7.62] Model notes:
1. Model is restricted to May 2010 ↑ because schedule does not start until December 2006. Estimation period starts in April to account for the lag in the error term. In the underlying regression database (ORALLOTHERSCHEDULES), data prior to December 2006 is recorded as -9999 going back to January 2004. Failure to move up the estimation period will cause the -9999 values to be incorporated into the estimation.
2. SC dummies controls for a step-up in UPC starting in May 20219 and a step-down in UPC staring in January 2020.
3. Starting in in 2019 there appears to be an upward shift in base-load and it’s unclear why. This will have to be monitored for next load forecast.

[7.63] $THM/C_{t,y,MED424.i} = \frac{1}{12}\sum_{j=1}^{12} THM/C_{t-j}$

[7.63] Model notes:
1. In the fall 2019 forecast, the forecast was switched from an ARIMA based model to a 12-month moving average. Due to customer changes, UPC has become highly erratic with shifting seasonality. Multiple regression models failed to produce acceptable diagnostic tests.
2. Starting in March 2018, the customer count fell from 2 to 1. See also [7.65] model notes regarding the movement of our Biomass facility customer (schedule 447b) to this schedule.

[7.64] $THM/C_{t,y,MED440.i} = \alpha_0 + \boldsymbol{\omega_{SD} D_{t,y}} + \omega_{SC} D_{Nov\ 2020\uparrow=1} + ARIMA\epsilon_{t,y}\,(3,0,0)(0,0,0)_{12} \; for \; y =$

$September\ 2015 \uparrow$

[7.64] Model notes:
1. [7.64] has experienced repeated structural changes due to firms moving in and out of the schedule. This has required repeated adjustments to the time-series to adjust for this movement—the adjusted time series is ORMED440ADJupc.i. The timeline of customer in- and out-migrations follows: In March 2019, Carestream Health moved from schedule 456 to this schedule. However, in November 2023 Carestream returned to schedule 456. In January 2020, Timber Products of Medford and Grants Pass moved from schedule 456 to this schedule. In March 2024 Plychem moved from schedule 456 to this schedule. In each of these movements, ORMED440ADJupc.i was adjusted to try to create a time-series consistent with the current customer mix. This reduces the likelihood of losing load because of customer migrations. The same can be said for Medford schedule 456, which has also been impacted by the in- and out-migrations from schedule 440. The ORMED440ADJupc.i series is restricted to September 2015↑ to arrive at a the most consistent time-series. See also model notes for Medford schedule 456.
2. SC dummy controls for a step-up in UPC staring in November 2020. This could reflect new customers entering this schedule or existing firms expanding load.

[7.65] $THM/C_{t,y,MED447b.i} = moved\ to\ schedule\ 424\ industrial$

[7.65] Model notes:
1. This special contract customer is a biomass generator that burns wood products to generate electricity. This customer did not become consistently active until August 2017.
2. After the June 2018 forecast, it was initially determined that this firm will most likely use gas to fire its operations during maintenance periods when biofuel cannot be burned. As a result, any observed gas usage would be temporary. However, early in 2019 the company decided to move to schedule 424 industrial. The decision to move to 424 reflects contractual problems with its gas supplier (which was not Avista) as a special contract gas customer. See also notes to schedule 424 industrial.

[7.66] $THM/C_{t,y,OR456.i} = \alpha_0 + \delta_1 IP_{t,y} + \omega_{OL} D_{Feb\ 2019=1} + ARIMA\epsilon_{t,y}\,(5,0,0)(2,1,0)_{12}\ \ for\ y = 2010 \uparrow$

[7.66] Model notes:
1. Originally, the majority of 456 industrial customers in Medford are in the wood products industry; therefore, housing starts, instead of industrial production, was used as a predictor. However, as discussed below, several of these customers moved to schedule 440 and some have gone to schedule 440 and then back to schedule 456. This reduced the significance of WHS, which was replaced by IP in the spring 2020 forecast.
2. The timeline of customer in- and out-migrations follows: In March 2019, Carestream Health moved from schedule 456 to schedule 440. However, Carestream Health moved back to schedule 456 in November 2023. In January 2020, Timber Products of Medford and Grants Pass moved to schedule 440. In November 2023 Roseburg Forest products 447 special contract moved to schedule 456. To adjust for the in- and out-migrations, an adjusted time-series, ORMED456ADJupc.i. was created so that forecasts better reflected UPC behavior. See also [7.64] notes for Medford schedule 440.
5. In spring 2021 forecast, this model aggregated all 456 industrial schedules in Oregon. This was reduced the volatility and structural change of 456 UPC in individual city areas caused by the customer migrations discussed above. In addition, this aggregation is possible because (1) weather is not a forecast driver, and (2) most firms are wood product producers. Also, the Census X11 procedure indicates stable seasonality over the historical estimation period, which is restricted to 2010↑; including the full period to 2004 was producing estimating issues and generating forecasts that did not look like recent history.

Industrial Sector, Customers:

[7.67] $C_{t,y,MED420.i} = \frac{1}{12}\sum_{j=1}^{12} C_{t-j}$

[7.67] Model notes:
1. Data starts November 2006. Excluding outliers in November 2006, November 2009, and February 2011, the customer count fluctuates between 9 and 16 without any clear trend or seasonality. Changes in the customer count occur in steps between prolonged periods of stability.

[7.68] $C_{t,y,MED424.i} = \frac{1}{12}\sum_{j=1}^{12} C_{t-j}$

[7.68] Model notes:
1. Data starts January 2009. Excluding a January 2009 outlier, the customer count fluctuates between 1 and 3 without any clear trend or seasonality. In March 2019, the schedule 447b (biomass plant) moved to schedule 424.

[7.69] $C_{t,y,MED440.i} = \frac{1}{12}\sum_{j=1}^{12} C_{t-j}$

[7.69] Model notes:
1. See also [7.64] and [7.66] for additional information. This is often adjusted based on recent in- and out-migrations by customers.

[7.70] $C_{t,y,MED447b.i} = \textit{schedule has died}$

[7.70] Model notes:
1. This special contract customer (Biomass) is a bio-gas provider. This customer did not become consistently active until August 2017. In early 2019, the customer shifted to schedule 424.
2. In Oregon, special contract customers have a different regulatory status during rate case proceedings.

[7.71] $C_{t,y,MED456.i} = C_{t-1}$

[7.71] Model notes:
1. See also [7.64] and [7.66] for additional information. This is often adjusted based on recent in- and out-migrations by customers.

4b. Roseburg, OR Forecasting Models

The forecasting models for the Roseburg region (Douglas County) are given below for the residential, commercial, and industrial sectors:

Residential Sector, THM:

[7.72] $THM/C_{t,y,ROS410.r} = \varphi_0 + \varphi_1 HDD_{t,y}^{AVA} + \varphi_2 (HDD_{t,y}^{AVA})^2 + \lambda RAP_{t,y-1,OR410} + \omega_{OL} D_{Dec\ 2015=1} + \omega_{OL} D_{Dec\ 2016=1} + \omega_{OL} D_{Jan\ 2018=1} + \omega_{OL} D_{Dec\ 2019=1} + \omega_{OL} D_{Nov\ 2022=1} + ARIMA\epsilon_{t,y}\ (2,0,0)(1,1,0)_{12}\ for\ y = 2010 \uparrow$

[7.72] Model notes:
1. AHS is not included because its impact is unstable and not statistically significant. As of the fall 2023 forecast, RAP's impact is still negative, but is no longer statistically significant.
2. Model is restricted to 2010 ↑ because including earlier periods produces models with very poor diagnostics; in particular, including earlier time periods generates error series that are both non-white noise and not normally distributed. The new series is white noise, but not quite normally distributed.
3. In the fall 2024 forecast a new residential schedule was introduced—schedule 411—for multifamily gas users. Previously, the customers in schedule 411 where part of schedule 410. Because there is very little historical usage data available, the fall 2024 forecast had to estimate the load profile and annual usage to derive monthly UPC forecasts. This method is described in the fall 2024 forecast sheet for Oregon. See also the notes to [7.73]. The removal of schedule 411 customers from schedule 410 did not appear to significantly change the existing schedule UPC.

Residential Sector, Customers:

[7.73] $C_{t,y,ROS410+411.r} = \varphi_0 + \varphi_1 POP_{t,y,DOUGLAS} + \boldsymbol{\omega_{SD} D_{t,y}} + \omega_{SC} D_{Jan\ 2020\uparrow=1} + \omega_{OL} D_{Oct\ 2018=1} + \omega_{OL} D_{Mar\ 2019=1} + \omega_{OL} D_{Nov\ 2020=1} + ARIMA\epsilon_{t,y}\ (12,1,0)(0,0,0)_{12}\ for\ y = 2010 \uparrow$

[7.73] Model notes:
1. POP is population for Douglas County, OR.
2. SC dummy control for step-up in customers starting in November 2020.
3. Model restricted to 2010↑ because of changing seasonality; including the full time series back to 2004 was creating estimation issues.
4. The creation for schedule 411 in fall 2024 caused a step down in customers in schedule 410 because historically the schedule 411 customers where part of schedule 410. To control for this, the fall 2024 forecast continued to forecast schedules 410 and 411 together and then subtract off the forecast of the schedule 411 customers. In the fall 2024 forecast, the schedule 411 customer forecast was the average number of customers for June, July, and August 2024. Even though schedule 411 started in April 2024, the number customers estimated for April 2024 was too high and dropped in May 2024. It appears to have stabilized over the Jun-Aug period.

Commercial Sector, THM:

[7.74] $THM/C_{t,y,ROS420.c} = \varphi_0 + \varphi_1 HDD_{t,y}^{AVA} + \varphi_2 (HDD_{t,y}^{AVA})^2 + \boldsymbol{\omega_{SD} D_{t,y}} + \omega_{OL} D_{Dec\ 2014=1} + \omega_{OL} D_{Mar\ 2019=1} + \omega_{OL} D_{Nov\ 2022=1} + +ARIMA\epsilon_{t,y}\ (3,0,0)(0,0,0)_{12}\ for\ y = 2013 \uparrow$

[7.74] Model notes:
1. Model restricted to 2013↑ because the inclusion of the pre-2013 produces unstable models, including base load forecasts that were substantially higher that more current values. More generally, as a city area, Roseburg's schedules are becoming increasingly difficult to model using the entire available data series.

[7.75] $THM/C_{t,y,ROS424.c} = \varphi_0 + \varphi_1 HDD_{t,y}^{AVA} + \boldsymbol{\omega_{SD} D_{t,y}} + \omega_{SC} D_{Jun\ 2020\uparrow=1} + \omega_{OL} D_{Oct\ 2007=1} + \omega_{OL} D_{Jul\ 2009=1} + \omega_{OL} D_{Mar\ 2011=1} + \omega_{OL} D_{Feb\ 2013=1} + \omega_{OL} D_{Jan\ 2014=1} + \omega_{OL} D_{Mar\ 2019=1} + ARIMA\epsilon_{t,y}\ (11,0,0)(2,0,0)_{12}\ for\ y = 2007 \uparrow$

[7.75] Model notes:
1. Model restricted to 2007↑ because inclusion of earlier periods significant error autocorrelation issues that cannot be resolved.
2. SC dummy controls for a step-down in UPC starting in June 2020.

[7.76] $THM/C_{t,y,ROS440.c} = \varphi_0 + \boldsymbol{\omega_{SD} D_{t,y}} + \omega_{SC} D_{Apr\ 2022\uparrow=1} + \omega_{OL} D_{Oct\ 2019=1} + \omega_{OL} D_{Jun\ 2020=1} + \omega_{OL} D_{Aug\ 2020=1} + \omega_{OL} D_{Sep\ 2020=1} + \omega_{OL} D_{Oct\ 2020=1} + \omega_{OL} D_{Aug\ 2021=1} + \omega_{OL} D_{Apr\ 2022=1} + \omega_{OL} D_{Nov\ 2022=1} + ARIMA\epsilon_{t,y}\ (7,0,0)(0,0,0)_{12}\ for\ t,y = 2019 \uparrow$

[7.76] Model notes:
1. Model is restricted to September 2019 ↑ because the entrance of new customers has changed seasonality. Including earlier periods results in a forecast that does not match more recent behavior.
2. SC dummies controls for a step-down in UPC staring in April 2022. OL dummies for June 2020 through October 2020 control for UPC instability caused by some combination of the COVID shock and wildfires. It's unclear if this impacted firm operations or the ability to read meters.

[7.77] $THM/C_{t,y,ROS456.c} = \varphi_0 + \boldsymbol{\omega_{SD} D_{t,y}} + \omega_{SC} D_{Jan\ 2017\uparrow=1} + \omega_{OL} D_{Feb\ 2018=1} + \omega_{OL} D_{Mar\ 2023=1} + ARIMA\epsilon_{t,y}\ (4,1,0)(0,0,0)_{12}\ for\ y = 2011 \uparrow$

[7.77] Model notes:
1. SC dummy controls for a step-up in UPC starting in January 2017. Data continues to be difficult to model due to ongoing changes in behavior. In the June 2019 forecast, the data was restricted to January 2011↑ to isolate a period of more consistent behavior. Inclusion of data prior to 2011 resulted in forecasts that bore little resemblance to more recent behavior.
2. Inclusion of HDD variables produces a model with significant error structure problems.

Commercial Sector, Customers:

[7.78] $C_{t,y,ROS420.c} = \varphi_0 + \varphi_1 POP_{t,y,DOUGLAS} + \boldsymbol{\omega_{SD} D_{t,y}} + \omega_{SC} D_{Dec\ 2005\uparrow=1} + \omega_{OL} D_{Jan\ 2005=1} + \omega_{OL} D_{Jan\ 2008=1} + \omega_{OL} D_{Mar\ 2019=1} + ARIMA\epsilon_{t,y}\ (9,1,0)(0,0,0)_{12}\ for\ y = 2005 \uparrow$

[7.78] Model notes:
1. Model does not use schedule 410 customers as driver. This reflects the lack of correlation between residential 410 and commercial 420 customer growth. However, POP was added for the 2018 gas IRP and was significant at the 10% level; however, by the time of the spring 2022 forecast it had become insignificant but still consistently positive, so it was left in.
2. The lack of correlation noted above could reflect Roseburg's position between larger cities that offer a range of commercial activities. Competition from these cities may be inhibiting commercial growth in Roseburg. However, as noted above, it now appears the linkage to population is also weakening.
3. Model restricted to 2005↑ because the inclusion of the pre-2005 period produced unstable models starting in the spring 2022 forecast.4. SC dummy controls for a significant step-up in customers starting in December 2005.

[7.79] $C_{t,y,ROS424.c} = C_{y-1} + (\widehat{\varphi_0} + \widehat{\varphi_1} \Delta EMP_{y-1,4County})$

[7.79] Model notes:
1. This model reflects a recommendation by Oregon staff in the 2016 rate case to include employment as an economic driver for schedule 424 commercial customers. The estimated equation in parenthesis reflects the regression estimated of $\Delta C_{y,ROS424.c} = \alpha_0 + \alpha_1 \Delta EMP_{y-1,4County} + \varepsilon_t$ using annual customer data since 2004. Annual data is used to smooth over the sometimes volatile changes in the monthly customer number. In addition, customer increases and decreases around the long-run trend tend to occur in steps. The combination of steps and month-to-month volatility creates significant economic problems when trying to model around the monthly data. For example, even with intervention variables, tests for error normality always indicated non-normal error terms with the use of monthly data.
2. $\Delta C_{y,ROS424.c}$ is the change in customers in year y (customer change between year y and y-1) and $\Delta EMP_{y-1,4County}$ is the change in total non-farm employment in Jackson+Josephine, Klamath, and Douglas counties in year y-1 (employment change between year y-1 and y-2).

Staff originally suggested lagged total employment for Oregon, but the correlation between schedule 424 customers and employment for the four-county area is higher. The forecasted employment values for Jackson+Josephine County are derived from the employment growth forecasts used in the Jackson+Josephine County population forecast. The forecasts for Douglas and Klamath counties come from IHS. In IRP years, IHS forecasts for all counties will be used for the out years.
3. The annual forecast value for each year, F(·), is assumed to hold for each month of that year. That is: $F(C_{y,ROS424.c}) = F(C_{t,y,ROS424.c})$. Given the step-like behavior of the monthly series, this is a reasonable assumption.
4. The forecast and regressions for this schedule can be found in the Excel file folder "OR 4County Sch 424c Cus."

[7.80] $C_{t,y,ROS440.c} = \frac{1}{12}\sum_{j=1}^{12} C_{t-j}$

[7.80] Model notes:
1. Data starts February 2007. Excluding outliers in February 2007, August 2008, and November 2009, the customer count fluctuates between 3 and 7 without any clear seasonality. There is slight downward trend in the customer count that has occurred in steps following prolonged periods of stability.

[7.81] $C_{t,y,ROS456.c} = C_{t-1}$

[7.81] Model notes:
1. Data starts in January 2004. Only one recorded customer since November 2009.

Industrial Sector, THM:

[7.82] $THM/C_{t,y,ROS420.i} = \varphi_0 + \boldsymbol{\omega_{SD} D_{t,y}} + \omega_{SC} D_{Mar\ 2018\uparrow =1} + \omega_{OL} D_{Jan\ 2018\ =1} + \omega_{OL} D_{Apr\ 2018\ =1} + \omega_{OL} D_{May\ 2019\ =1} + \omega_{OL} D_{Aug\ 2024\ =1} + ARIMA\epsilon_{t,y}\ (1,0,0)(0,0,0)_{12}\ for\ y = 2014\uparrow$

[7.82] Model notes:
1. Model is restricted to 2014 ↑ since including earlier periods results in forecasts that are not consistent with more recent UPC behavior. In the underlying regression database (ORALLOTHERSCHEDULES), data prior to September 2009 is recorded as -9999 going back to January 2004. Failure to move up the estimation period will cause the -9999 values to be incorporated into the estimation.
2. SC dummy controls for a step-up in UPC starting in March 2018. OL dummy for August 2024 may be a billing error; this will be reviewed in the next forecast.
3. IP has been removed from the model; the sign had become insignificant and unstable.

[7.83] $THM/C_{t,y,ROS424.i} = See\ SAS\ output\ file\ for\ additon\ of\ new\ customers\ in\ fall\ 2024$

[7.83] Model notes:
1. New customer added in the fall 2024 forecast. Last significant activity was nearly a decade ago.

[7.84] $THM/C_{t,y,ROS440.i} = \varphi_0 + \delta_1 WHS_{t,y-1} + \boldsymbol{\omega_{SD} D_{t,y}} + \omega_{OL} D_{Mar\ 2015\ =1} + \omega_{OL} D_{Jun\ 2021\ =1} + ARIMA\epsilon_{t,y}\ (6,0,0)(0,0,0)_{12}\ for\ t,y = November\ 2011\uparrow$

[7.84] Model notes:
1. Model restricted to January 2011 forward to account for the movement of C&D Lumber and Superior Studs as discussed below.
2. In January 2019, C&D Lumber and Superior Studs moved from Roseburg schedule 446 to schedule 440. For the fall 2019 forecast, the Roseburg 440 UPC series was recalculated by adding in the two companies. This means [7.84] is used to forecast UPC as if C&D Lumber and Superior Studs was always a part of the UPC time series. However, the addition of the two companies caused a structural change in the data series at the end of 2011. As a result, the model is restricted to November 2011 ↑. In the regression database (ORALLOTHERSCHEDULES) the recalculated series is ORROS440ADJupc.i See also model notes for Roseburg schedule 456.
3. Due to the movement of C&D Lumber and Superior Studs, in the June 2019 forecast WHS replaces IP as the economic driver in model [7.84].

[7.85] $THM_{t,y,ROS447m.i} = \varphi_0 + \boldsymbol{\omega_{SD} D_{t,y}} + \omega_{OL} D_{Aug\ 2015\ =1} + \omega_{OL} D_{Jan\ 2019\ =1} + \omega_{OL} D_{Jan\ 2023\ =1} + \omega_{OL} D_{Aug\ 2024\ =1} + \omega_{OL} D_{Jun\ 2021\ =1} + ARIMA\epsilon_{t,y}\ (6,1,0)(2,0,0)_{12}\ for\ t,y = 2013\uparrow$

[7.85] Model notes:
1. In September 2024, Murphy Plywood moved to schedule 440 industrial. See the fall 2024 forecast file. In the spring 2024 forecast, it was expected that Murphy Plywood would move to schedule 456. However, just before the fall 2024 forecast, the company decided to move to schedule 440. The above forecast model was used to forecast Murphy Plywood individually and then adjust the schedule 440 forecast (without Murphy Plywood) accordingly. See the SAS output file for Roseburg's schedule 440 industrial forecast for the adjustment process.

This model will not be used for the spring 2025 forecast.

[7.86] $THM/C_{t,y,ROS447r} = \ schedule\ has\ died$

[7.86] Model notes:
1. Roseburg Forest Products moved to schedule 456 in November 2023. See the spring 2024 forecast file.

[7.87] $THM/C_{t,y,ROS456.i} = \ see\ notes\ to$ [7.66]

[7.87] Model notes:
1. In January 2019, C&D Lumber and Superior Studs left schedule 456 and moved to Roseburg schedule 440. At the end of February 2019, Swanson Lumber closed. For the June 2019 forecast, the Roseburg 456 UPC series was recalculated to remove the three companies. This means [7.87] is used to forecast UPC as if the three customers were never part of the UPC time series. See also model notes for [7.84]

Industrial Sector, Customers:

[7.88] $C_{t,y,ROS420.i} = \frac{1}{12}\sum_{j=1}^{12} C_{t-j}$

[7.88] Model notes:
1. Data starts September 2009. Excluding a February 2015 outlier, the customer count fluctuates between 1 and 2 without any clear trend or seasonality.

[7.89] $C_{t,y,ROS424.i} = \ schedule\ has\ died$

[7.89] Model notes:
1. Schedule appears to have died. No customers are currently being reported.

[7.90] $C_{t,y,ROS440.i} = \frac{1}{12}\sum_{j=1}^{12} C_{t-j}$

[7.90] Model notes:
1. Data starts January 2004. From January 2004 to December 2006, the customer count fluctuated between 11 and 13. Afterwards, the customer count stepped down to fluctuate between 1 and 7 until January 2019. There is no clear seasonality, but there is a slight downward trend that occurs in steps following prolonged periods of stability.
2. In January 2019, C&D Lumber and Superior Studs moved from schedule 456 to schedule 440. See also notes model notes for Roseburg schedule 456.

[7.91] $C_{t,y,ROS456.i} = C_{t-1}$

[7.91] Model notes:
1. Data starts January 2004. From January 2004 to November 2009, the customer count increased to 7 by December 2009. Afterwards, the customer count has fluctuated between 7 and 10 until January 2019. There is no clear trend or seasonality.
2. In January 2019, C&D Lumber and Superior Studs moved from Roseburg schedule 456 to schedule 440. At the end of 2019, Swanson Lumber closed. In the spring 2021 forecast, customers where set to 5 until August 2021 then the forecast switches to a moving average. Some small amount of Swanson's load may remain in 2020, but it will be negligible. See also notes model notes for Roseburg schedule 440.
3. In the spring 2024 forecast two schedule 447 special contract customers in Roseburg, Roseburg Forest Products and Murphy Plywood, were added to the customer count. See spring 2024 forecast sheet.

4c. Klamath Falls, OR Forecasting Models

The forecasting models for the Klamath Falls region (Klamath County) are given below for the residential, commercial, and industrial sectors:

Residential Sector, THM:

[7.92] $THM/C_{t,y,KLM410.r} = \ \beta_0 + \beta_1 HDD_{t,y}^{AVA} + \beta_2 (HDD_{t,y}^{AVA})^2 + \boldsymbol{\omega_{SD} D_{t,y}} + \omega_{OL} D_{Nov\ 2009=1} + \omega_{OL} D_{Dec\ 2009=1}\ +$

$$\omega_{OL}D_{Feb\ 2011=1} + \omega_{OL}D_{Mar\ 2017=1} + \omega_{OL}D_{Nov\ 2017=1} + \omega_{OL}D_{Mar\ 2018=1} + ARIMA\epsilon_{t,y}\ (3,0,0)(0,,0,0)_{12}\ for\ y = 2009\uparrow$$

[7.92] Model notes:
1. AHS is not included because its impact is unstable and not statistically significant.
2. RAP does not appear in the model because the coefficient's sign is unstable and statistically insignificant.
3. SC dummy controls for a step-down in UPC starting in June 2007.
4. In the spring 2024 this model was re-worked and the estimation period restricted to 2009↑ because the model diagnostics improved considerably. In particular it greatly reduced the number of outliers that need to be controlled for and improved the model's error term diagnostics.
5. In the fall 2024 forecast a new residential schedule was introduced—schedule 411—for multifamily gas users. Previously, the customers in schedule 411 where part of schedule 410. Because there is very little historical usage data available, the fall 2024 forecast had to estimate the load profile and annual usage to derive monthly UPC forecasts. This method is described in the fall 2024 forecast sheet for Oregon. See also the notes to [7.73]. The removal of schedule 411 customers from schedule 410 did not appear to significantly change the existing schedule UPC.

Residential Sector, Customers:

[7.93] $C_{t,y,KLM410+411.r} = \beta_0 + \beta_1 POP_{t,y,KLAMATH} + ARIMA\epsilon_{t,y}\ (6,1,0)(3,1,0)_{12}$

[7.93] Model notes:
1. POP is for Klamath County, OR.
2. The creation for schedule 411 in fall 2024 caused a step down in customers in schedule 410 because historically the schedule 411 customers where part of schedule 410. To control for this, the fall 2024 forecast continued to forecast schedules 410 and 411 together and then subtract off the forecast of the schedule 411 customers. In the fall 2024 forecast, the schedule 411 customer forecast was the average number of customers for June, July, and August 2024. Even though schedule 411 started in April 2024, the number customers estimated for April 2024 was too high and dropped in May 2024. It appears to have stabilized over the Jun-Aug period.

Commercial Sector, THM:

[7.94] $THM/C_{t,y,KLM420.c} = \beta_0 + \beta_1 HDD_{t,y}^{AVA} + \beta_2 (HDD_{t,y}^{AVA})^2 + \boldsymbol{\omega_{SD} D_{t,y}} + \omega_{SC} D_{Jan\ 2012\uparrow=1} +$
$\omega_{OL}D_{Dec\ 2005=1} + \omega_{OL}D_{Dec\ 2006=1} + \omega_{OL}D_{Mar\ 2008=1} + \omega_{OL}D_{Dec\ 2008=1} + \omega_{OL}D_{Nov\ 2009=1} +$
$\omega_{OL}D_{Dec\ 2009=1} + \omega_{OL}D_{Feb\ 2011=1} + \omega_{OL}D_{Feb\ 2018=1} + \omega_{OL}D_{Mar\ 2019=1} + + ARIMA\epsilon_{t,y}\ (10,0,0)(2,0,0)_{12}$

[7.94] Model notes:
1. SC dummies control for a step-down in UPC starting in January 2012.
2. Error structure is white noise, but not quite normally distributed.

[7.95] $THM/C_{t,y,KLM424.c} = \beta_0 + \beta_1 HDD_{t,y}^{AVA} + ARIMA\epsilon_{t,y}\ (3,0,0)(2,1,0)_{12}\ for\ y = 2010\uparrow$

[7.95] Model notes:
1. Model restricted to January 2019↑ because of a structural change in behavior starting in 2010. Inclusion of the 2004-2009 period results in base-load (peak) usage months that are considerably higher (lower) than the 2010↑ period. Current data series has a downward trend. It's unclear if UPC will stabilize or continue to fall in the future.

[7.96] $THM/C_{t,y,KLM440.c} = \frac{1}{N}\sum_{j=1}^{N}(THM/C_{t,y-j})\ for\ y-j\ = 2020\uparrow$
$up\ to\ the\ most\ recent\ month, then\ repeat\ forecast\ values$
[7.96] Model notes:
1. Model is restricted to 2020↑ because the schedule had a significant load increase starting in 2019. Schedule starts in February 2007.
2. UPC behavior continues to evolve and change. In particular 2023 saw a significant decline in both UPC and load. Additional data will be required before any econometric model can be estimated.

Commercial Sector, Customers:

[7.97] $C_{t,y,KLM420.c} = \beta_0 + \beta_1 C_{t,y,KLM410+411.r} + \boldsymbol{\omega_{SD} D_{t,y}} + ARIMA\epsilon_{t,y}\ (11,1,0)(1,0,0)_{12}$

[7.97] Model notes:

1. $C_{t,y,KLM410+411.r}$ are residential customers from residential schedule 410. They are being used as a forecast driver because of the historical positive correlation between residential and commercial customer growth. See Tables 5.1 and 5.2.

[7.98] $C_{t,y,KLM424.c} = C_{y-1} + (\hat{\beta}_0 + \hat{\beta}_1 \Delta EMP_{y-1,4County})$

[7.98] Model notes:
1. This model reflects a recommendation by Oregon staff in the 2016 rate case to include employment as an economic driver for schedule 424 commercial customers. The estimated equation in parenthesis reflects the regression estimated of $\Delta C_{y,KLM424.c} = \alpha_0 + \alpha_1 \Delta EMP_{y-1,4County} + \varepsilon_t$ using annual customer data since 2004. Annual data is used to smooth over the sometimes volatile changes in the monthly customer number. In addition, customer increases and decreases around the long-run trend tend to occur in steps. The combination of steps and month-to-month volatility creates significant economic problems when trying to model around the monthly data. For example, even with intervention variables, tests for error normality always indicated non-normal error terms with the use of monthly data.
2. $\Delta C_{y,KLM424.c}$ is the change in customers in year y (customer change between year y and y-1) and $\Delta EMP_{y-1,4County}$ is the change in total non-farm employment in Jackson, Josephine, Klamath, and Douglas counties in year y-1 (employment change between year y-1 and y-2). Staff originally suggested lagged total employment for Oregon, but the correlation between schedule 424 customers and employment for the four-county area is higher. The forecasted employment values for Jackson+Josephine County are derived from the employment growth forecasts used in the Jackson+Josephine County population forecast. The forecasts for Douglas and Klamath counties come from IHS. In IRP years, IHS forecasts for all counties will be used for the out years.
3. The annual forecast value for each year, F(·), is assumed to hold for each month of that year. That is: $F(C_{y,KLM424.c}) = F(C_{t,y,KLM424.c})$. Given the step-like behavior of the monthly series, this is a reasonable assumption.
4. The forecast and regressions for this schedule can be found in the Excel file folder “OR 4County Sch 424c Cus.”

[7.99] $C_{t,y,KLM440.c} = \frac{1}{12}\sum_{j=1}^{12} C_{t-j}$

[7.99] Model notes:
1. Data starts February 2007. The customer count fluctuates between 0 and 3 without any clear trend or seasonality.

Industrial Sector, THM:

[7.100] $THM/C_{t,y,KLM420.i} = \frac{1}{N}\sum_{j=1}^{N}(THM/C_{t,y-j})\ for\ y-j = 2022 \uparrow$ $up\ to\ the\ most\ recent\ month, then\ repeat\ forecast\ values$

[7.100] Model notes:
1. Due to declining customers and UPC (total load is moving toward zero), the econometric model was replaced by the simple averaging forecast shown above. An econometric approach was producing forecasts that were either negative or unlikely given the number of customers.

[7.101] $THM/C_{t,y,KLM424.i} = \beta_0 + \omega_{SC} D_{Sep\ 2015\uparrow=1} + \omega_{SC} D_{Sep\ 2020\uparrow=1} + \omega_{OL} D_{Sep\ 2019=1} + \omega_{OL} D_{Nov\ 2019=1} + \omega_{OL} D_{Nov\ 2020=1} + \omega_{OL} D_{Dec\ 2020=1} + ARIMA\epsilon_{t,y}\ (1,0,0)(2,1,0)_{12}\ for\ t,y = Februrary\ 2013 \uparrow$

[7.101] Model notes:
1. Model is restricted to February 2013 ↑ because schedule does not start until April 2009; in addition, April and May 2009 are extreme outliers due to billing issues. Also, starting in 20219 there was a change in seasonality with much higher loads in December and January. Therefore, the estimation period starts in February 2013 to account for both the lag in the seasonal error term; the April and May extreme outliers; and the structural change in December and January. In the underlying regression database (ORALLOTHERSCHEDULES), data prior to June 2009 are recorded as -9999 going back to January 2004. Failure to move up the estimation period will cause the -9999 values to be incorporated into the estimation.
2. SC dummy controls for a step-up in UPC starting in September 2015 and September 2020.
3. Spikes in winter load started in 2019; currently unclear if this is a permanent change. However, as of the spring 2022 forecast there are only 2 customers, and this is creating some issues with the forecast reflecting recent history. The model will need to be revaluated in future forecasts.

[7.102] $THM/C_{t,y,KLM440.i} = \frac{1}{12}\sum_{j=1}^{12} THM/C_{t-j}$

[7.102] Model notes:
1. Model is restricted to 2008 ↑ because of a significant change in seasonality compared to earlier periods. Repeated attempts to include the earlier periods resulted in model diagnostics that failed to show a stable error structure and forecasts that did not appear to fit recent patterns of behavior.
2. For months where UPC has been historically zero, the model will sometimes forecast negative values; these are set to zero in the final forecast.

3. Load activity stopped after November 2014 and then returned September 2021.

[7.103] $THM/C_{t,y,KLM456.i} = \ see\ notes\ to$ [7.66]

[7.103] Model notes:
1. The majority of 456 industrial customers in Klamath are in the wood products industry; therefore, housing starts, instead of industrial production, are used as a predictor.
2. SC dummy controls for a step-down in UPC starting in August 2005.
3. As of February 2016 Collins moved from schedule 447 to schedule 456. A control dummy was determined to be unnecessary because of this move.

Industrial Sector, Customers:

[7.104] $C_{t,y,KLM420.i} = \ \frac{1}{12}\sum_{j=1}^{12} C_{t-j}$

[7.104] Model notes:
1. Data starts December 2006. The customer count fluctuates between 4 and 9 without any clear trend or seasonality.

[7.105] $C_{t,y,KLM424.i} = \ \frac{1}{12}\sum_{j=1}^{12} C_{t-j}$

[7.105] Model notes:
1. Data starts April 2009. The customer count fluctuates between 1 and 4 without any clear trend or seasonality.

[7.106] $C_{t,y,KLM440.i} = \ \ 1\ if\ THM/C_{KLM,440.i} > 0$

[7.106] Model notes:
1. Data starts January 2004. The customer forecast is a derivative of the schedule's load forecast. Since 2007, typically only one customer has been registered.
2. Schedule appears to have died. See also [7.102] model notes.

[7.107] $C_{t,y,KLM456.i} = \ \frac{1}{12}\sum_{j=1}^{12} C_{t-j}$

[7.107] Model notes:
1. Data starts January 2004. Excluding outliers in April 2004 and January 2012, the customer count fluctuates between 3 and 5 without any clear trend or seasonality.

4d. La Grande, OR Forecasting Models

The forecasting models for the La Grande region (Union County) are given below for the residential, commercial, and industrial sectors:

Residential Sector, THM:

[7.108] $THM/C_{t,y,LaG410.r} = \ \theta_0 + \theta_1 HDD_{t,y}^{AVA} + \theta_2 (HDD_{t,y}^{AVA})^2 \ + \boldsymbol{\omega_{SD} D_{t,y}} + \omega_{SC} D_{Jan\ 2005\uparrow=1} + \omega_{SC} D_{Jan\ 2009\uparrow=1} + \omega_{OL} D_{Jan\ 2007=1} + \ \omega_{OL} D_{Feb\ 2007=1} + \omega_{OL} D_{Dec\ 2008=1} + \ \omega_{OL} D_{Dec\ 2016=1} + \omega_{OL} D_{Feb\ 2017=1} + ARIMA\epsilon_{t,y}\ (12,0,0)(0,0,0)_{12}$

[7.108] Model notes:
1. RAP does not appear in the model because the coefficient's sign is unstable and statistically insignificant.
2. The variable lnT reflects a general control for multiple factors lowering UPC, including the impact of RAP.
3. SC dummies control for a step-down in UPC starting in January 2005 and 2009.
4. AHS is not included because its impact is unstable and not statistically significant.
5. In the fall 2024 forecast a new residential schedule was introduced—schedule 411—for multifamily gas users. Previously, the customers in schedule 411 where part of schedule 410. Because there is very little historical usage data available, the fall 2024 forecast had to estimate

the load profile and annual usage to derive monthly UPC forecasts. This method is described in the fall 2024 forecast sheet for Oregon. See also the notes to [7.73]. The removal of schedule 411 customers from schedule 410 did not appear to significantly change the existing schedule UPC.

Residential Sector, Customers:

[7.109] $C_{t,y,LaG410.r} = \theta_0 + \theta_1 POP_{t,y,UNION} + \omega_{OL} D_{Oct\ 2004=1} + \omega_{OL} D_{Jul\ 2006=1} + \omega_{OL} D_{Dec\ 2009=1} + ARIMA\epsilon_{t,y}\ (9,1,0)(3,1,0)_{12}$

[7.109] Model notes:
1. POP is population for Union County, OR. The coefficient on POP is positive, but no longer statistically significant.
2. The creation for schedule 411 in fall 2024 caused a step down in customers in schedule 410 because historically the schedule 411 customers where part of schedule 410. To control for this, the fall 2024 forecast continued to forecast schedules 410 and 411 together and then subtract off the forecast of the schedule 411 customers. In the fall 2024 forecast, the schedule 411 customer forecast was the average number of customers for June, July, and August 2024. Even though schedule 411 started in April 2024, the number customers estimated for April 2024 was too high and dropped in May 2024. It appears to have stabilized over the Jun-Aug period.

Commercial Sector, THM:

[7.110] $THM/C_{t,y,LaG420.c} = \theta_0 + \theta_1 HDD_{t,y}^{AVA} + \theta_2 (HDD_{t,y}^{AVA})^2 + \boldsymbol{\omega_{SD} D_{t,y}} + \omega_{SC} D_{Jan\ 2005\uparrow=1} + \omega_{OL} D_{Feb\ 2007=1} + \omega_{OL} D_{Dec\ 2008=1} + \omega_{OL} D_{Feb\ 2017=1} + \omega_{OL} D_{Dec\ 2016=1} + \omega_{OL} D_{May\ 2017=1} + ARIMA\epsilon_{t,y}\ (12,0,0)(0,0,0)_{12}$

[7.110] Model notes:
1. SC dummy controls for a significant step-down in UPC in January in 2005.
2. Similar to this schedule in the other city areas, the forecast baseload is slightly higher than what has been observed in recent actuals. It's unclear why given the model's diagnostics look reasonable. This will need to be evaluated in the next forecast.

[7.111] $THM/C_{t,y,LaG424.c} = \theta_0 + \theta_1 HDD_{t,y}^{AVA} + \boldsymbol{\omega_{SD} D_{t,y}} + \omega_{SC} D_{Jul\ 2020\uparrow=1} + \omega_{OL} D_{Dec\ 2010=1} + \omega_{OL} D_{Jan\ 2011=1} + ARIMA\epsilon_{t,y}\ (1,0,0)(0,0,0)_{12}\ for\ t,y = June\ 2010\uparrow$

[7.111] Model notes:
1. Model is restricted to June 2010 ↑ because of a significant change in seasonality compared to earlier periods. A statistical test for moving seasonality was applied using the CensusX12 procedure in SAS. The test confirmed the presence of moving seasonality at the 1% level.
2. SC dummy controls for a step-down in UPC starting in July 2020.

[7.112] $THM_{t,y,LaG444.c} = schedule\ has\ died$

[7.112] Model notes:
1. Model is restricted to 2011 ↑ because the schedule does not start until 2011.
2. No identifiable trend in the data.
3. Virtually no load activity since September 2018. Load set to zero in the spring 2021 forecast.

[7.113] $THM/C_{t,y,LaG440.c} = \theta_0 + \theta_1 HDD_{t,y}^{AVA} + \boldsymbol{\omega_{SD} D_{t,y}} + \omega_{OL} D_{Sep\ 2013=1} + \omega_{OL} D_{Nov\ 2013=1} + \omega_{OL} D_{Jul\ 2014=1} + \omega_{OL} D_{Aug\ 2014=1} + \omega_{OL} D_{Oct\ 2014=1} + \omega_{OL} D_{Nov\ 2014=1} + \omega_{OL} D_{Jan\ 2015=1} + \omega_{OL} D_{Jul\ 2015=1} + \omega_{OL} D_{Sep\ 2015=1} + \omega_{OL} D_{May\ 2017=1} + \omega_{OL} D_{Dec\ 2023=1} + ARIMA\epsilon_{t,y}\ (11,0,0)(0,0,0)_{12}\ for\ t,y = September\ 2009\uparrow$

[7.113] Model notes:
1. Model is restricted to September 2009↑ because schedule does not start until February 2007. The estimation period starts in September 2009 to account for the lag in the error term and a change in seasonality. In the underlying regression database (ORALLOTHERSCHEDULES), data prior to May 2007 is recorded as -9999 going back to January 2004. Failure to move up the estimation period will cause the -9999 values to be incorporated into the estimation.

[7.114] $THM/C_{t,y,LaG456.c} = \theta_0 + \theta_1 HDD_{t,y}^{AVA} + \omega_{SC} D_{Jan\ 2006\uparrow=1} + \omega_{SC} D_{Jul\ 2013\uparrow=1} + \omega_{OL} D_{Apr\ 2004=1} + \omega_{OL} D_{Jan\ 2006=1} + \omega_{OL} D_{Mar\ 2009=1} + \omega_{OL} D_{May\ 2013=1} + \omega_{OL} D_{Jun\ 2013=1} + \omega_{OL} D_{Dec\ 2016=1} + ARIMA\epsilon_{t,y}\ (0,0,0)(3,0,0)_{12}$

[7.114] Model notes:
1. This reflects only one customer—Eastern Oregon University. Base load forecast months are often negative; these are set to zero in the

forecast spreadsheet. This often occurs when base load months show zero usage. If the forecast is negative, the average of the previous three years of actuals (for the negative forecast months) are used to replace the actuals.
2. SC dummies controls for step-downs in UPC starting in January 2006 and July 2013.

Commercial Sector, Customers:

[7.115] $C_{t,y,LaG420.c} = \theta_0 + \boldsymbol{\omega_{SD} D_{t,y}} + \omega_{SC} D_{Jan\ 2016\uparrow =1} + \omega_{SC} D_{Nov\ 2019\uparrow =1} + \omega_{OL} D_{Jul\ 2005\ =1} + \omega_{OL} D_{Dec\ 2008\ =1} + \omega_{OL} D_{Mar\ 2011\ =1} + \omega_{OL} D_{May\ 2011\ =1} + \omega_{OL} D_{Nov\ 2021\ =1} + ARIMA\epsilon_{t,y}\ (12,1,0)(0,0,0)_{12}$

[7.115] Model notes:
1. $C_{t,y,LaG410.r}$, residential customers from residential schedule 410, are no longer used as a forecast driver. The estimated coefficient on $C_{t,y,LaG410.r}$ was no longer statistically significant and its sign flips between positive and negative, depending on the form of the model. POP for union county was also tried as a driver, but had the same issues as $C_{t,y,LaG410.r}$.
2. SC dummies control for step-ups in customers starting in January 2016 and November 2019.

[7.116] $C_{t,y,LaG424.c} = \frac{1}{12}\sum_{j=1}^{12} C_{t-j}$

[7.116] Model notes:
1. Data starts January 2007. The customer count fluctuates between 2 and 4 without any clear trend or seasonality. Changes in the customer count appear as steps after prolonged periods of stability.

[7.117] $C_{t,y,LaG444.c} = \frac{1}{N}\sum_{j=1}^{N} C_{t,y-j}\ \ if\ (THM_{t,y})_{Lag,444.c} \geq 0$

[7.117] Model notes:
1. Data starts September 2011. The customer forecast is a derivative of the schedule's load forecast.
2. Schedule appears to have died. Almost not load activity since September 2018.

[7.118] $C_{t,y,LaG440.c} = \frac{1}{12}\sum_{j=1}^{12} C_{t-j}$

[7.118] Model notes:
1. Data starts February 2007. The customer count fluctuates between 1 and 5 without any clear trend or seasonality.

[7.119] $C_{t,y,LaG456.c} = C_{t-1}$

[7.119] Model notes:
1. Excluding outliers in April 2004, December 2004, and May 2013, only one customer has been registered since January 2004.

Industrial Sector, THM:

[7.120] $THM/C_{t,y,LaG440.i} = \theta_0 + \omega_{SC} D_{Jan\ 2005\uparrow=1} + \omega_{SC} D_{Jan\ 2021\uparrow=1} + \omega_{OL} D_{Sep\ 2008\ =1} + \omega_{OL} D_{Sep\ 2017\ =1} + \omega_{OL} D_{Sep\ 2014\ =1} + \omega_{OL} D_{May\ 2019\ =1} + ARIMA\epsilon_{t,y}\ (11,0,0)(2,1,0)_{12}$

[7.120] Model notes:
1. IP was removed as a forecast driver in the spring 2022 forecast because the coefficient had become negative and insignificant.
1. SC dummies control for steps-down in UPC starting in January 2005 and January 2021.
2. Error structure white noise, but not quite normally distributed.

[7.121] $THM/C_{t,y,LaG444.i} = \frac{1}{N}\sum_{j=1}^{N}(THM/C_{t,y-j})\ \ for\ y - j = 2004\uparrow$ $up\ to\ the\ most\ recent\ month, then\ repeat\ forecast\ values$

[7.121] Model notes:
1. Model was moved to an average based forecast due to poor regression diagnostics due to frequent extreme outliers and months with zero load.

[7.122] $THM/C_{t,y,LaG456.i} = \text{ see notes to }$ [7.66]

[7.122] Model notes:
1. Housing starts was found to be an unstable predictor; therefore, it is not included in the model.
2. Multiple modeling attempts using the entire time series produced forecasts with seasonality that did not match the seasonality since the end of 2013. A statistical test for moving seasonality was applied using the CensusX12 procedure in SAS. The test confirmed the presence of moving seasonality at the 5% level. As a result, the model is restricted to July 2013↑.

Industrial Sector, Customers:

[7.123] $C_{t,y,LaG420.i} = \frac{1}{12}\sum_{j=1}^{12} C_{t-j}$

[7.123] Model notes:
1. Since these customers appeared approximately, there has been no load activity. As a result, they have never been included in a forecast prior to fall 2021; it was assumed this schedule was simply a revenue reporting error. However, subsequent research of billing activity indicates the customers are paying fixed charges. The current forecast assumes no load over the forecast horizon.

[7.124] $C_{t,y,LaG440.i} = \frac{1}{12}\sum_{j=1}^{12} C_{t-1}$

[7.124] Model notes:
1. Historically, even in the presence of seasonality, customer count can be highly erratic. Regression models produced poor diagnostics. As a result, a historical monthly average is used as the forecast. However, as of the spring 2022 forecast, the customer count has stabilized at 5. It is unclear if this will be the case in the long-term.
2. Restricted to 2012 ↑ because of a significant change in behavior starting in 2012.

[7.125] $C_{t,y,LaG444.i} = \frac{1}{N}\sum_{j=1}^{N} C_{t,y-j} \; for \; y-j = 2018 \uparrow$
$up \; to \; the \; most \; recent \; month, then \; repeat \; forecast \; values.$

[7.125] Model notes:
1. Even in the presence of seasonality, customer count can be highly erratic. Regression models produced poor diagnostics and required many OL dummies. As a result, a historical monthly average is used as the forecast.
2. Restricted to 2018 ↑ because of significant behavior changes starting in 2014.

[7.125] $C_{t,y,LaG456.i} = \frac{1}{12}\sum_{j=1}^{12} C_{t-j}$

[7.125] Model notes:
1. Data starts January 2004. Since November 2004, only three customers have been registered.

4e. Calendar Core Load for Unbilled Calculation

This section discusses the Company's method to calendarize the retail load forecast. This method uses the historic average monthly difference between calendar load and retail load to convert the retail forecast to an approximate calendar forecast of firm and interruptible load, called "Core Load."

The historical Core Load data come from Avista's Nucleus database and reflects aggregated city-gate firm and interruptible load, which is not broken out by revenue class. The first step derives the calendar Core Load forecast from the core retail load forecast (firm plus interruptible) using the historic average monthly difference between the calendar Core Load and retail core load. The next step is to approximate the firm-only calendar load. This is done by subtracting the retail interruptible forecast for the same 12-month period.

To understand how the method is applied, it is necessary to understand the difference between core load (L_C) and billed load. It is useful to first to discuss how Avista measures L_C. L_C is defined as:

[7.144] $L_{C,t,y} \equiv L_{G,t,y} - L_{T,t,y}$

Where $L_{G,t,y}$ is the total load measured at all Avista city gates in ***calendar*** month t, year y and $L_{T,t,y}$ is all metered transport customer load in ***calendar*** month t, year y. The subtraction of $L_{T,t,y}$ means the $L_{C,t,y}$ reflects all firm ($L_{F,t,y}$) and interruptible load ($L_{T,t,y}$).[1] Therefore:

[7.145] $L_{C,t,y} \equiv L_{F,t,y} + L_{I,t,y} \equiv L_{G,t,y} - L_{T,t,y} \Rightarrow L_{G,t} \equiv L_{F,t,y} + L_{I,t,y} + L_{T,t,y}$

Because of losses and other unaccounted for load (l_y) resulting from moving the gas through the distribution process, annual core load in year y ($L_{C,y}$) with be slightly larger than annual core billed load in year y ($L_{B,y}$). Therefore:

[7.146] $L_{B,y} = L_{C,y} - l_y$

Note the switch to an annual rather than monthly time-domain reflects that on a monthly basis $L_{B,t,y}$ is on a non-calendar billing cycle. Which means in any given month the difference between $L_{C,t,y}$ and $L_{B,t,y}$ is defined as net-unbilled load ($L_{NUB,t,y}$):

[7.147] $L_{NUB,t,y} = L_{C,t,y} - L_{B,t,y}$

For example, if t = calendar December for $L_{C,t}$, then $L_{B,t}$ for December will actually include part of November and December. The lack of time synchronicity creates $L_{NUB,t}$. Analysis of this historical data shows that over a 12-month period:[2]

[7.148] $\frac{\sum_{t=1}^{12} L_{NUB,t}}{12} \approx 0$

Given the discussion above, it is now useful to discuss the approach used for forecasting unbilled load to convert the billed load forecast to a calendar basis. This can be done by converting the core billed load forecast into an equivalent core load forecast using the historical difference between the two. Start with the monthly ratio between the core load and the non-calendar billed load for month t in year y in jurisdiction k:

[7.149] $\omega_{t,y,k} \equiv \frac{L_{C,t,y,k}}{L_{B,t,y,k}}\ for\ k = OR\ or\ WA - ID$

The ratio $\omega_{t,y,k}$ can be averaged for month t over N years:

[1] Firm load includes interdepartmental schedules.

[2] Based on a distributional analysis of a 12 month moving average, the calendar 12 month average and the calendar sum of $L_{NUB,t}$ where not statistically different from zero.

[7.150] $\bar{\omega}_{t,k} = \frac{\sum_{y}^{N} \omega_{t,y,k}}{N} \quad for\ y = 1, \ldots, N\ years$

The average $\omega_{t,y,k}$ can be applied to the forecast for $L_{B,t,y,k}$, to generate a forecast for $L_{C,t,y,k}$:

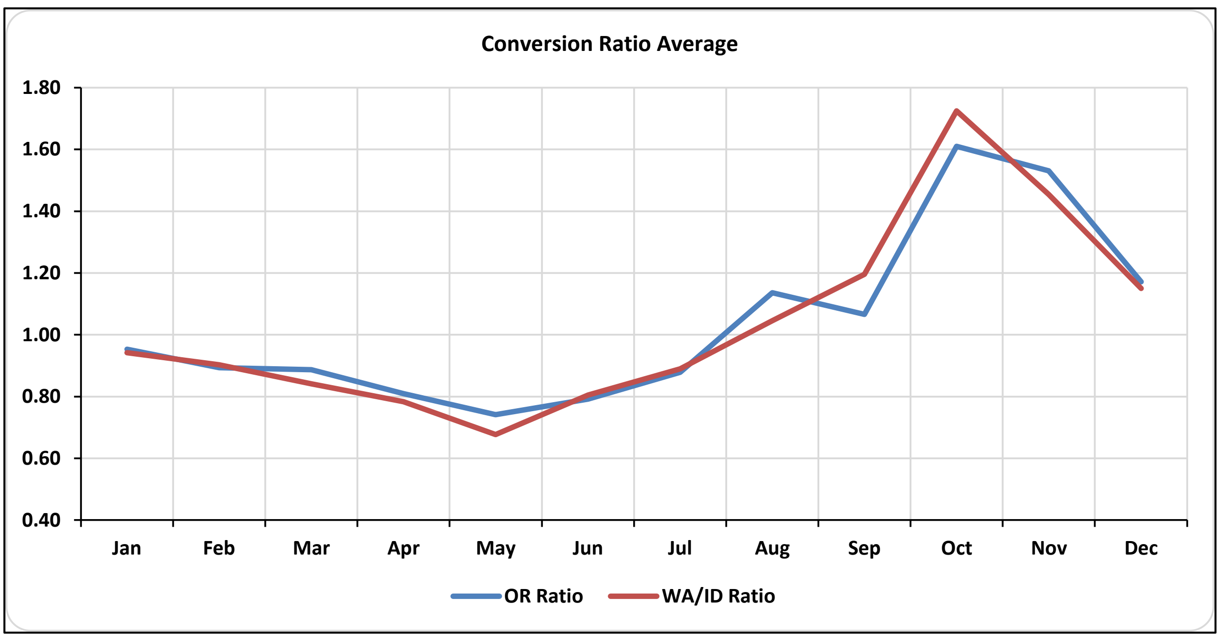
[7.151] $F(L_{C,t,y_c+j,k}) = \bar{\omega}_{t,k} \cdot F\left(L_{B,t,y_c+j,k}\right) \ for\ j = 1, \ldots, 5$

With the forecasted series for $L_{C,t,y,k}$, a forecast for $L_{NUB,t,y,k}$:

[7.152] $F(L_{NUB,t,y_c+j,k}) = F(L_{C,t,y_c+j,k}) - F(L_{B,t,y_c+j,k}) = F(L_{B,t,y_c+j,k}) \cdot (\bar{\omega}_{t,k} - 1)$

The average shown by [7.150] is essentially equivalent to the values that would be produced by a linear regression model with monthly dummy variables and no ARMIA error correction. The figure below shows the monthly average of $\omega_{t,y,k}$ for the 2008-2023 time period used for the Fall 2024 forecast:

<u>Average Ratio of Monthly Calendar Core Load to Non-Calendar Billed Load, $\bar{\omega}_{t,k}$</u>



The average goes back to 2008 because of data limitations related to the historical values of interdepartmental load.