Contents

TO:	**Center for Sustainable Energy, Marketing and Outreach Team**
FROM:	**Francis Alvarez, Keir Havel, James Tamerius, Colin Evans, John Gartner, Center for Sustainable Energy**
Date:	**10/13/2022**
Subject:	**Analysis of CVRP Leading Indicators and Their Effect on Forecasts**

Summary

Center for Sustainable Energy (CSE) has evaluated market, demographic, and economic indicators that impact the rate of electric vehicle (EV) adoption. CSE reviewed the data during the period in which the COVID-19 pandemic influenced the EV and larger automotive markets, as well as the overall economy in California. From this analysis CSE has developed forecasts that predict the number of future vehicles receiving rebates through the Clean Vehicle Rebate Project (CVRP), which helps gauge California's progress in achieving the state's goals for zero-emission vehicles (ZEVs) on the road.[1] This analysis aimed to identify variables that improve CSE's predictions of light-duty vehicle market trends, allowing CSE to better forecast CVRP participation.

The hundreds of indicators originally considered were reduced based on their importance in relation to the number of program rebates; the remaining indicators we analyzed include the unemployment rate, the number of public charging ports, and the number of households with income over $200,000. Additionally, CSE analyzed the number of vehicle models eligible for rebates during each month and categorized them into low-, medium-, and high-rebated groups as indicators in the forecast models. The results from our analysis of these indicators are:

- Households with higher income were statistically significant in a few models, but it is best to exclude this indicator in future models as subsequent program changes include lowered income caps, which would deem this group of applicants as ineligible.
- The unemployment rate correlates highly with CVRP rebates due to the overall timing of the program start date and initial growth, but the unemployment rate does not improve rebate forecast accuracy and, therefore, should be excluded in future models.
- Public charging availability improves forecast accuracy in some models, but the improvement coincides with the endogenous relationship between charging and the number of rebated vehicles. Public charging as an indicator should also be excluded in future models.
- The low-, mid-, and high-volume vehicle availability indicators improved forecast accuracy, but the mid- and high-volume vehicle groups had the largest impact on improving the forecasts. Overall, the high-volume indicator is the preferred indicator to include as it has the greatest contribution to the number of rebates.

Approach

Based on our understanding of the drivers of the light-duty electric vehicle market and the CVRP, we identified numerous variables that had the potential to improve CVRP rebate forecasts. We identified

[1] https://www.cpuc.ca.gov/zev/

publicly available data sources, and then selected the variables with a sufficient quality level required to improve the models. We then evaluated the impact of each of these variables on the predictive skill of the models. We then developed a set of final models that consists of the baseline model plus leading indicator variables that improved the forecast models.

Model Description

The forecast models we develop are at a monthly resolution and are for the entire state of California. Specifically, we forecast approved monthly CVRP rebate applications from March 2010 through July 2021. Each model is generated using Facebook's Prophet[2] forecasting library to increase forecast accuracy and simplify implementation. This modeling framework decomposes the forecast into three components: trend, seasonality, and holidays. Leading indicator variables were added to the forecast model as "regressors" to improve the skill of the model.

To quantify the impact of these regressors, we first conducted baseline forecasts (with the regressors excluded) and then evaluated the performance of these forecasts on a few metrics. We then incorporated the regressors into the models, reevaluated the forecast based on the same metrics, and compared the resulting metrics to the baseline to determine if there was improved forecast accuracy. Specifically, the models are categorized by technology type (battery electric vehicles [BEVs] versus plug-in hybrid vehicles [PHEVs]) and by applicant rebate type.[3] This allows us to forecast disparate trends across each category more accurately than a single forecast model. Further, this allows us to provide category-specific information when communicating forecast results. In this analysis, we further disaggregated the data by separating Tesla and non-Tesla BEV rebates (see Figure 18 and Figure 17, respectively). We did this due to the large share of rebates that Tesla vehicles represent, their distinct quarterly distribution pattern, and the large increases in the number of rebates associated with some of their vehicle releases.

The six different combinations of technology type and applicant rebates that we modeled are listed in Table 1, and their respective plots are available in the Baseline Forecast Models section in the Appendix. Each baseline forecast includes a waitlist regressor to adjust forecasts for periods where the program was suspended due to insufficient funds. We also included a seasonal regressor to adjust the model for any cyclical trends throughout the program.

Table 1. The combinations of technology and rebate type were used in our forecasting analysis.

Technology Type	Applicant Rebate Types
Non-Tesla Specific BEVs	Individual Standard Rebates
Non-Tesla Specific BEVs	Individual Increased Rebates
Tesla Specific BEVs	Individual Standard Rebates
Tesla Specific BEVs	Individual Increased Rebates
PHEVs	Individual Standard Rebates
PHEVs	Individual Increased Rebates

[2] https://facebook.github.io/prophet/

[3] We only assess models for individual standard and individual increased rebates in this analysis, but the approach can be easily applied to fleet rebates in the future.

Leading Indicator Data

To improve the baseline models, we identified numerous leading indicator variables for inclusion in the forecasting models. These variables represent economic, demographic, vehicle, and environmental factors that could reasonably be expected to impact EV adoption and CVRP rebates. The variables were grouped into distinct topics and are summarized in Table 2. The table also depicts the number of variables collected on several themes, with additional fields describing each dataset's spatial and temporal characteristics. The variables that were not spatially comprehensive (i.e., were not available across all of California), timely (i.e., significant lag time before data becomes available), or temporally coarse (i.e., were at an annual or coarser level) were not evaluated further. Moreover, we then evaluated each variable using a random forest (multiple decision tree) model[4] where the top three variables were selected for further assessment.

Table 2. Topics present in data, sorted by utility. Temporal unit of analysis is abbreviated as "Y" for yearly and "M" for monthly. Note that the FEG data includes statewide data in addition to county-level data, so it is labeled as containing 59 counties. Multiple rows with the same topic and source value but different min or max dates are variables within the same category but do not have the same date windows of available information.

Topic	Source	Temporal Resolution	Min Date	Max Date	Number of Counties Covered	Variable Count
General Demographics	ACS	Y	6/30/2005	6/30/2019	41	49
EV Charging	EV Atlas, NREL	M	8/31/1995	1/31/2021	58	9
Rebate Applicant Characteristics	CVRP	M	3/31/2010	1/31/2021	56	13
Rebate Applicant Characteristics	CVRP	M	4/30/2010	1/31/2021	57	37
Rebate Applicant Characteristics	CVRP	M	5/31/2010	1/31/2021	56	1
Rebate Applicant Characteristics	CVRP	M	3/31/2016	1/31/2021	58	1
Rebate Applicant Characteristics	CVRP	M	4/30/2016	1/31/2021	57	1
Rebate Applicant Characteristics	CVRP	M	1/31/2018	1/31/2021	8	3

[4] https://link.springer.com/article/10.1023/A:1010933404324. A decision tree is a model that can classify or predict based on set decision rules, similar to that of a flow chart. A random forest model implements multiple decision trees and pools their decisions relying on the fact that the collective decision is often more accurate than an individual decision tree. As we applied the random forest model across the potential regressors, we ranked the variables based on their resulting mean decrease in impurity and permutation importance. Mean decrease in impurity is a measure of how much removing a variable affects the accuracy of the model, which helps gauge the importance of that variable as a predictor. Permutation importance is a similar method but shuffles the order of variables that are utilized for classification and then reevaluates the importance of each predictor.

Rebate Applicant Characteristics	CVRP	M	3/31/2018	12/31/2020	3	3
Household Income	ACS	Y	6/30/2005	6/30/2019	41	18
Vehicle Counts	IHS	M	3/31/2010	1/31/2021	58	5
Vehicle Counts	IHS	M	4/30/2010	1/31/2021	57	1
Application Type or Status	CVRP	M	3/31/2010	1/31/2021	58	26
Application Type or Status	CVRP	M	4/30/2010	1/31/2021	57	26
Application Lags	CVRP	M	3/31/2010	1/31/2021	57	9
Application Lags	CVRP	M	4/30/2010	1/31/2021	57	9
Household Type	ACS	Y	6/30/2005	6/30/2019	40	9
Household Type	ACS	N/A	6/30/2019	6/30/2019	41	2
Housing Cost	ACS	Y	6/30/2005	6/30/2019	40	16
Housing Cost	ACS	Y	6/30/2015	6/30/2019	41	3
Unemployment	EDD	M	1/31/2000	11/30/2020	58	1
Unemployment	EDD	M	1/31/2007	11/30/2020	58	2
Applicant Utility	CVRP	M	3/31/2010	1/31/2021	58	52
Applicant Utility	CVRP	M	4/30/2010	1/31/2021	57	52
Tax Credit	FEG	M	1/31/2010	1/31/2022	59	9
CVRP Eligible Vehicles	CVRP	M	3/31/2010	4/30/2021	58	79
Cancellation Reasons	CVRP	M	3/31/2010	1/31/2021	58	21
Rebate Vehicle Category	CVRP	M	3/31/2010	1/31/2021	58	6
Rebate Vehicle Category	CVRP	M	4/30/2010	1/31/2021	57	6
Vehicles Used in Commute	ACS	Y	6/30/2005	6/30/2019	41	3
Applicant Demographics	CVRP	M	3/31/2010	1/31/2021	58	12
Applicant Demographics	CVRP	M	4/30/2010	1/31/2021	57	10
Per Capita Income	ACS	Y	6/30/2005	6/30/2019	41	1
CO2 Emissions	EIA	M	1/31/2000	12/31/2022	59	1
GINI and GDP	ACS	Y	6/30/2006	6/30/2019	41	1
GINI and GDP	EIA	M	1/31/2000	12/31/2022	59	1
Fuel Prices	EIA	M	6/30/2000	12/31/2020	58	1
Fuel Prices	EIA	M	1/31/2001	10/31/2020	58	1
Income Verification	CVRP	M	3/31/2010	1/31/2021	58	2
Population	ACS	Y	6/30/2006	6/30/2018	40	1

Vehicle Miles Travelled	BTS	N/A	6/30/2009	6/30/2017	58	2

We also included the number of program-eligible vehicles by month as a variable to further assess since this showed a demonstrable impact on rebate volume. Each program-eligible vehicle was assigned to a low-, medium-, or high-volume class since each vehicle differs in terms of rebate volume (more information on how the program-eligible vehicles were classified is provided in the Vehicle Volume Clustering section in the Appendix). A table of the leading indicators we analyzed is presented in Table 3.

Table 3. Summary table of the leading indicator regressors that we analyzed with a brief description of the regressor.

Forecast Group	Regressor Name	Description
Base Regressors	Waitlist	Used to model periods in the CVRP where rebate processing was temporarily delayed.
Leading Indicator Regressors	Household Income	The number of households in California with income over $200,000.
Leading Indicator Regressors	Public Charging	The total number of public charging ports available in California.
Leading Indicator Regressors	Unemployment Rate	The unemployment rate in California.
Leading Indicator Regressors	Low-volume Vehicles	Vehicle models that have a low impact on the number of rebates in CVRP.
Leading Indicator Regressors	Mid-volume Vehicles	Vehicle models that have a medium impact on the number of rebates in CVRP.
Leading Indicator Regressors	High-volume Vehicles	Vehicle models that have a high impact on the number of rebates in CVRP.

Assessing Leading Indicators Impacts on Forecasts

We used several methods to evaluate the impact of each leading indicator on the model's predictive skill. The first method is cross-validation, which begins by designating a portion of the initial data as "training" data based on a specific cutoff date. Once the training data were established, we forecasted six months and evaluated the results with the available observed data. After the first forecast was produced, the training data were then extended by two months, and another six-month forecast was created. This process then was repeated until there were no observed data remaining. The preliminary training data begins from March 2010 through December 2016, and the 13 six-month forecasts are produced and evaluated. A figure containing the CVRP rebate data and the dates that are used for cross-validation[5] can be found in the Cross-Validation Dates section in the Appendix.

The second method to assess the impact of adding each leading indicator to the model is by computing uncertainty bounds through Markov chain Monte Carlo (MCMC) sampling.[6] A lower and upper bound for the coefficient of the leading indicator component is provided, which allows us to determine

[5] https://robjhyndman.com/hyndsight/tscv/

[6] http://stat.wharton.upenn.edu/~stjensen/stat542/lecture14.mcmchistory.pdf

whether each regressor is statistically significant at the 0.05 level. The lower and upper coefficient bounds also provide an approximation of the effect of the regressor in the form of number of rebates.[7]

Models were also assessed using classical metrics of model fit, including root mean square error (RMSE) and mean absolute percentage error (MAPE). RMSE measures the absolute distance between the model's predictions and observed data, while MAPE is more relative and measures the percent difference between the prediction and the observed data.

Results

Non-Tesla BEV Individual Standard Rebates

The following indicators improved the forecast model significantly for non-Tesla BEV standard rebates: household income, mid-volume vehicles, and high-volume vehicles. While household income was considered significant, the predicted impact was inversely related to what is intuitively expected. For example, as the number of households with an income over $200,000 increases, the number of rebates is expected to decline.

The expected impact of mid-volume vehicles on rebate volume during peak release is 286 per month, while high-volume vehicles are expected to increase rebates by 697 per month. In Table 5, when comparing both the mid-volume and high-volume regressors to the baseline forecast model, the RMSE and MAPE values are smaller, indicating that each regressor improved the forecast accuracy within the cross-validation testing period.

In the scenario of adding both mid-volume and high-volume vehicles into the model, both regressors remain significant. The estimated rebate effect of the mid-volume and high-volume vehicles both decreased, with mid-volume rebates decreasing from 286 to 227 monthly rebates and high-volume rebates reducing from 697 to 602 rebates. Although there was a decrease in the estimated impact of each indicator, the two-vehicle volume model also improved the cross-validation metrics relative to the baseline model and more parsimonious models. Based on the resulting metrics, we recommend using a mid- and high-volume regressor in future forecasting predictions.

Table 4. MCMC sampling results for non-Tesla BEV individual standard rebates. The addition of a new vehicle and how it will affect rebates during its peak release is provided below the coefficient value in parentheses.

Forecast Group	Regressor Combinations	Regressor Name	Lower Coefficient	Coefficient	Upper Coefficient	Significant Regressor
Single Leading Indicator Models	Household Income	Household Income	-0.87	-0.46	-0.09	TRUE
Single Leading Indicator Models	Public Charging	Public Charging	-0.02	0.02	0.07	FALSE

[7] Note that when interpreting the coefficients for eligible vehicles some additional transformations must be performed to directly interpret how one new vehicle will affect the number of rebates (transformations are explained for converting eligible vehicle volume in the Vehicle Volume Clustering section in the Appendix).

Single Leading Indicator Models	Unemployment Rate	Unemployment Rate	-128.99	-7.84	110.04	FALSE
Single Leading Indicator Models	Low-volume Vehicles	Low-volume Vehicles	-120.48 (-162.49)	-52.0 (-70.13)	13.64 (18.4)	FALSE
Single Leading Indicator Models	Mid-volume Vehicles	Mid-volume Vehicles	63.3 (90.55)	199.95 (286.04)	351.88 (503.37)	TRUE
Single Leading Indicator Models	High-volume Vehicles	High-volume Vehicles	286.59 (313.32)	637.49 (696.95)	1006.36 (1100.22)	TRUE
Mixed Leading Indicator Models	Mid- & High-volume Vehicles	Mid-volume Vehicles	4.57 (6.54)	158.45 (226.67)	303.34 (433.93)	TRUE
Mixed Leading Indicator Models	Mid- & High-volume Vehicles	High-volume Vehicles	170.32 (186.2)	550.24 (601.56)	907.64 (992.29)	TRUE

Table 5. Cross-validation scores for non-Tesla BEV individual standard rebates.

Forecast Group	Forecast Version	RMSE	RMSE Rank	MAPE	MAPE Rank
Base Model	Base	301.947	6	0.22	7
Single Leading Indicator Models	Household Income	247.954	2	0.181	2
Single Leading Indicator Models	Public Charging	302.132	7	0.218	6
Single Leading Indicator Models	Unemployment Rate	314.334	8	0.238	8
Single Leading Indicator Models	Low-volume Vehicles	279.265	5	0.19	4
Single Leading Indicator Models	Mid-volume Vehicles	276.756	4	0.195	5
Single Leading Indicator Models	High-volume Vehicles	251.584	3	0.187	3
Mixed Leading Indicator Models	Mid- & High-volume Vehicles	235.66	1	0.163	1

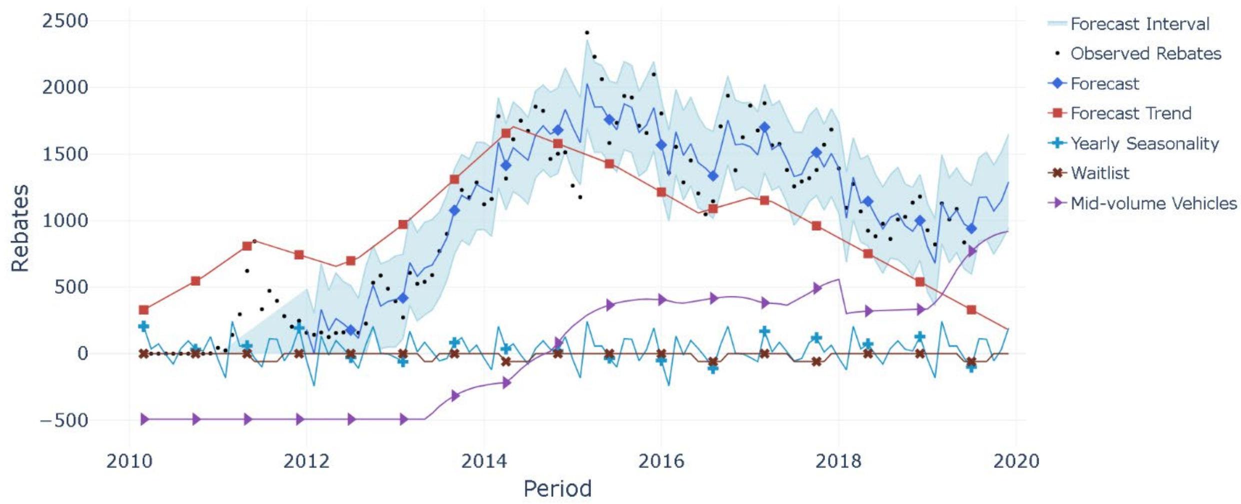

Figure 1. Non-Tesla BEV individual standard rebate model with a mid-volume vehicle model regressor.

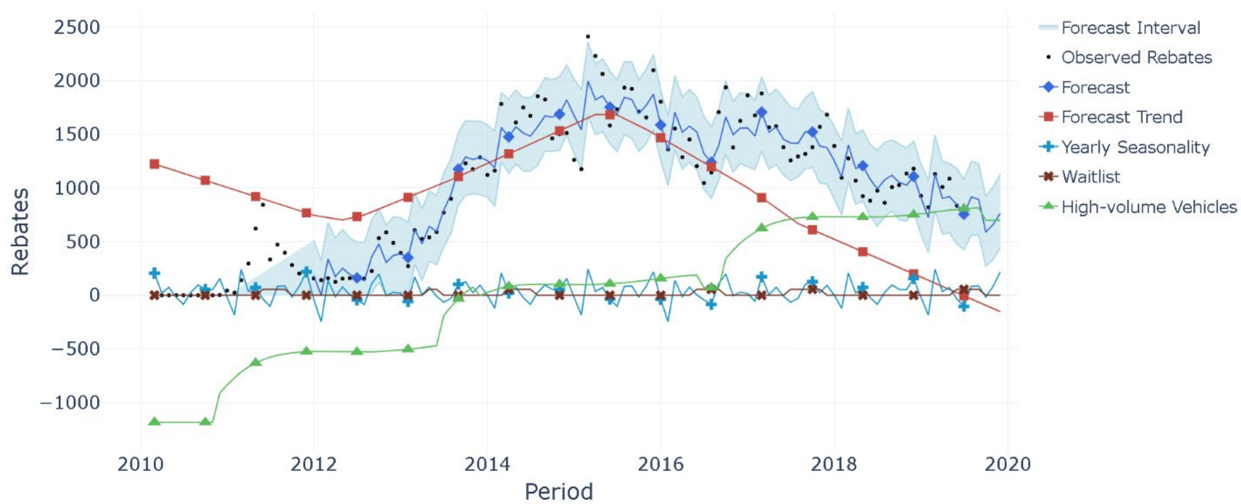

Figure 2. Non-Tesla BEV individual standard rebate model with a high-volume vehicle model regressor.

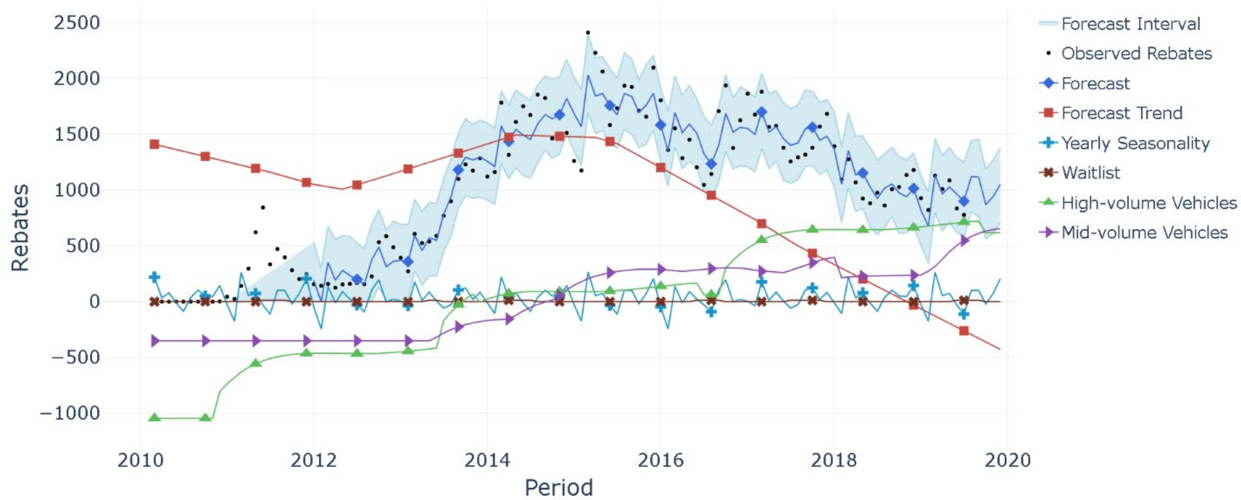

Figure 3. Non-Tesla BEV individual standard rebate model with a mid-volume & high-volume vehicle model regressor.

Non-Tesla BEV Individual Increased Rebates

The following indicators were shown to be statistically significant using MCMC for the non-Tesla BEV individual increased rebate: public charging, mid-volume vehicles, and high-volume vehicles (Table 6). At its peak demand during its sales life cycle, a new mid-volume vehicle is expected to increase rebates by 75 per month, while a high-volume vehicle is expected to increase rebates by 77 per month. Public charging is also expected to positively impact rebates; with every increase in 100 charging ports, there is a predicted increase of one rebate.

Based on the cross-validation metrics listed (Table 7), each of the three significant leading indicator regressors led to greater forecasting accuracy compared to the baseline model. When each leading indicator was individually incorporated into its own model, the mid-volume vehicle indicator resulted in the most accurate cross-validation forecast, followed by the high-volume indicator and then the public charging indicator. In comparison to the base model, the mid-volume regressor reduced the MAPE value by approximately 35 percentage points, while the high-volume regressor similarly reduced the MAPE value by 32 percentage points. Public charging, with a less sizable impact, also improved the accuracy of the forecast by reducing the MAPE value by 5 percentage points.

We then analyzed all combinations of these three significant regressors and found that of these combinations, the only regressor groups that remained significant were the following pairs: public charging with mid-volume vehicles and public charging with high-volume vehicles. Furthermore, the joint public charging indicator with a mid-volume indicator was the only mixed indicator model that had an RMSE metric that performed better than any of the single leading indicator models. In this circumstance, the mixed model performed slightly better than the single mid-volume indicator model, which suggests that the mid-volume indicator is the main driving indicator in improving forecast accuracy. For the MAPE metric, most of the mixed models performed slightly better than the best single indicator model. While using a mixture of the leading indicators may improve the forecast accuracy, there is the added complexity of introducing more variables that will also need to be forecasted for future predictions. As such, since the public charging regressor has minimal impact, we recommend using a more parsimonious model that only includes the mid-volume and high-volume regressors.

Table 6. MCMC sampling results for non-Tesla BEV individuals increased rebates. The addition of a new vehicle and how it will affect rebates during its peak release is provided below the coefficient value in parentheses.

Forecast Group	Regressor Combinations	Regressor Name	Lower Coefficient	Coefficient	Upper Coefficient	Significant Regressor
Single Leading Indicator Models	Household Income	Household Income	-0.1	-0.05	-0.01	TRUE
Single Leading Indicator Models	Public Charging	Public Charging	0	0.01	0.01	TRUE
Single Leading Indicator Models	Unemployment Rate	Unemployment Rate	8.33	20.49	33.32	TRUE
Single Leading Indicator Models	Low-volume Vehicles	Low-volume Vehicles	-12.32 (-16.61)	-4.25 (-5.73)	4.83 (6.51)	FALSE
Single Leading Indicator Models	Mid-volume Vehicles	Mid-volume Vehicles	36.89 (52.77)	52.52 (75.12)	66.55 (95.21)	TRUE
Single Leading Indicator Models	High-volume Vehicles	High-volume Vehicles	3.21 (3.5)	70.1 (76.63)	131.82 (144.12)	TRUE
Mixed Leading Indicator Models	Mid- & High-volume Vehicles	Mid-volume Vehicles	34.93 (49.97)	49.21 (70.4)	63.79 (91.25)	TRUE
Mixed Leading Indicator Models	Mid- & High-volume Vehicles	High-volume Vehicles	-12.02 (-13.14)	41.64 (45.53)	100.95 (110.36)	FALSE
Mixed Leading Indicator Models	Public Charging & Mid-volume Vehicles	Public Charging	0	0.01	0.01	TRUE
Mixed Leading Indicator Models	Public Charging & Mid-volume Vehicles	Mid-volume Vehicles	31.21 (44.65)	47.57 (68.05)	62.95 (90.05)	TRUE
Mixed Leading Indicator Models	Public Charging & High-volume Vehicles	Public Charging	0	0.01	0.01	TRUE
Mixed Leading Indicator Models	Public Charging & High-volume Vehicles	High-volume Vehicles	11.59 (12.67)	82.02 (89.67)	147.21 (160.93)	TRUE
Mixed Leading Indicator Models	Public Charging, Mid- & High-	Public Charging	0	0.01	0.01	TRUE

	volume Vehicles					
Mixed Leading Indicator Models	Public Charging, Mid- & High-volume Vehicles	Mid-volume Vehicles	29.42 (42.09)	44.2 (63.23)	58.83 (84.15)	TRUE
Mixed Leading Indicator Models	Public Charging, Mid- & High-volume Vehicles	High-volume Vehicles	-5.64 (-6.16)	44.68 (48.85)	96.06 (105.02)	FALSE

Table 7. Cross-validation scores for non-Tesla BEV individual increased rebates.

Forecast Group	Forecast Version	RMSE	RMSE Rank	MAPE	MAPE Rank
Base Model	Base	106.846	10	0.747	11
Single Leading Indicator Models	Household Income	107.524	11	0.728	10
Single Leading Indicator Models	Public Charging	101.371	8	0.697	8
Single Leading Indicator Models	Unemployment Rate	102.25	9	0.719	9
Single Leading Indicator Models	Low-volume Vehicles	92.322	7	0.688	7
Single Leading Indicator Models	Mid-volume Vehicles	55.79	2	0.397	4
Single Leading Indicator Models	High-volume Vehicles	69.206	6	0.432	6
Mixed Leading Indicator Models	Mid- & High-volume Vehicles	61.037	4	0.419	5
Mixed Leading Indicator Models	Public Charging & Mid-volume Vehicles	55.15	1	0.395	3
Mixed Leading Indicator Models	Public Charging & High-volume Vehicles	66.575	5	0.393	2
Mixed Leading Indicator Models	Public Charging, Mid- & High-volume Vehicles	59.32	3	0.392	1

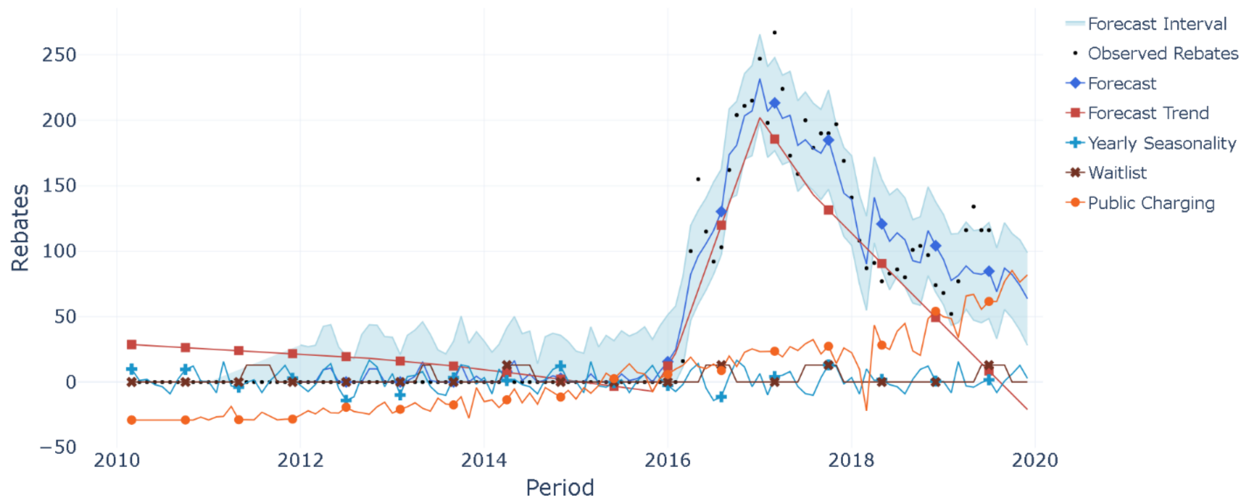

Figure 4. Non-Tesla BEV individual increased rebate model with a public charging vehicle model regressor.

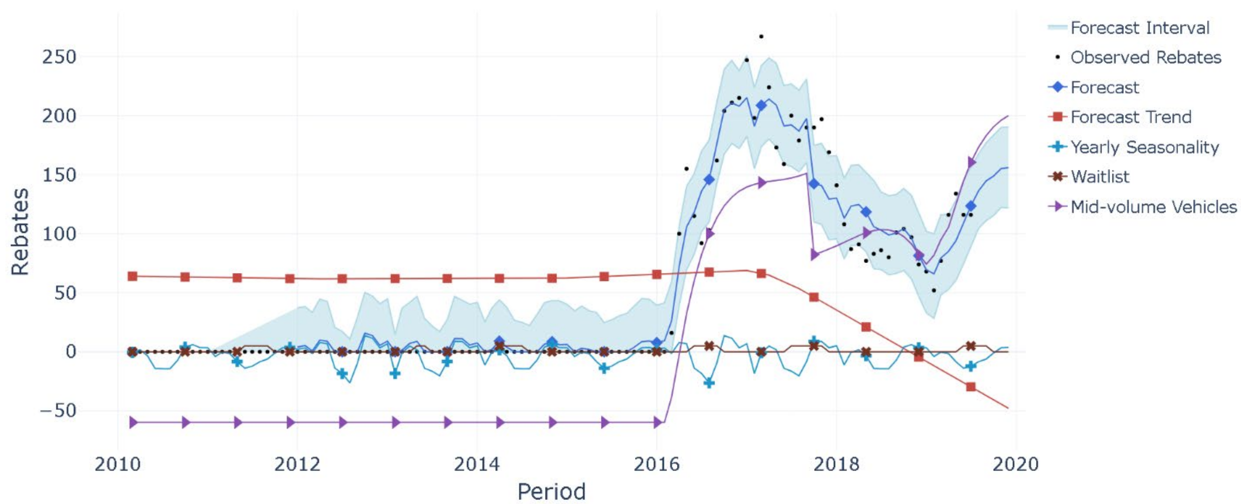

Figure 5. Non-Tesla BEV individual increased rebate model with a mid-volume vehicle model regressor.

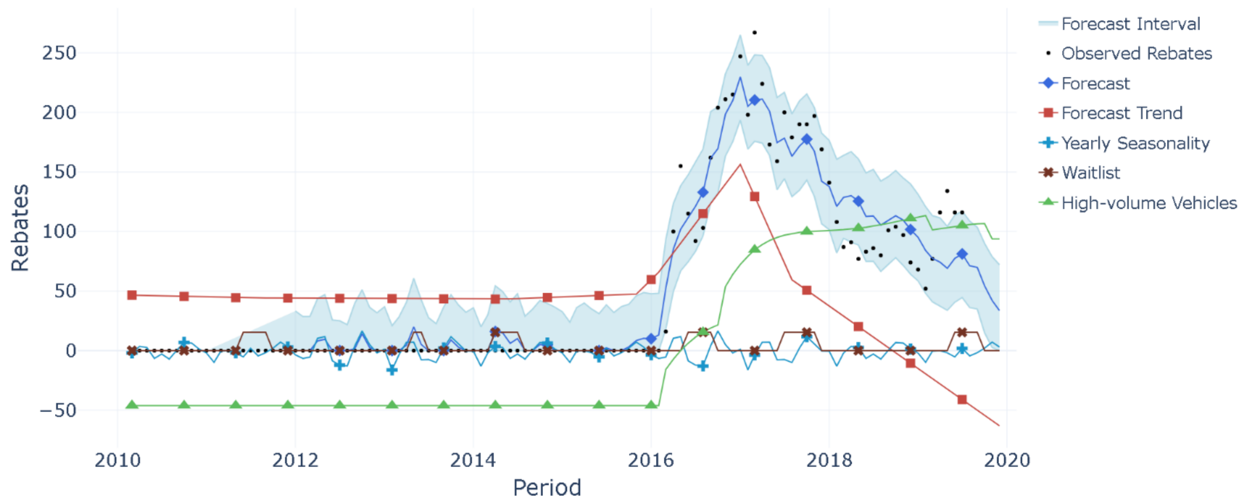

Figure 6. Non-Tesla BEV individual increased rebate model with a high-volume vehicle model regressor.

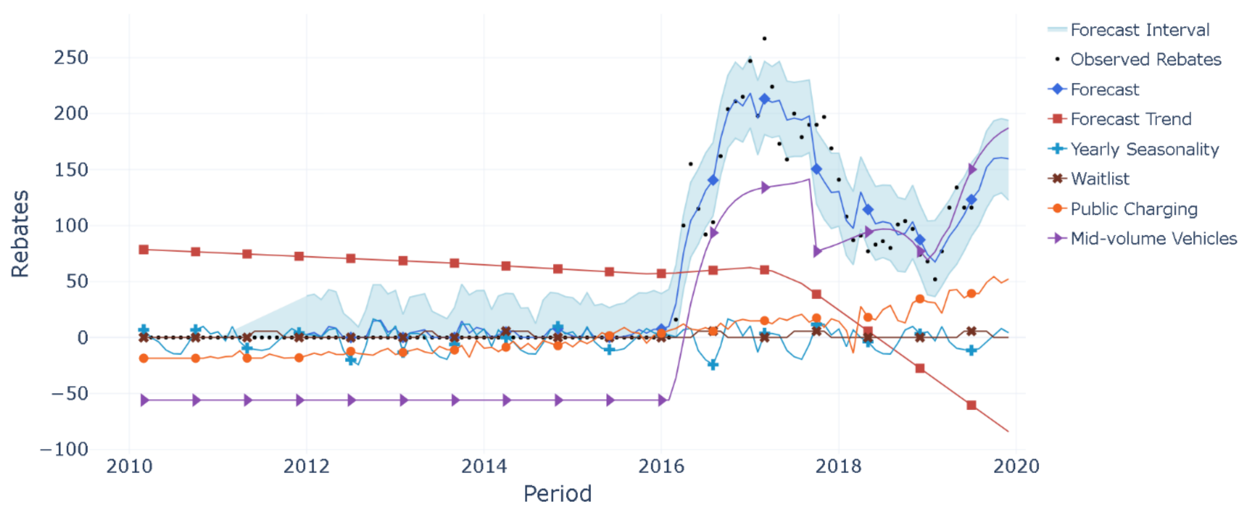

Figure 7. Non-Tesla BEV individual increased rebate model with a mid-volume vehicle and public charging model regressor.

Tesla BEV Individual Standard Rebates

In recent years Tesla-specific models have grown to become the majority of total CVRP rebates, which led to our decision to analyze Tesla-specific BEV rebates separately. Additionally, Tesla's distinct California delivery seasonality exemplifies their difference in comparison to non-Tesla BEV rebate trends, further supporting Tesla BEVs into their own respective model. This model (and other Tesla-specific models below) did not include low-, mid-, and high-volume indicators since there is an insufficient number of Tesla models to generate meaningful categories. From our analysis, household income was a significant leading indicator based on the MCMC results in Table 8. The estimated impact of household income is that for every 100 additional households with an income over $200,000, there will be an increase of 46 rebates per month. The results in Table 9 show that the household income indicator is also the best-performing single leading indicator in terms of the RMSE and MAPE metrics. Compared to the baseline model, the model with household income has a MAPE value roughly 5 percentage points lower than the base model. This decrease translates to an improved forecast of about 5 percent within the cross-validation period.

The significance of this household income variable makes sense since the economic environment has created an increasing trend in households with incomes higher than $200,000 that largely mirrors the increasing trend in BEVs over the past decade. However, it is likely that this general association is more coincidental rather than causal and would likely result in misleading forecasts in the future if it were included in the forecast model (see Figure 8). Another concern with this regressor is that any changes that lower the household income cap for applicants will decrease its impact. For instance, the CVRP income eligibility requirements were lowered twice in 2016,[8] which reduced the number of households eligible for the program. We suggest excluding this variable from future models despite significantly improving the forecast here.

Table 8. MCMC sampling results for Tesla BEV individual standard rebates.

Forecast Group	Regressor Combinations	Regressor Name	Lower Coefficient	Coefficient	Upper Coefficient	Significant Regressor
Single Leading Indicator Models	Household Income	Household Income	0.01	0.46	0.89	TRUE
Single Leading Indicator Models	Public Charging	Public Charging	-0.07	0.01	0.09	FALSE
Single Leading Indicator Models	Unemployment Rate	Unemployment Rate	-226.3	-12.16	203.22	FALSE

Table 9. Cross-validation scores for Tesla BEV individual standard rebates.

Forecast Group	Forecast Version	RMSE	RMSE Rank	MAPE	MAPE Rank
Base Model	Base	1239.368	2	0.489	2
Single Leading Indicator Models	Household Income	1164.348	1	0.44	1
Single Leading Indicator Models	Public Charging	1244.162	3	0.491	3
Single Leading Indicator Models	Unemployment Rate	1257.398	4	0.503	4

[8] https://cleanvehiclerebate.org/sites/default/files/attachments/Disruptions_Fact_Sheet_9_2021.pdf

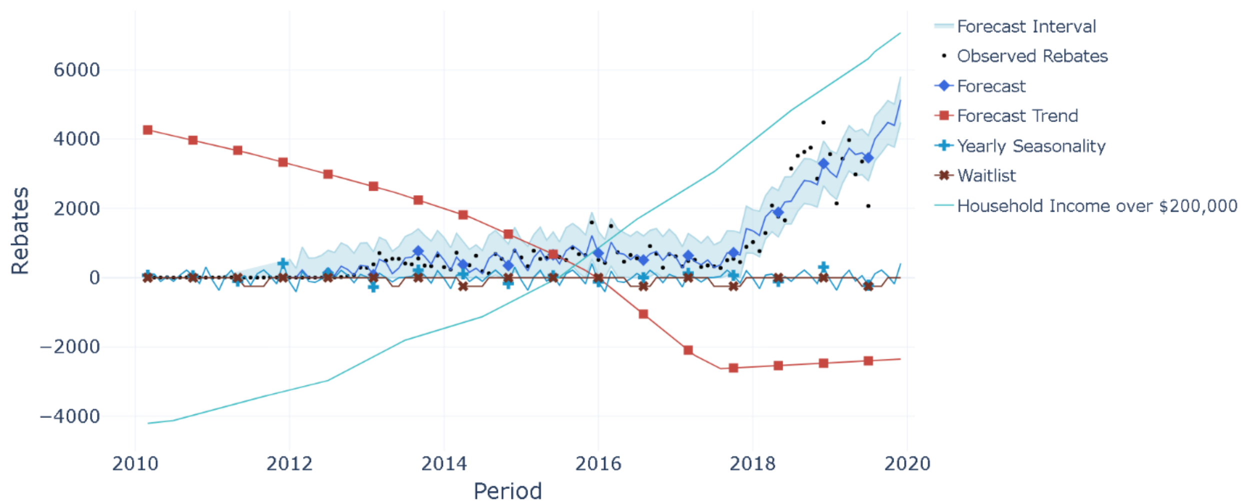

Figure 8. Tesla BEV individual standard rebate model with a household income over $200,000 model regressor.

Tesla BEV Individual Increased Rebates

For forecast models for Tesla individual increased rebates, public charging was the only variable that significantly improved the forecast. In contrast, higher income is understandably not significant in predicting increased rebates as increased rebates are only available to those in lower-income households. The estimated impact of public charging is an increase in five rebates for every 1,000 new charging ports. Additionally, the public charging indicator also has a lower RMSE and MAPE value in comparison to the baseline model (see Table 11). Still, this improvement is minimal, with forecast accuracy improving by less than 1 percentage point.

Although the metrics suggest incorporating public charging, there are challenges to distinguishing the endogenous relationship between the number of charging ports and the number of rebated vehicles. It is possible there is not a direct causal relationship between public chargers to the number of Tesla increased rebates, but rather a reciprocated relation as public chargers are also being introduced to meet increases in vehicle demand.

Table 10. MCMC sampling results for Tesla BEV individual increased rebates.

Forecast Group	Regressor Combinations	Regressor Name	Lower Coefficient	Coefficient	Upper Coefficient	Significant Regressor
Single Leading Indicator Models	Household Income	Household Income	-0.08	-0.04	0	FALSE
Single Leading Indicator Models	Public Charging	Public Charging	0	0	0.01	TRUE
Single Leading Indicator Models	Unemployment Rate	Unemployment Rate	-3.42	8.1	19.08	FALSE

Table 11. Cross-validation scores for Tesla BEV individual increased rebates.

Forecast Group	Forecast Version	RMSE	RMSE Rank	MAPE	MAPE Rank
Base Model	Base	60.166	4	0.594	4
Single Leading Indicator Models	Household Income	59.319	3	0.592	3
Single Leading Indicator Models	Public Charging	56.783	2	0.586	2
Single Leading Indicator Models	Unemployment Rate	56.221	1	0.567	1

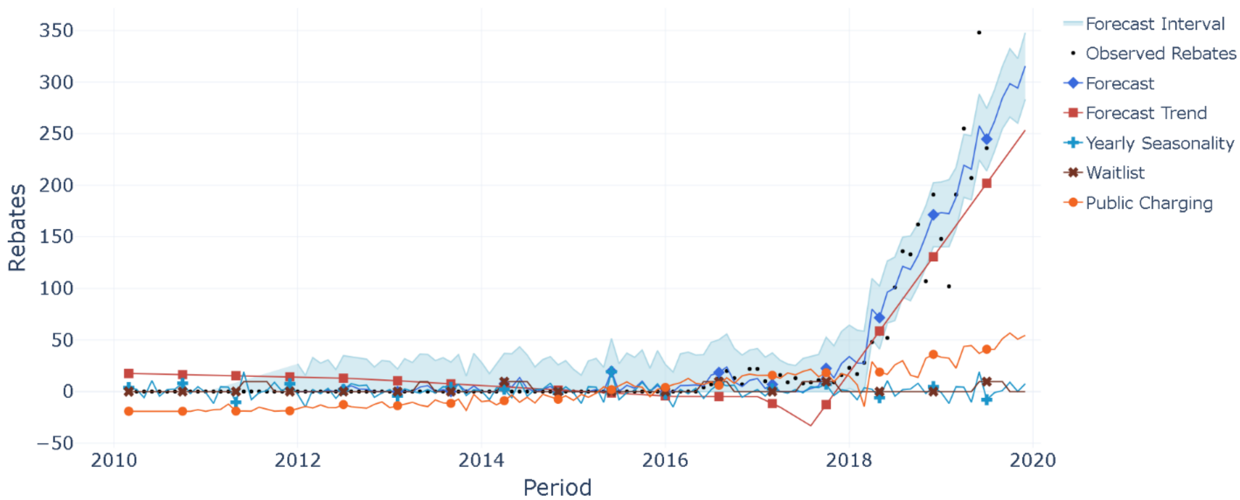

Figure 9. Tesla BEV individual increased rebate model with a public charging model regressor.

PHEV Individual Standard Rebates

Analysis of PHEV individual standard rebates showed that low-, mid-, and high-volume eligibility regressors are all statistically significant; the modeling suggests that at their peak impact, a single unit increase in eligible vehicles is expected to increase rebates by 108, 860, and 442 per month, respectively (see Table 12). The reason that mid-level volume vehicles are expected to contribute more than high-volume vehicles is visible in Figure 11 and Figure 12. When comparing the mid-volume and high-volume vehicle release trend to the overall CVRP rebate pattern, the increases in the number of mid-volume vehicles coincide more with rebates than the amount of high-volume eligible vehicles, causing the impact of this regressor to be inflated. However, the wide confidence intervals (Table 12) show the large uncertainty of this impact and should approach more realistic values with time.

Based on the cross-validation metrics in Table 13, the mid-volume vehicle regressor is the only regressor that improved the forecast within the testing time frame. The low- and high-volume regressors did not increase the forecast accuracy compared to the baseline since the cross-validation period overlaps with an enacted waitlist period near the end of December 2017. The low- and high-volume regressors predict an increase in rebates in addition to the predicted waitlist decline, which reduces the forecast accuracy of these models in this testing time frame. Conversely, the amount of eligible mid-volume vehicles plateaued during this 2017 waitlist period and then immediately increased afterward, just as the number of rebates did, which improved the forecasting accuracy for this model compared to the base model.

We also tested the model by jointly incorporating all three of the eligible vehicle regressors in various combinations, and in this setting, each of the individual vehicle regressors remained significant based on the MCMC results. These joint models did not improve the forecast accuracy for similar reasons as previously mentioned with the waitlist period. Still, it is worthwhile to incorporate eligible vehicles in a parsimonious model with only a high-volume vehicle regressor as that vehicle group is expected to have the largest impact on the number of rebated vehicles. As more data becomes available, the possibility of incorporating a joint mid- and high-volume vehicle model should continue to be considered.

Table 12. MCMC sampling results for PHEV individual standard rebates. The addition of a new vehicle and how it will affect rebates during its peak release is provided below the coefficient value in parentheses.

Forecast Group	Regressor Combinations	Regressor Name	Lower Coefficient	Coefficient	Upper Coefficient	Significant Regressor
Household Income	Household Income	Household Income	-0.17	0.18	0.55	FALSE
Public Charging	Public Charging	Public Charging	-0.09	-0.04	0	TRUE
Unemploymen t Rate	Unemployment Rate	Unemployment Rate	-178.55	-75.19	26.4	FALSE
Low-volume Vehicles	Low-volume Vehicles	Low-volume Vehicles	37.76 (50.93)	79.88 (107.74)	122.34 (165.01)	TRUE
Mid-volume Vehicles	Mid-volume Vehicles	Mid-volume Vehicles	286.91 (410.43)	601.17 (859.98)	908.57 (1299.72)	TRUE
High-volume Vehicles	High-volume Vehicles	High-volume Vehicles	13.31 (14.55)	404.32 (442.03)	797.75 (872.15)	TRUE
Mixed Leading Indicator Models	Low- & Mid-volume Vehicles	Low-volume Vehicles	14.52 (19.58)	55.08 (74.29)	96.88 (130.66)	TRUE
Mixed Leading Indicator Models	Low- & Mid-volume Vehicles	Mid-volume Vehicles	107.06 (153.15)	435.28 (622.67)	760.22 (1087.51)	TRUE
Mixed Leading Indicator Models	Low- & High-volume Vehicles	Low-volume Vehicles	61.57 (83.05)	103.0 (138.92)	140.71 (189.78)	TRUE
Mixed Leading Indicator Models	Low- & High-volume Vehicles	High-volume Vehicles	351.0 (383.73)	763.83 (835.07)	1179.18 (1289.16)	TRUE

Mixed Leading Indicator Models	Mid- & High-volume Vehicles	Mid-volume Vehicles	389.95 (557.82)	724.8 (1036.84)	1010.12 (1444.99)	TRUE
Mixed Leading Indicator Models	Mid- & High-volume Vehicles	High-volume Vehicles	257.1 (281.08)	670.96 (733.53)	1050.81 (1148.82)	TRUE
Mixed Leading Indicator Models	Low-, Mid- & High-volume Vehicles	Low-volume Vehicles	31.5 (42.48)	70.56 (95.17)	113.89 (153.61)	TRUE
Mixed Leading Indicator Models	Low-, Mid- & High-volume Vehicles	Mid-volume Vehicles	228.83 (327.34)	493.3 (705.67)	786.84 (1125.59)	TRUE
Mixed Leading Indicator Models	Low-, Mid- & High-volume Vehicles	High-volume Vehicles	407.69 (445.71)	822.92 (899.67)	1235.65 (1350.89)	TRUE

Table 13. Cross-validation scores for PHEV individual standard rebates.

Forecast Group	Forecast Version	RMSE	RMSE Rank	MAPE	MAPE Rank
Base Model	Base	300.06	2	0.171	2
Single Leading Indicator Models	Household Income	389.357	6	0.236	8
Single Leading Indicator Models	Public Charging	309.18	3	0.175	3
Single Leading Indicator Models	Unemployment Rate	315.397	4	0.178	4
Single Leading Indicator Models	Low-volume Vehicles	404.958	8	0.205	6
Single Leading Indicator Models	Mid-volume Vehicles	280.006	1	0.156	1
Single Leading Indicator Models	High-volume Vehicles	437.94	9	0.267	10
Mixed Leading Indicator Models	Low- & Mid-volume Vehicles	400.479	7	0.2	5
Mixed Leading Indicator Models	Low- & High-volume Vehicles	526.875	11	0.288	11
Mixed Leading Indicator Models	Mid- & High-volume Vehicles	371.787	5	0.226	7
Mixed Leading Indicator Models	Low-, Mid- & High-volume Vehicles	486.367	10	0.263	9

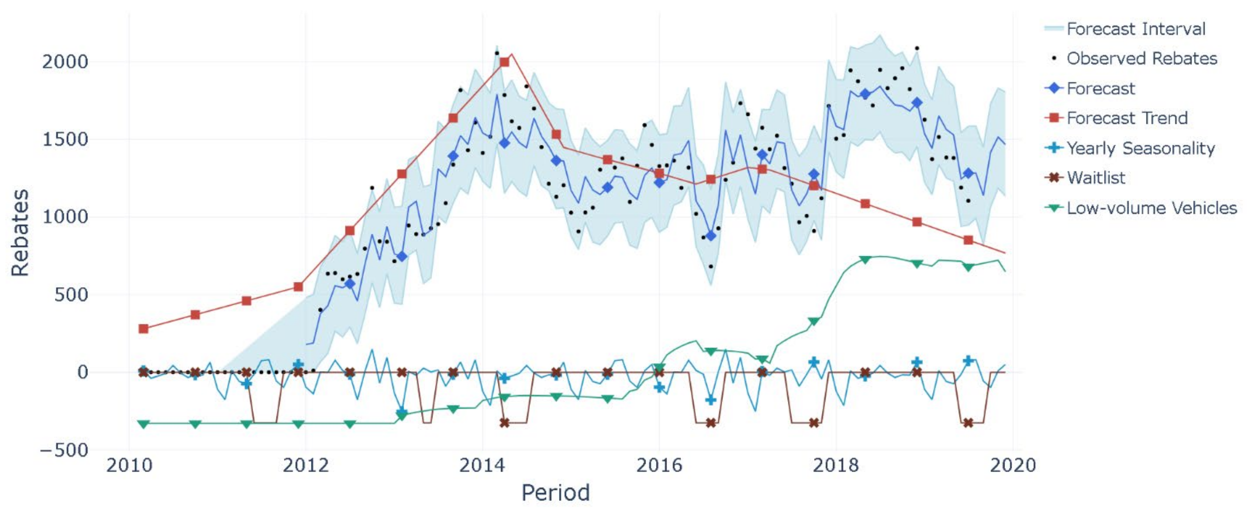

Figure 10. PHEV individual standard rebate model with a low-volume model regressor.

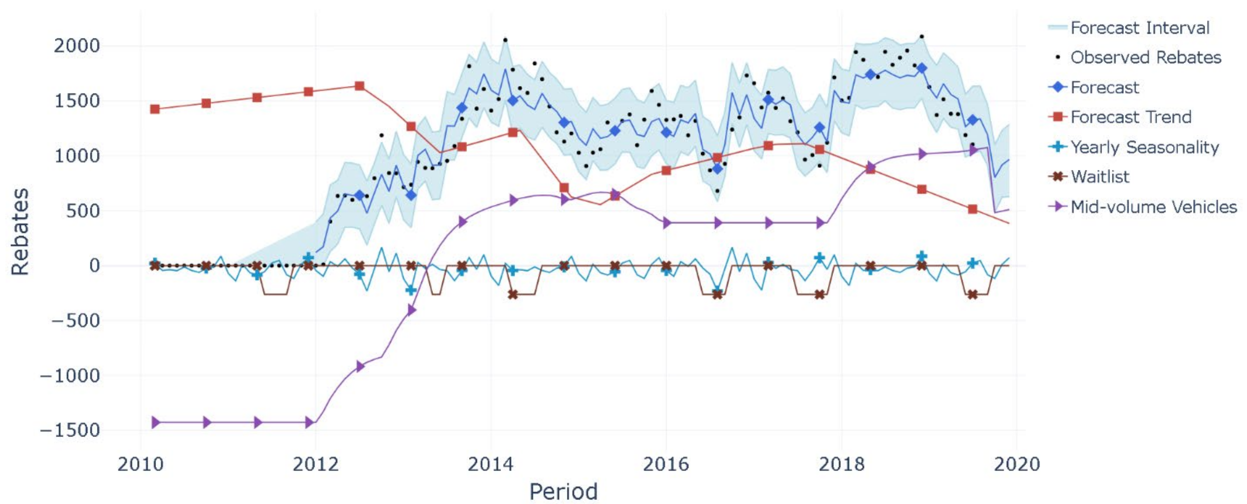

Figure 11. PHEV individual standard rebate model with a medium-volume model regressor.

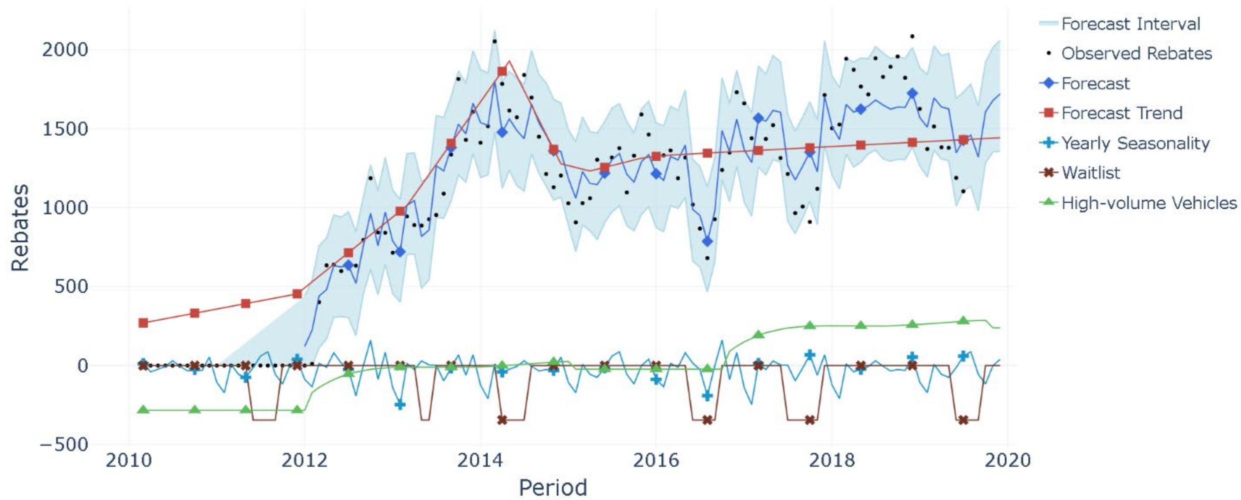

Figure 12. PHEV individual standard rebate model with a high-volume model regressor.

PHEV Individual Increased Rebates

The low-volume and mid-volume eligible vehicle regressors are statistically significant in the PHEV individual increased rebates model. The impact of adding a low-volume vehicle during its release peak is expected to increase rebates by 19 per month. In contrast, the addition of an eligible mid-volume vehicle is expected to increase rebates by 118 per month (see Table 14). The low-volume and mid-volume models also have the lowest RMSE and MAPE values for all the other single leading indicator models (see Table 15). Comparing the mid-volume and low-volume regressor models to the base model, both indicators reduce the MAPE score by nearly 10 percentage points.

In the scenario of jointly adding both low-volume and mid-volume vehicles into the model, both regressors are still significant, but their single leading indicator counterparts perform better based on MAPE and RMSE values. We also evaluated a joint mid- and high-volume indicator, which remained statistically significant and performed second best out of all the possible single and mixed leading indicator models. While the joint mid- and high-volume indicator improves the cross-validation metrics, the overall trendline of the model has strongly decreased, which may lead to a loss of forecast accuracy in future predictions. For these reasons, we recommend a parsimonious model only utilizing the mid-volume vehicle regressor but also consider reevaluating a high-volume vehicle regressor as more data becomes available.

Table 14. MCMC sampling results for PHEV individual increased rebates. The addition of a new vehicle and how it will affect rebates during its peak release is provided below the coefficient value in parentheses.

Forecast Group	Regressor Combinations	Regressor Name	Lower Coefficient	Coefficient	Upper Coefficient	Significant Regressor
Single Leading Indicator Models	Household Income	Household Income	0.02	0.06	0.1	TRUE

Single Leading Indicator Models	Public Charging	Public Charging	-0.01	0	0	FALSE
Single Leading Indicator Models	Unemployment Rate	Unemployment Rate	-18.3	-7.44	2.75	FALSE
Single Leading Indicator Models	Low-volume Vehicles	Low-volume Vehicles	9.93 (13.4)	14.37 (19.39)	18.54 (25.0)	TRUE
Single Leading Indicator Models	Mid-volume Vehicles	Mid-volume Vehicles	60.33 (86.31)	82.9 (118.6)	103.82 (148.52)	TRUE
Single Leading Indicator Models	High-volume Vehicles	High-volume Vehicles	-49.21 (-53.8)	16.97 (18.55)	80.63 (88.15)	FALSE
Mixed Leading Indicator Models	Low- & Mid-volume Vehicles	Low-volume Vehicles	-8.61 (-11.61)	1.18 (1.59)	11.31 (15.26)	FALSE
Mixed Leading Indicator Models	Low- & Mid-volume Vehicles	Mid-volume Vehicles	25.25 (36.12)	76.75 (109.79)	127.0 (181.68)	TRUE
Mixed Leading Indicator Models	Mid- & High-volume Vehicles	Mid-volume Vehicles	75.21 (107.59)	95.86 (137.13)	118.61 (169.68)	TRUE
Mixed Leading Indicator Models	Mid- & High-volume Vehicles	High-volume Vehicles	43.4 (47.45)	104.74 (114.51)	173.64 (189.84)	TRUE
Mixed Leading Indicator Models	Low-, Mid- & High-volume Vehicles	Low-volume Vehicles	-13.91 (-18.76)	-3.52 (-4.74)	5.99 (8.08)	FALSE
Mixed Leading Indicator Models	Low-, Mid- & High-volume Vehicles	Mid-volume Vehicles	59.53 (85.16)	111.69 (159.78)	163.05 (233.24)	TRUE
Mixed Leading Indicator Models	Low-, Mid- & High-volume Vehicles	High-volume Vehicles	50.0 (54.67)	108.95 (119.11)	178.22 (194.84)	TRUE

Table 15. Cross-validation scores for PHEV individual increased rebates.

Forecast Group	Forecast Version	RMSE	RMSE Rank	MAPE	MAPE Rank
Base Model	Base	48.443	7	0.242	7
Single Leading Indicator Models	Household Income	48.295	6	0.242	9
Single Leading Indicator Models	Public Charging	49.145	8	0.242	8

Single Leading Indicator Models	Unemployment Rate	49.295	9	0.238	6
Single Leading Indicator Models	Low-volume Vehicles	28.261	1	0.129	1
Single Leading Indicator Models	Mid-volume Vehicles	33.368	3	0.147	3
Single Leading Indicator Models	High-volume Vehicles	63.322	10	0.317	10
Mixed Leading Indicator Models	Low- & Mid-volume Vehicles	35.123	4	0.166	4
Mixed Leading Indicator Models	Low-, Mid- & High-volume Vehicles	43.235	5	0.208	5

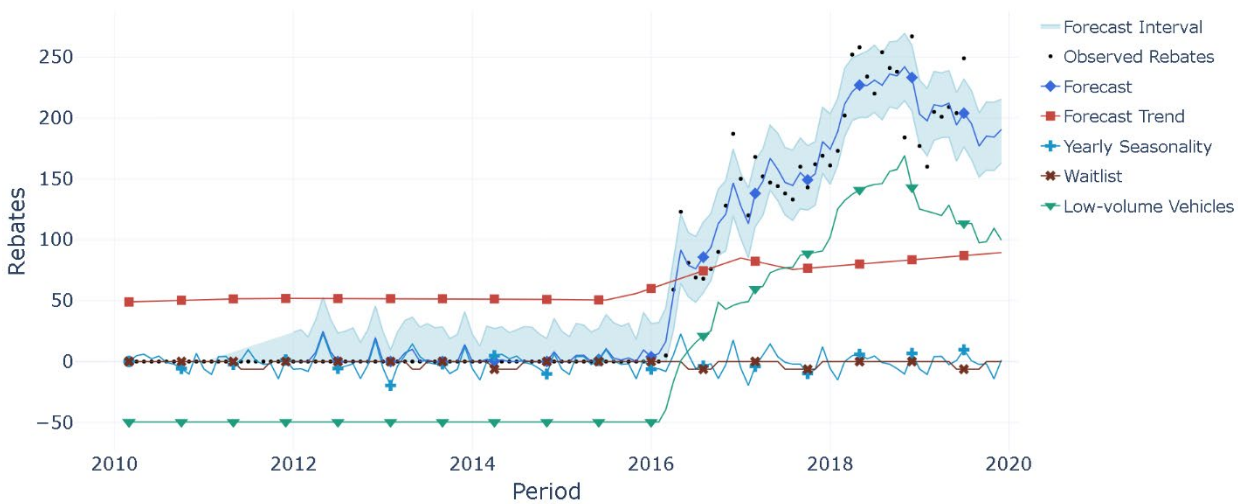

Figure 13. PHEV individual increased rebate model with a low-volume model regressor.

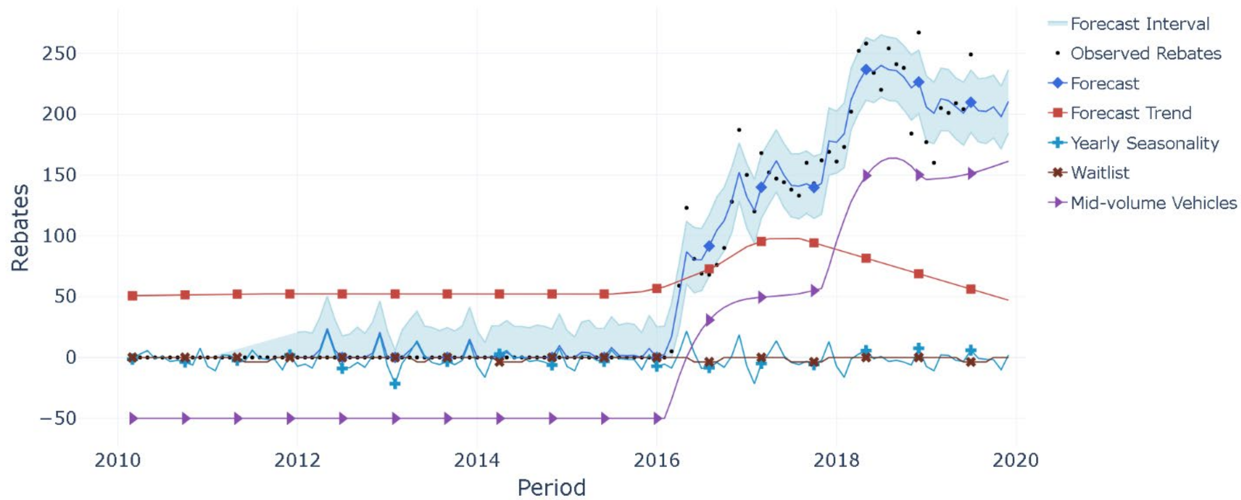

Figure 14. PHEV individual increased rebate model with a mid-volume model regressor.

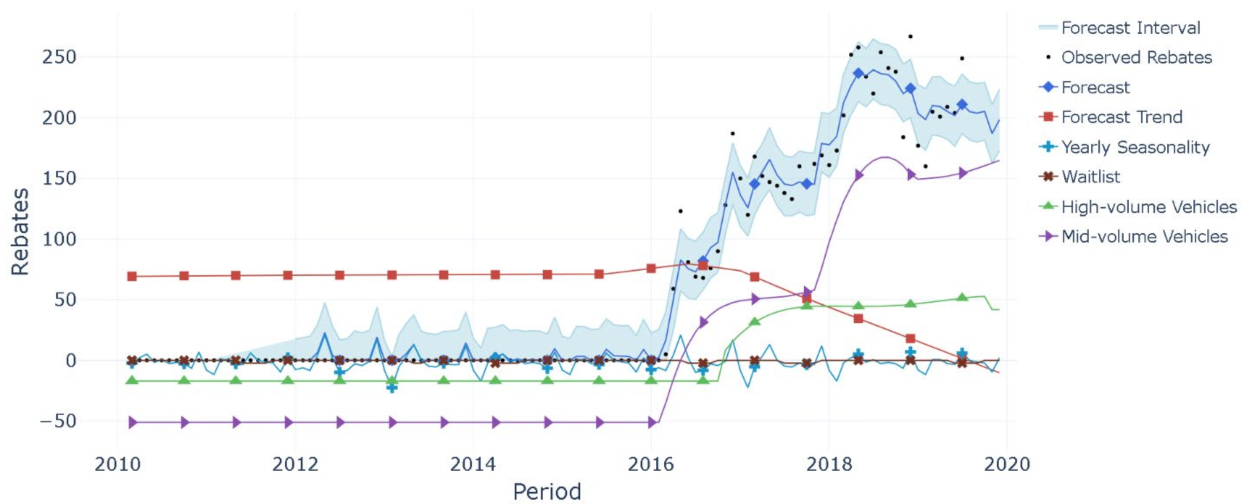

Figure 15. PHEV individual increased rebate model with a mid-volume and high-volume model regressor.

Discussion

After analyzing multiple leading indicators in various model settings, noteworthy results were found that will be utilized for future research and forecasting methodology. Conversely, a few indicators had challenges and will not be utilized in future work: income, public charging, and household income. For example, while the unemployment rate has a high correlation with program participation, the indicator was not considered significant and did not improve forecast accuracy. Alternatively, household income and public charging were significant in the Tesla-specific models, but we recommend their exclusion in future models. While the households with incomes over $200,000 regressor were significant, CVRP continues to implement future income caps, which will eventually nullify the direct relationship that high-income household applicants have on the number of rebates. Furthermore, public charging has an endogenous relationship with rebated vehicles, and using this variable would require knowledge of future charging station locations, which is not readily available.

Comparatively, incorporating CVRP eligible vehicles as a regressor is a much more feasible indicator. The benefit of modeling future release vehicles is they are consistently analyzed from outside sources, and predictions about their release date are available with sufficient notice. The predicted vehicle release date also becomes more accurate as the confirmed release date nears, especially with new vehicles now administering pre-orders. The pre-order component of new vehicles also gives a better intuition of which volume group the vehicle should fall into.

In addition to their simplicity in getting future release information, the program-eligible vehicle regressors were significant in multiple models and improved forecast accuracy. The low-, medium-, and high-volume regressors all had instances where they were not significant in certain models, but the medium- and high-volume model indicators were more often significant in the models they were analyzed in. For non-Tesla BEV models, the joint use of a medium- and high-volume vehicle regressor is the best choice, while for PHEV models the use of a single medium- or high-volume regressor led to better forecast accuracy. Although, when deciding between the medium- or high-volume regressor, the better choice is the high-volume regressor as that vehicle classification regressor is anticipated to have the largest impact on the number of rebates.

Conclusion

The executive branch and state agencies in California continue to push for increasing adoption of EVs to reach the state's climate objectives.[9] After an extensive review of the many intrinsic and extrinsic variables that may influence adoption, CSE has found that the number of vehicles eligible for the CVRP is the most important variable to include in future models for predicting future CVRP rebates. The implementation of EV eligibility within CVRP and its use in modeling efforts should continue in future years as more data becomes available, and the model is reevaluated to optimize performance further. Implementing these factors will improve the accuracy of future forecasting exercises for both EV demand and the rate at which CVRP rebate funds will be exhausted. The growing number of EVs eligible for the CVRP rebate, which appeals to a wide variety of consumers, will contribute to the expected rapid expansion of the market. Tracking the announcements of future production EVs and quantifying their likely appeal is an area of further study that will enhance forecast accuracy.

[9] https://www.cpuc.ca.gov/zev/

Data References

A. Includes content from IHS Markit© 2021
B. https://afdc.energy.gov
C. https://www.census.gov
D. https://www.edd.ca.gov

Appendix

Vehicle Volume Clustering

Table 16. Low-volume categorization clustering results for individual standard rebate CVRP eligible non-Tesla BEVs. Tesla BEVs are excluded from clustering due to the limited number of models to make meaningful clusters.

Low-volume non-Tesla BEVs	Average Monthly Rebates
Wheego LiFe	0
Th!nk City	1
smart Electric Fortwo Cabriolet	2
Mitsubishi i-MiEV	3
CODA	3
BMW i3s	6
Polestar 2	6
BMW i3s REx	7
Jaguar I-PACE	9
Honda Fit EV	10
MINI Cooper SE	13
Volvo XC40 Recharge	14
Audi e-tron	15
Hyundai Ioniq Electric	24
Kia Soul EV	27
smart Electric Fortwo Coupe	30
Chevrolet Bolt EUV	34
BMW i3	35
Mercedes-Benz B250e	37
Toyota RAV4 EV	38
Ford Focus Electric	42
Honda Clarity Electric	49
Kia Niro Electric	57

Table 17. Mid-volume categorization clustering results for individual standard rebate CVRP eligible non-Tesla BEVs. Tesla BEVs are excluded from clustering due to the limited number of models to make meaningful clusters.

Mid-volume non-Tesla BEVs	Average Monthly Rebates
Hyundai Kona Electric	70
BMW i3 REx	80
Chevrolet Spark EV	88

Ford Mustang Mach-E	120
Volkswagen ID.4	122
Volkswagen e-Golf	127

Table 18. High-volume categorization clustering results for individual standard rebate CVRP eligible non-Tesla BEVs. Tesla BEVs are excluded from clustering due to the limited number of models to make meaningful clusters.

High-volume non-Tesla BEVs	Average Monthly Rebates
FIAT 500e	218
Nissan Leaf	311
Chevrolet Bolt EV	372

Table 19. Low-volume categorization clustering results for individual standard rebate CVRP eligible PHEVs.

Low-volume PHEVs	Average Monthly Rebates
Volvo S90 T8	1
Ford Escape Plug-In Hybrid	1
Volvo S60 T8	2
BMW 530e xDrive iPerformance	2
Mercedes-Benz S-Class 550e	3
Volvo XC90 T8	5
Volvo XC60 T8	7
Cadillac ELR	7
Honda Accord Plug-In	9
Kia Optima Plug-in Hybrid	11
Hyundai Sonata Plug-in Hybrid	16
Mitsubishi Outlander PHEV	18
Subaru Crosstrek Hybrid PHEV	26
Hyundai Ioniq PHEV	43
Audi A3 e-tron	52
Kia Niro Plug-in Hybrid	55
Chrysler Pacifica	57
BMW 530e iPerformance	75

Table 20. Mid-volume categorization clustering results for individual standard rebate CVRP eligible PHEVs.

Mid-volume PHEVs	Average Monthly Rebates
Ford C-MAX Energi	104
Toyota RAV4 Prime	113
Ford Fusion Energi	149
Honda Clarity Plug-In Hybrid	224

Table 21. High-volume categorization clustering results for individual standard rebate CVRP eligible PHEVs.

High-volume PHEVs	Average Monthly Rebates
Toyota Prius Plug-in Hybrid	269
Toyota Prius Prime	343
Chevrolet Volt	408

To determine how to model the release of a new vehicle, California registration data was used. Data on each ZEV registration was aligned by the first recorded month it had a registration. Since the number of registrations varied per month, the results were then scaled with respect to their average monthly registrations to make each vehicle comparable. After the figures were aligned and scaled, the best fit curve to predict the vehicle release pattern was created. See Figure 16 for an example of this process. In the figure, the gray lines are recorded registrations for vehicles after they have been shifted by their first release month and after they have been scaled. The overlaid dashed blue line is the estimated release pattern of a new vehicle. Based on our analysis of the data, when the vehicle is initially released, we expect a low number of registrations, but as the vehicle has been on the market for more months, we expect the number of registrations to increase for 30 months. After 30 months, we expect the vehicle registrations to begin to taper off.

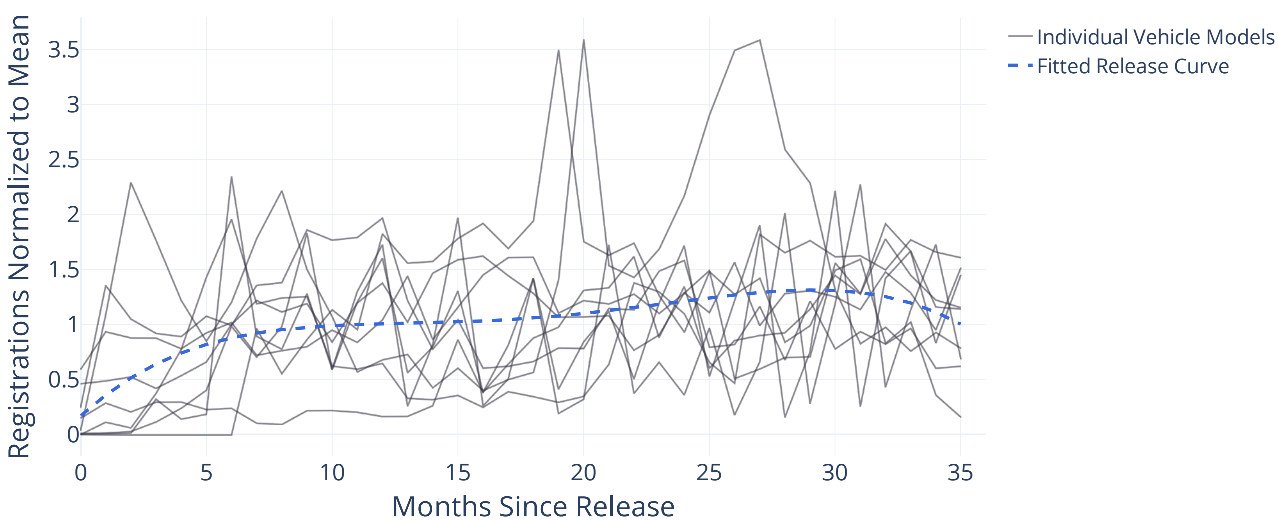

Figure 16. Fitted curve method of predicting new eligible vehicles.

When using this vehicle release estimator as a regressor in Prophet, the coefficients that are provided must be adjusted to return an interpretable result of how a new vehicle will affect the number of rebates. The figure above depicts that the greatest frequency of registrations occurs 30 months after the vehicle has been released, which corresponds to 1.31 on the y-axis.

The coefficients that are returned in MCMC samples need to be scaled by the factors provided in Table 22. For example, if a mid-volume BEV regressor has a coefficient of 100, then during the max release peak of that vehicle we would expect 131 (100*1.31=131) additional monthly rebates.

Table 22. Modeling the release pattern of vehicles and where they reach their max.

Rebate Volume Group	BEV & PHEV Volume Models Max Values	Non-Tesla BEV Volume Model Max Values
Low-volume Vehicle	1.40	1.41
Mid-volume Vehicle	1.31	1.55
High-volume Vehicle	1.68	1.15

Baseline Forecast Models

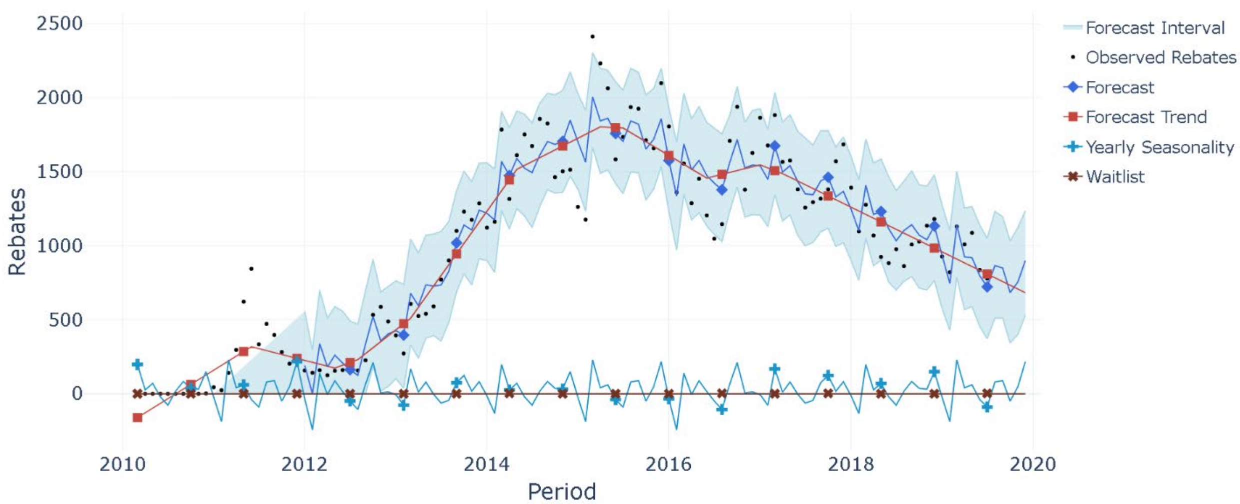

Figure 17. Baseline non-Tesla BEV individual standard rebate forecast.

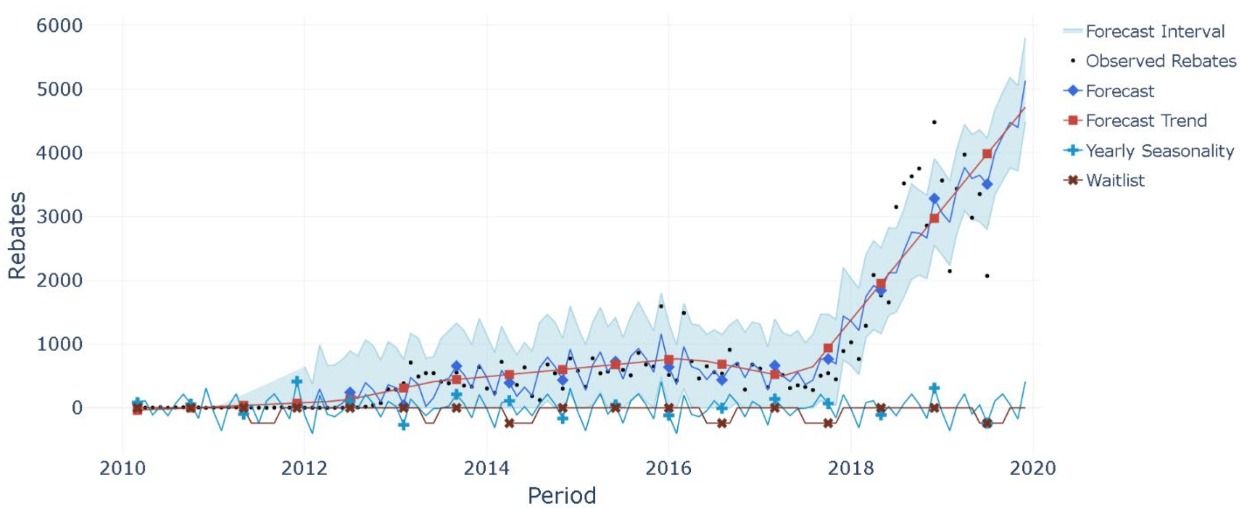

Figure 18. Baseline Tesla BEV individual standard rebate forecast.

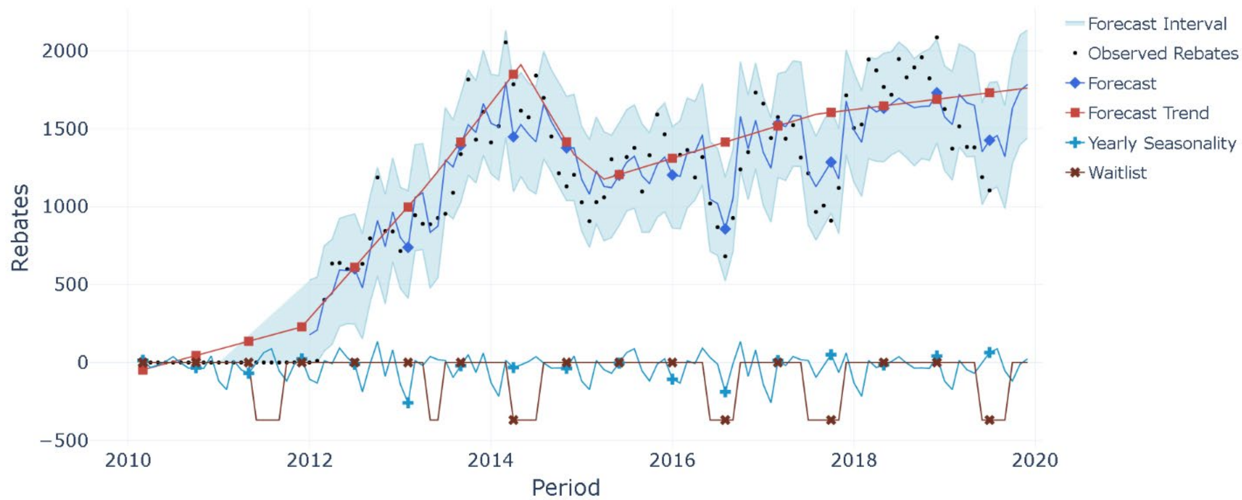

Figure 19. Baseline PHEV individual standard rebate forecast.

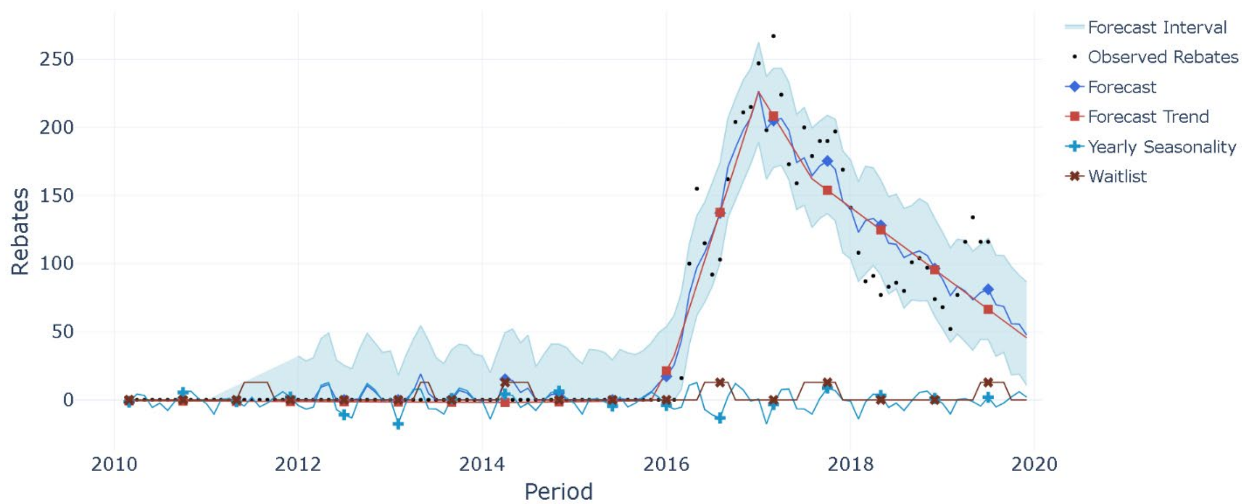

Figure 20. Baseline non-Tesla BEV individual increased rebate forecast.

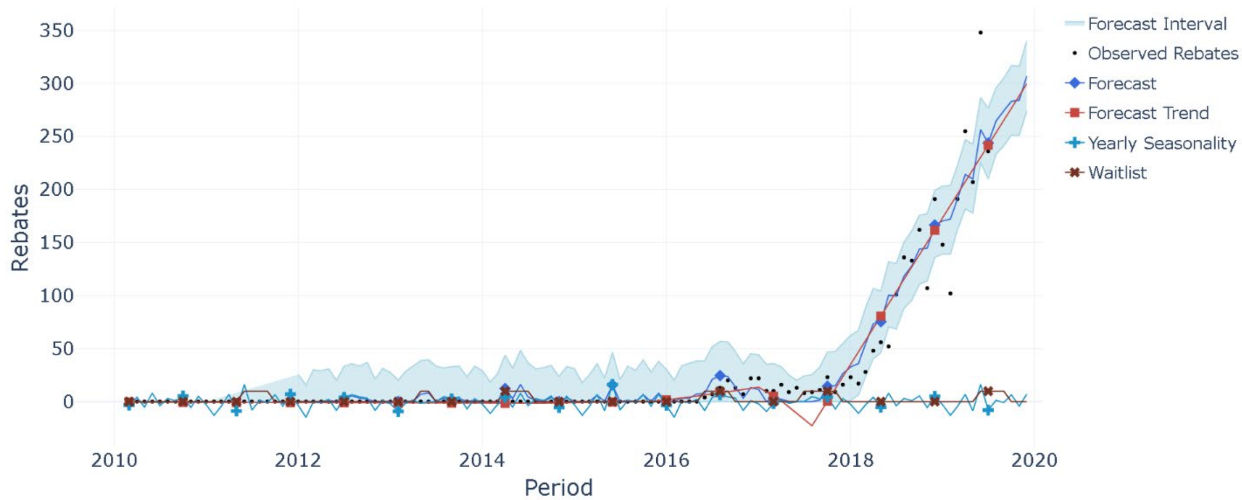

Figure 21. Baseline Tesla BEV individual increased rebate forecast.

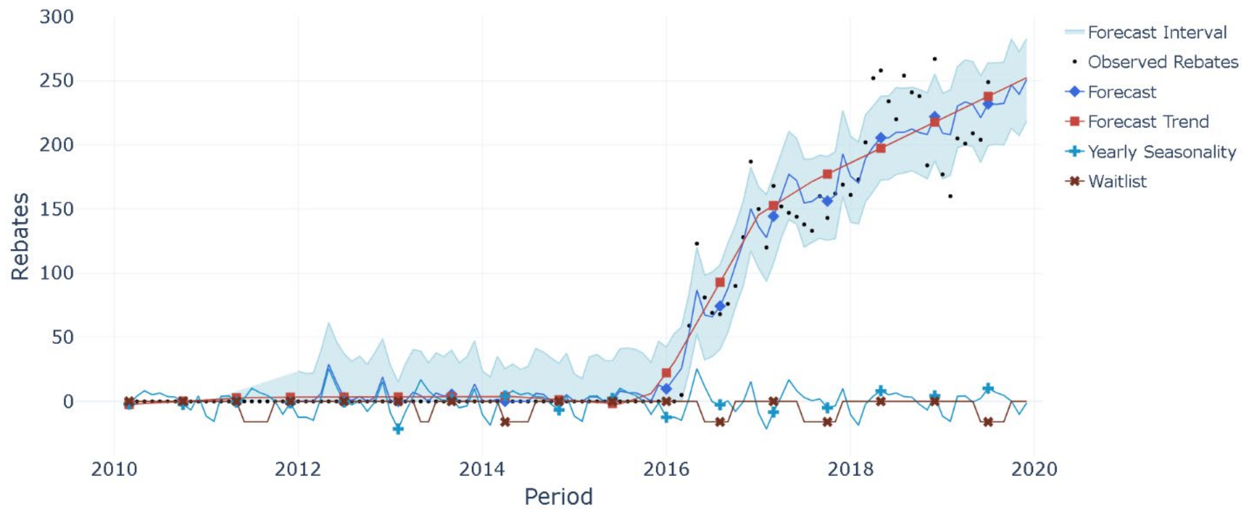

Figure 22. Baseline PHEV individual increased rebate forecast.

Cross-Validation Dates

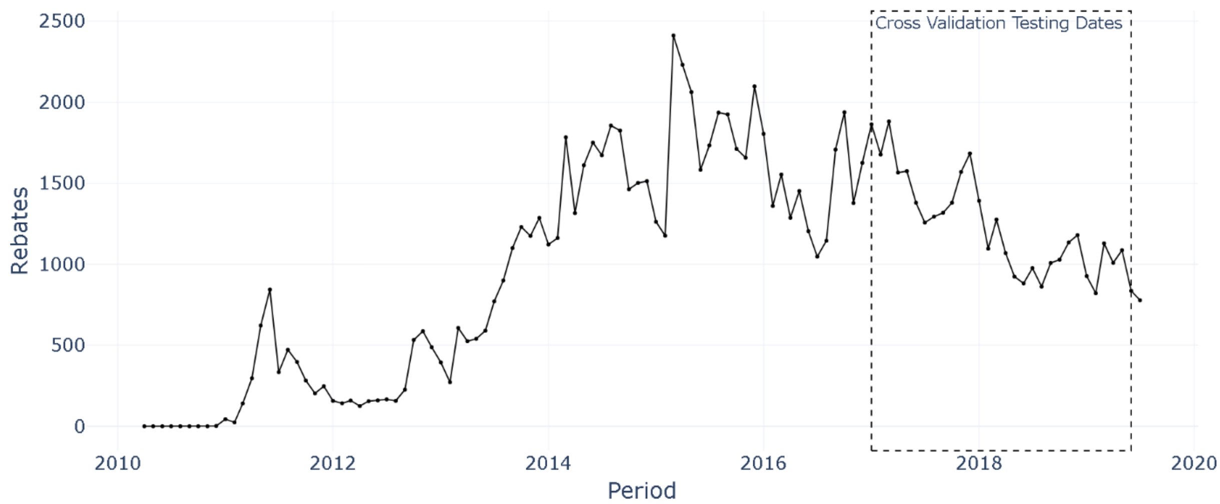

Figure 23. Individual standard non-Tesla BEV rebates with the cross-validation date range boxed with a dashed line.

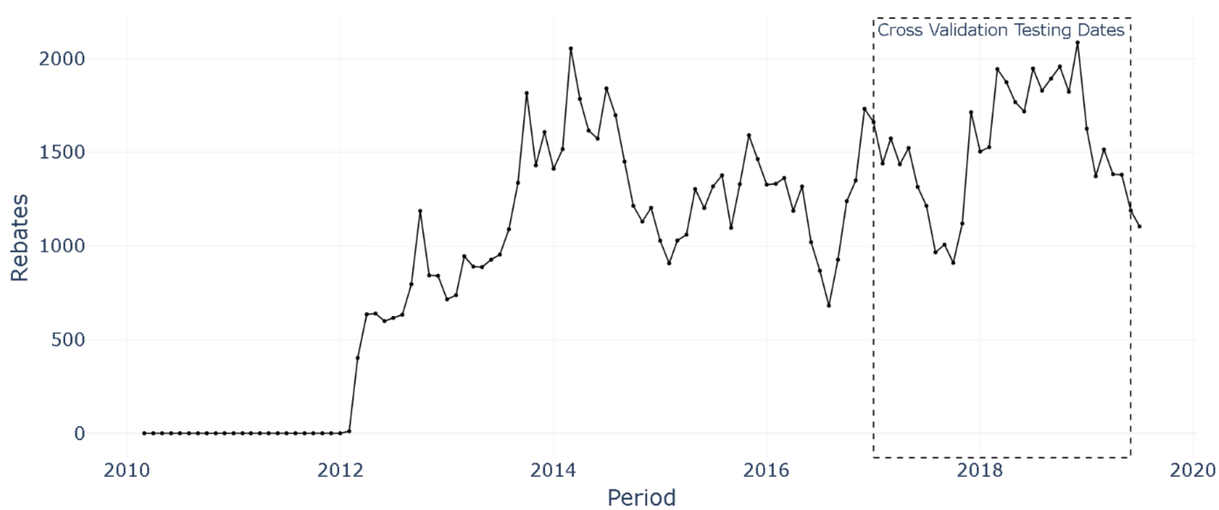

Figure 24. Individual standard PHEV rebates with the cross-validation date range boxed with a dashed line.