Forecasting the Post-COVID Recovery of Washington DC Public Transportation Usage

Steven VanOmmeren   Predictive Analytics Midterm*

April 6, 2025

I. Introduction

The COVID-19 lockdowns of early 2020 saw the use of public transportation plummet. Since lockdown restrictions have loosened, the use of public transportation has slowly increased and in some cases exceeds pre-pandemic levels. Residents of the DC-Maryland-Virginia (DMV) area heavily rely on Washington DC's public rail system, run by the Washington Metropolitan Area Transit Authority (WMATA). I use daily ridership data provided by WMATA to analyze rail transport usage from March 2020 through March 2024. I then make predictions of average daily ridership for April 2024 through March 2025. I employ a range of statistical models and compare their predictive accuracy, finding that ridership can be reasonably predicted using simple models. I also compare my model predictions with results from the academic literature and WMATA's own budget forecasts.

Forecasting is the prediction of future events using historical information. Accurate forecasts are vitally important to ensuring that appropriate business and government decisions are made. When sufficient data are available, statistical forecasting methods can approximate the time-series behavior of relevant variables using historical data of both the variable

*A complete replication package of this project is available at https://github.com/svanomm/forecasting-midterm-2025.

of interest and any additional relevant variables that may increase a model's performance. In this paper, I use a variety of statistical forecasting methods, some of which make use of an exogenous control variable, to predict WMATA ridership.

Forecasting long-term ridership patterns is critical to transportation systems. WMATA is required to predict its future costs and revenues in its annual budgets for government approval; revenue predictions require an accurate estimate of future ridership. Additionally, there has been much discussion of the lingering effects of COVID-19 on public transportation usage, and whether the current state is still trending upwards or a "new normal."

With 5 years passed since the start of COVID-19 lockdowns, we have sufficient data to analyze the post-COVID recovery of train ridership in Washington DC. I analyze the average daily train ridership per month, leaving a prediction of bus rides for future work.[1]

II. Literature Review

Because predicting the use of public transportation is vitally important to government budgeting, there is a substantial literature on modeling ridership. However, very few papers discuss the out-of-sample fit of their chosen models. Even fewer papers have analyzed post-COVID recovery given its relative recency.

Mucci and Erhardt (2018) evaluate the out-of-sample forecasting accuracy of a popular model of transportation volume, called a direct ridership model (DRM). As they explain, DRMs are commonly used to evaluate a city's public transportation volume, and can accommodate all modes of transportation (including rail). DRMs continue to be used even in more recent research.[2] The authors train a DRM model of rail ridership using San Francisco data from 2009, and they then use the model to predict ridership in 2016. They compare the model's predictions against actual ridership data for 2016 and find that the model underpre-

[1] While bus ride data are easily available and would likely improve my in-sample models, it is impractical to include them for out-of-sample predictions. This is because out-of-sample forecasts would require a separate predictive model of bus rides, which is beyond the scope of this paper.

[2] For example, see Pinho et al., 2024, published in December of 2024.

dicts ridership by 18%.[3] The model also underpredicts ridership for 2009 (i.e., in-sample) by 13%.[4]

In a March, 2023 news issue, Jared Brey reviews the increased uncertainty inherent in ridership predictions for major cities in the post-COVID world. The author highlights that San Francisco, Washington DC, and New York City governments predict ridership significantly below pre-pandemic levels as late as the end of 2026.

The article also notes that control variables that used to be informative predictors before COVID, are no longer as useful: "Some things, like the price of gas, now seem to have less of an impact on transit ridership than they did before the pandemic."[5] Finally, the author cites recent research that indicates government forecasts are biased upwards. Hoque et al. (2024) review 164 transit infrastructure projects, finding that the vast majority overpredicted ridership, with actual volume nearly 25% less than predicted, on average.

Finally, WMATA itself has posted its forecasts of future ridership.[6] In its proposed budget for fiscal year 2026 (Jul. 2025 through Jun. 2026), WMATA predicted that FY2025 would see 113.7 million Metrorail trips, or a 7.7% decrease from FY2024. While rail ridership is expected to flatten in mid-2024 to mid-2025, WMATA predicts that ridership will increase in the following years above FY2024 levels. But the forecasted increase could be subject to the bias discussed in Hoque et al. (2024).

[3]Mucci and Erhardt (2018), Table 5. Predicted ridership: 131,721. Actual: 161,387.
[4]Mucci and Erhardt (2018), Table 5. Predicted ridership: 128,042. Actual: 147,470.
[5]Brey (2023).
[6]https://www.wmata.com/initiatives/budget/upload/FINAL-FY2026-Proposed-Budget-REMEDIATED.pdf.

Figure 1: WMATA Forecasts of Rail Ridership (FY2026 Budget, p. 6)

RIDERSHIP BY SERVICE

(Trips in Thousands)	FY2023 Actual	FY2024 Actual	FY2025 Budget[2]	FY2026 Budget[2]	FY2027 Forecast[4]
Metrorail[1]	95,813	123,166	113,713	133,717	136,392
Metrobus[2]	102,477	117,540	111,408	133,069	135,730
MetroAccess	1,394	1,397	1,481	1,630	1,630
Total Ridership[3]	**199,684**	**242,103**	**226,603**	**268,416**	**273,752**

1. Ridership statistics beginning in January 2024 include both tapped and non-tapped ridership for Metrorail.
2. Metrobus ridership reflects Automated Passenger Count (APC) data
3. Metrorail ridership is based on linked trips; Metrobus ridership is based on unlinked trips from APC data; MetroAccess ridership is based on total passengers. Unlinked trips are total boardings, while linked trips are total number of complete trips from origin to destination, including transfers
4. Forecast is **subject to change** and is for planning purposes only. Based on current economic and ridership growth assumptions.

III. Methods

I downloaded daily rail rider data from WMATA and then aggregated the data to the monthly level, averaging the daily ridership counts. Separately, I obtained monthly average gasoline prices from the U.S. Energy Information Administration (EIA).[7] After combining and graphing the data, I did not see any extreme outliers that required additional consideration.

After combining the time series and formatting the appropriate date columns, I analyzed the seasonality of ridership using a seasonal subseries plot and STL decomposition.

Next, I split the data into training and test subsets, with the first 48 months (Apr. 2020 - Mar. 2024) used as training data. I used the training data to run 5 models of monthly job openings:

1. Seasonal Naïve: $\widehat{\text{Rides}_t} = \text{Rides}_{t-12}$
2. ETS: includes a damped-additive trend component and additive seasonal component.
3. TS Regression: regress Rides on gas price, seasonal indicators, and a piecewise-linear trend with knots at Jan. 2021, Jan. 2022, Jan. 2023, and Jan. 2024.

[7] Specifically, I use the monthly price of Regular grade gasoline in the Central Atlantic region. Data are available here: `https://www.eia.gov/dnav/pet/pet_pri_gnd_dcus_r1y_m.htm`. Because March price data was not available yet, I used the weekly price files to estimate an average price for March. The weekly data are available here: `https://www.eia.gov/dnav/pet/pet_pri_gnd_dcus_r1y_w.htm`.

4. ARIMAX: includes gas price as an exogenous regressor. The parameters are chosen automatically via an optimization technique.

5. Ensemble: a simple average of the above models.

Additionally, because forecasting the TS Regression and ARIMAX models requires predictions of the exogenous variables, I used a Seasonal ARIMA$(1, 1, 1)$ model to forecast the future price of gasoline.

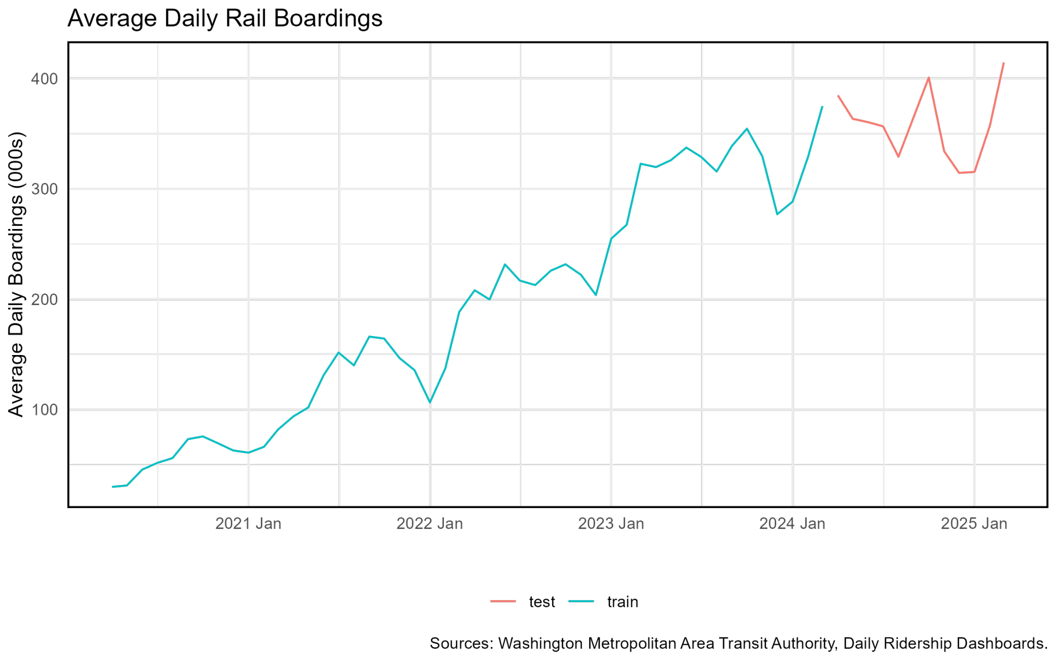
After running these models on the training data, I forecasted the next 12 months of ridership (Apr. 2024 - Mar. 2025) and compared the forecasts to the actual values in the test data.

IV. Results and Analysis

The time series of interest for my analysis is shown below. During the 5-year period following COVID, average daily rail use in Washington DC has consistently increased, though with noticeable seasonality: early Spring months appear to have higher volumes, while November through January appear to have lower volumes.

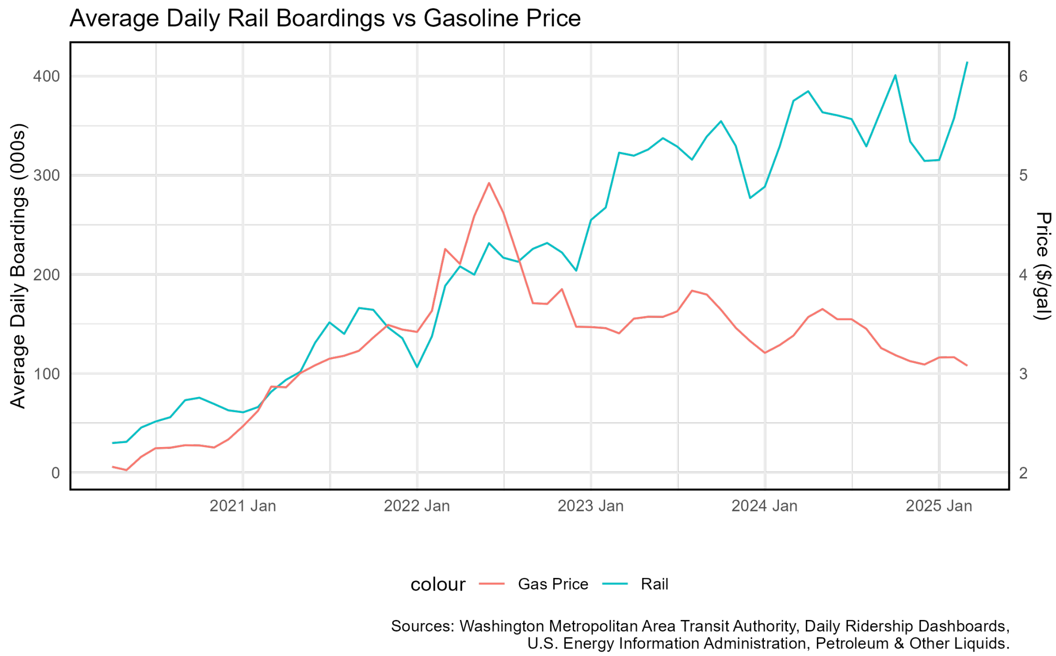
Interestingly, the most recent 12 months (in red) do not appear to follow the historical trend. While there is substantial monthly variation, ridership appears to be flattening its post-COVID recovery.

Figure 2: Variable of Interest, Training and Testing Periods



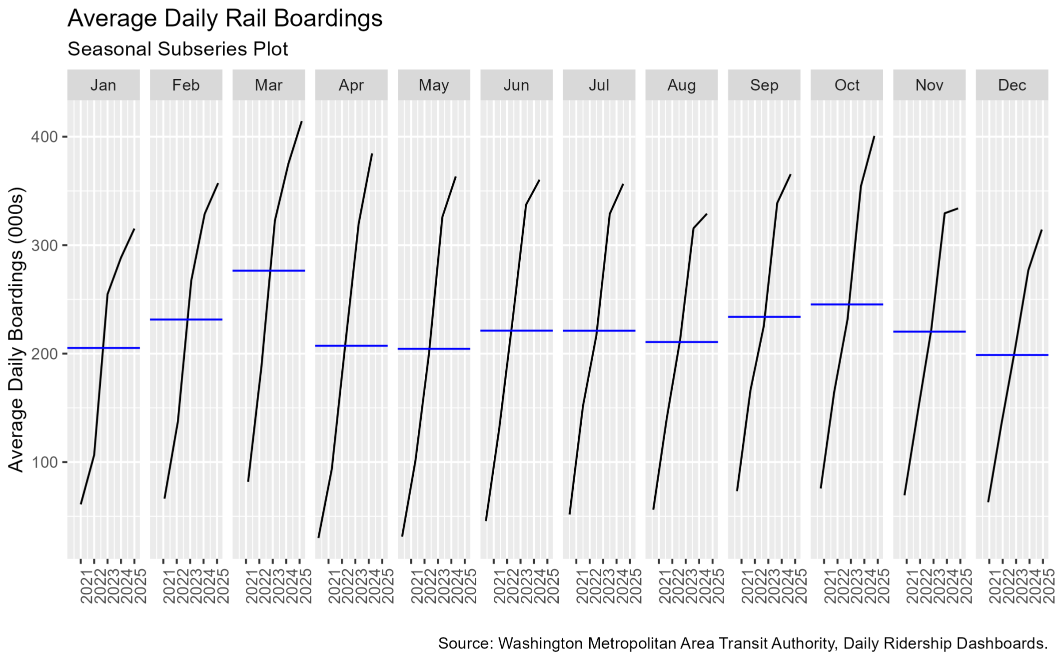
One variable that could help explain train ridership is gasoline prices. If gasoline prices are higher, public transportation becomes a relatively cheaper (and thus more attractive) option to some people who would otherwise drive to work. When overlaying the two time series, we see that while gas prices (in red) appear to have the same trend as rail ridership through mid-2022, the trend then reverses, decreasing to just over $3 per gallon in March 2025. This is consistent with the quote mentioned above that certain variables have become less useful for predicting ridership post-pandemic.

Figure 3: Average Ridership vs Gas Prices



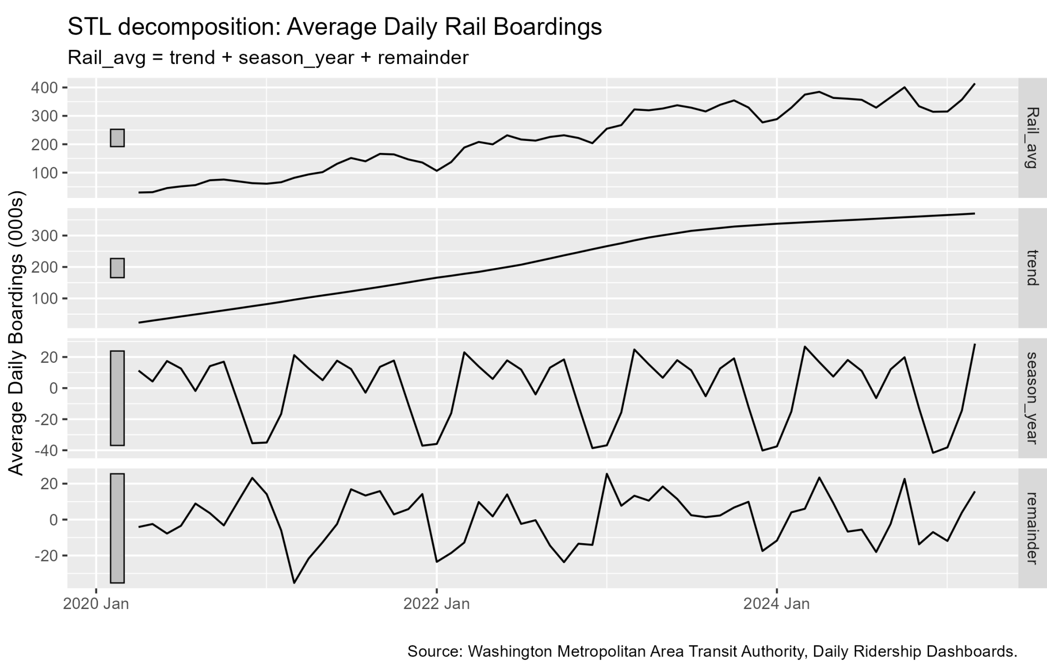
The seasonal subseries plot of rail volume confirms that there is ridership seasonality. Particularly, March is a noticeably higher-volume month than the others. This could correspond to the blooming of cherry blossom trees in Washington DC, which is a notably popular tourist activity.

Figure 4: Seasonal Subseries Plot, Rail Ridership



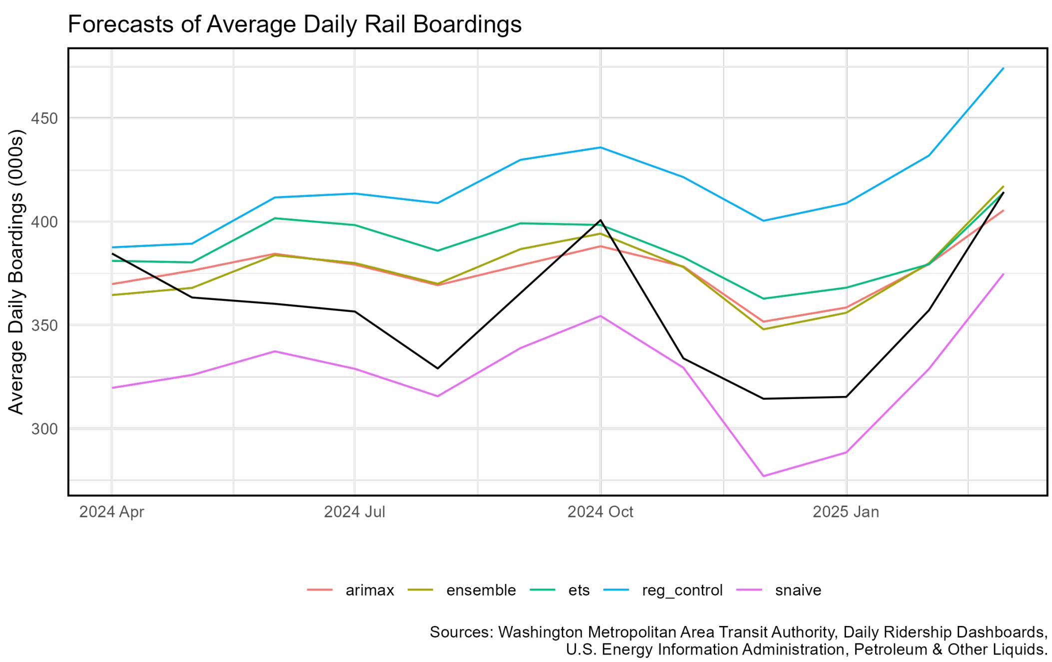
Applying STL decomposition to ridership shows the trend component has increased linearly until around 2023, where the trend increases at a slower rate.

Figure 5: STL Decomposition, Rail Ridership



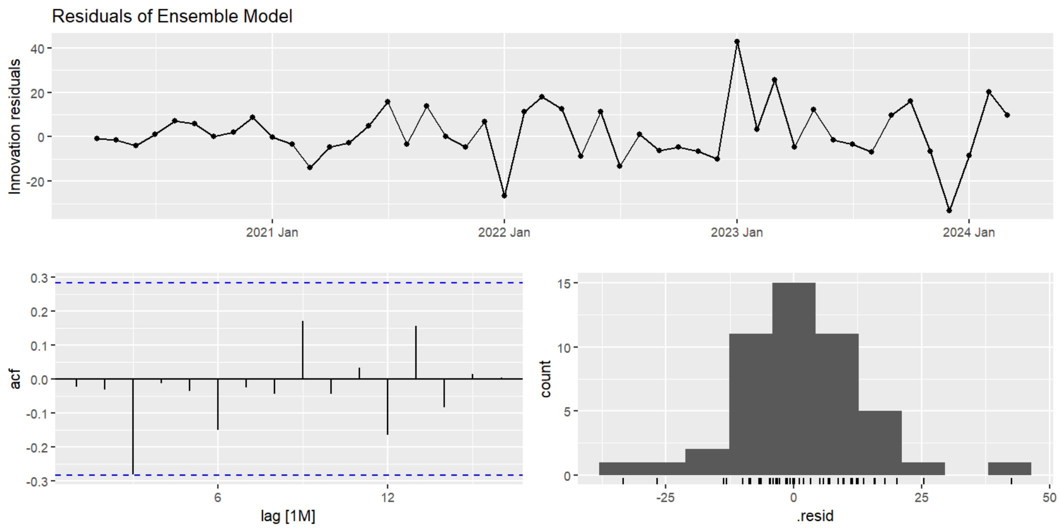
I now report my forecasts of ridership for the test data. All models capture the monthly seasonal pattern in the test data (black line), but they underestimate the strength of seasonality for the 2024-25 period. The TS Regression and Seasonal Naïve models appear furthest from the actual data on average, with "snaive" (purple line) consistently underestimating and "reg_control" (blue line) consistently overestimating ridership. Of the remaining models, "ets" (green line) appears to perform worst, consistently overestimating the truth, though it is significantly closer than the TS Regression. Finally, the ARIMAX and Ensemble models fit the data best, though they underestimate the strength of monthly seasonality in the test data.

Figure 6: Out-of-Sample Forecasts, Rail Ridership



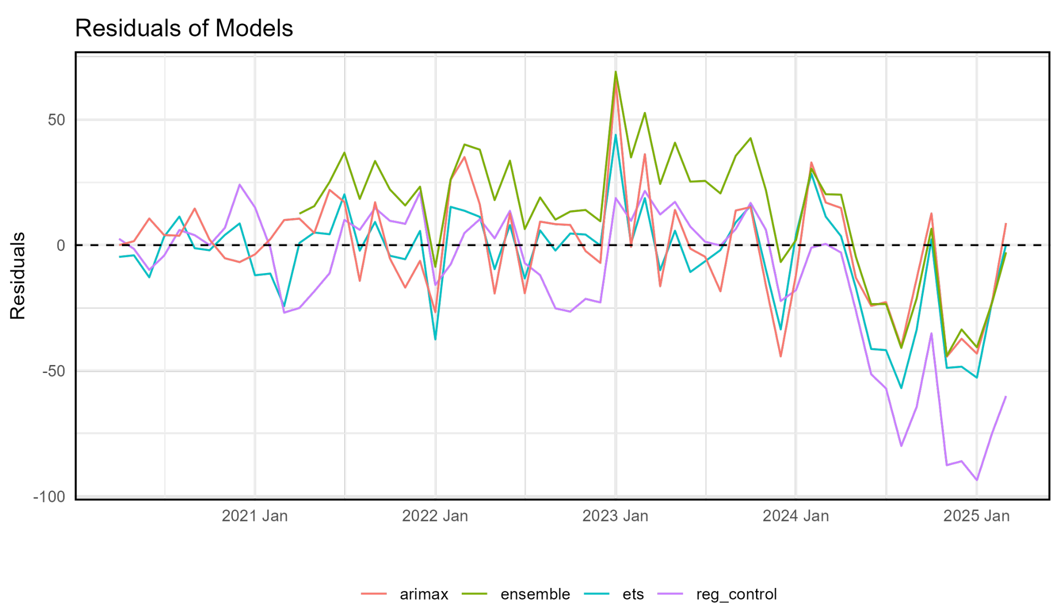
I analyzed the residuals of the Ensemble model and found that they do not exhibit significant autocorrelation, they do not display a discernible trend, and they appear to be normally distributed in the histogram below. This suggests that the Ensemble model can be used reliably for forecasting.

Figure 7: In-Sample Residuals of Ensemble Model



I then plotted the residuals of all models over the entirety of the data. Importantly, this includes the out-of-sample residuals as well. The below chart shows that the in-sample residuals are all quite similar, but by April 2024 they diverge significantly from the horizontal line at y = 0. Interestingly, while the Ensemble model's in-sample residuals appear to be among the worst, the out-of-sample residuals are closest to 0 and visually appear the best.

Figure 8: In- and Out-of-Sample Residuals



Sources: Washington Metropolitan Area Transit Authority, Daily Ridership Dashboards, U.S. Energy Information Administration, Petroleum & Other Liquids.

Finally, to quantify the accuracy of my models, I compared both the in- and out-of-sample accuracy metrics for each of them.

Table 1: In- and Out-of-Sample Performance Statistics

.model	.type	ME	RMSE	MAE	MPE	MAPE	MASE	RMSSE	ACF1
ets	Training	1.12	13.98	10.28	-1.02	7.59	0.12	0.15	-0.01
reg_control	Training	0.00	14.09	11.58	-0.26	8.47	0.13	0.15	0.56
arimax	Training	3.83	18.77	14.10	2.54	8.37	0.16	0.20	-0.01
ensemble	Training	23.95	28.41	24.81	10.96	11.54	0.28	0.31	0.16
snaive	Training	89.29	93.14	89.29	42.26	42.26	1.00	1.00	0.61
ensemble	Test	-19.27	27.32	23.70	-5.80	6.95			0.15
arimax	Test	-18.72	27.78	24.75	-5.71	7.23			0.09
snaive	Test	31.35	34.72	31.35	8.63	8.63			-0.13
ets	Test	-29.75	36.70	30.77	-8.82	9.08			0.17
reg_control	Test	-59.94	65.46	59.94	-17.31	17.31			0.43

V. Discussion

Overall, the Ensemble model is the best predictive model of the five models tested. This is a surprising result when reviewing the in-sample performance of the models. On the training data, the ETS model has the lowest RMSE, MAE, and MAPE, while the Seasonal Naïve model performs worst by a large margin. But when we look at out-of-sample performance, the results are reversed: ETS performs worse than Seasonal Naïve. This is because the Seasonal Naïve model implicitly assumes no growth year-over-year, which happened to be closer to reality than the ETS model's assumption that existing trends would continue going forward. The test data shows a decreased trend that is not inherent in the training data. When averaging the models with the Ensemble method, the underpredictions and overpredictions balance out and result in a generally well-fitting model.

Compared to Mucci and Erhardt (2018), my best model had a mean average percent error of 6.95%, versus their 18% out-of-sample error. Similarly, Hoque et al. (2024) reported an average of 25% overestimation of ridership in public transportation projects. While there are differences in estimation strategies, my Ensemble model is well within the out-of-sample forecasting accuracy of the academic literature.

WMATA forecasted ridership for the Jul. 2024 through Jun. 2025 period to be 113.7 million. The available data show a total ridership of 127.7 million during my test period (Apr. 2024 - Mar. 2025).[8] Using my ensemble model, I predicted ridership of 137.7 million, an overestimate of 7.8% compared to actual data.[9] Interestingly, my estimates from the ensemble model are closer to WMATA's forecasted ridership for FY2026, rather than 2025.[10] My results show that WMATA's forecasted decrease going into 2025 does not appear to be correct, and that post-COVID ridership recovery is still ongoing.

[8]While my test period is not exactly the same as WMATA's FY2025 forecast, June and July are historically average months for ridership, and so the actual total ridership for FY2025 will likely be higher than WMATA's forecast.

[9]My estimate of 137.7 million comes from multiplying the model's predicted average daily ridership by the number of days in each month, then summing across months.

[10]WMATA forecasted ridership for FY2026: 133.7 million trips.

While it is unclear what ridership forecasting methods WMATA uses, it appears that relatively simple forecasting techniques can meet or exceed the out-of-sample accuracy of their forecasts. While my results could have simply been due to chance, and I have only forecasted 12 months of future ridership versus the multiple years by WMATA, the government should formally review its forecasting techniques to find areas of improvement. Given their access to vast amounts of transportation data, WMATA should find ways to improve its forecasts, which will improve its budgeting accuracy.

My models used relatively simple forecasting techniques, but a more sophisticated analysis would further improve forecasting accuracy. To improve my model, I would source additional exogenous variables that influence rail ridership. Following Mucci and Erhardt (2018), I could modify my analysis to forecast train station-level ridership, rather than system-wide volume. Forecasting at the station level could result in more accurate predictions by incorporating geography-specific controls. Lastly, I would incorporate additional modes of transportation into the model to account for substitution from rail to buses, cars, bicycles, etc.

VI. Conclusion

Accurate forecasting of public transportation usage is a difficult task in the wake of COVID-19 recovery. Hypothesizing that rail ridership in the DMV area could be forecasted using time series techniques and gasoline prices, I trained five different time series models on 2020-24 data to forecast 2025 rail ridership. The most useful model was an Ensemble of the other four models, averaging their predictions. Even with relatively simple techniques, my models have out-of-sample predictive accuracy on par with academic research and WMATA's own forecasts. My results show that Washington DC rail ridership appears to still be on a post-COVID recovery path, with transportation volume continuing to increase in the short term.

References

Brey, Jared (2023). "Predicting Future Transit Ridership Is Trickier Than Ever". In: *Governing*. URL: https://www.governing.com/community/predicting-future-transit-ridership-is-trickier-than-ever.

Hoque, Jawad Mahmud et al. (2024). "Are public transit investments based on accurate forecasts? An analysis of the improving trend of transit ridership forecasts in the United States". In: *Transportation Research Part A: Policy and Practice* 186, p. 104142. ISSN: 0965-8564. DOI: https://doi.org/10.1016/j.tra.2024.104142. URL: https://www.sciencedirect.com/science/article/pii/S0965856424001903.

Mucci, Richard A. and Gregory D. Erhardt (2018). "Evaluating the Ability of Transit Direct Ridership Models to Forecast Medium-Term Ridership Changes: Evidence from San Francisco". In: *Transportation Research Record* 2672.46, pp. 21–30. DOI: 10.1177/0361198118758632.

Pinho, Paulo et al. (2024). "The application of direct ridership models in the evaluation of the expansion of the Porto Light Rail Transit". In: *Case Studies on Transport Policy* 18, p. 101282. ISSN: 2213-624X. DOI: https://doi.org/10.1016/j.cstp.2024.101282. URL: https://www.sciencedirect.com/science/article/pii/S2213624X24001378.

Appendix: Code

Below is the R code used to generate my results.

```
library(fpp3)
library(here)
library(stargazer)
library(dplyr)
library(readr)
library(lubridate)
library(readxl)

here()

rides      <- read_csv(here("./data/Daily Ridership - Data View_data.csv"))
gas_prices <- read_excel(here("./data/EMM_EPMR_PTE_R1Y_DPGm.xls"), sheet = "Data 1", range = "A3:B386")
```

```r
# reshape rides data wide by Mode
rides_wide <- rides %>%
  pivot_wider(names_from = Mode, values_from = "Entries Or Boardings")

# Convert the datetime column to a date
rides_wide$Date <- mdy(format(mdy_hms(rides_wide$Date), "%m/%d/%Y"))

# Divide columns by 1000
rides_wide$Bus  <- rides_wide$Bus  / 1000
rides_wide$Rail <- rides_wide$Rail / 1000

# Filter to date >= March 2020
rides_wide <- rides_wide %>%
  filter(Date >= mdy("04/01/2020"))

# Aggregate to monthly level
rides_monthly <- rides_wide %>%
  group_by(Date = floor_date(Date, "month")) %>%
  summarise(Rail_tot = sum(Rail),
            Bus_tot  = sum(Bus),
            Rail_avg = mean(Rail),
            Bus_avg  = mean(Bus))

# Change gas_prices to report day 01 instead of 15
gas_prices$Date <- floor_date(gas_prices$Date, "month")
gas_prices$Date <- as.Date(gas_prices$Date)
gas_prices <- gas_prices |> rename(gas_price = `Central Atlantic (PADD 1B) Regular All Formulations Retail Gasoline Prices (Dollars per Gallon)`)

# Merge gas prices with rides data
data <- left_join(rides_monthly, gas_prices, by = "Date")
data$Date <- yearmonth(data$Date)

# add March gas price data using https://www.eia.gov/dnav/pet/pet_pri_gnd_dcus_r1y_w.htm
# March weekly prices: 3.118 3.071 3.054 3.075. Average: 3.0795
data <- data |> mutate(
  gas_price = ifelse(Date == yearmonth("2025 Mar"), 3.0795, gas_price)
)

# Convert data to tsibble
data <- as_tsibble(data, index = Date)

# Seasonality graph
data |> gg_subseries(y=Rail_avg) +
  labs(title = "Average Daily Rail Boardings",
       y="Average Daily Boardings (000s)",
       subtitle = "Seasonal Subseries Plot",
       x="",
       caption = "Source: Washington Metropolitan Area Transit Authority, Daily Ridership Dashboards.")
ggsave(here("./analysis/output/graphs/Seasonality Plot.png"), width = 8, height = 5)

# Split the data
train <- data |> slice(1:48)
test  <- data |> slice(49:60)

# Create variable train_test which is "train" for observations 1:48 and "test" else
data <- data |> mutate(
```

```
  train_test = if_else(row_number() <= 48, "train", "test")
  )

# Plot Rail_avg over time, colored by train_test
ggplot(data, aes(x = Date)) +
  # Plot Rail_avg on primary scale
  geom_line(aes(y = Rail_avg, color = train_test)) +
  labs(title = "Average Daily Rail Boardings",
       y = "Average Daily Boardings (000s)",
       x = "",
       caption = "Sources: Washington Metropolitan Area Transit Authority, Daily Ridership Dashboards.") +
  theme_minimal() +
  theme(
    legend.position = "bottom",
    panel.border = element_rect(colour = "black", fill=NA, linewidth=1)
  ) +
  guides(colour = guide_legend(title = ""))
ggsave(here("./analysis/output/graphs/train_test.png"), width = 8, height = 5)


# Plot Rail vs gas price
ggplot(data, aes(x = Date)) +
  # Plot Rail_avg on primary scale
  geom_line(aes(y = Rail_avg, color = "Rail")) +
  # Scale up gas_price values for plotting on the primary scale
  geom_line(aes(y = gas_price * 100 - 200, color = "Gas Price")) +
  scale_y_continuous(
    name = "Average Daily Boardings (000s)",
    # Transform back to original gas price values for the secondary axis labels
    sec.axis = sec_axis(~ (.+ 200) / 100 , name = "Price ($/gal)")
  ) +
  labs(title = "Average Daily Rail Boardings vs Gasoline Price",
       x = "",
       caption = "Sources: Washington Metropolitan Area Transit Authority, Daily Ridership Dashboards,\n
       U.S. Energy Information Administration, Petroleum & Other Liquids.") +
  theme_minimal() +
  theme(
    legend.position = "bottom",
    panel.border = element_rect(colour = "black", fill=NA, linewidth=1)
  )
ggsave(here("./analysis/output/graphs/rides_vs_gas.png"), width = 8, height = 5)


STL_decomp <- data |>
  model(stl = STL(Rail_avg))

STL_decomp |> components() |>
  autoplot() +
  labs(title = "STL decomposition: Average Daily Rail Boardings",
       y="Average Daily Boardings (000s)",
       x="",
       caption = "Source: Washington Metropolitan Area Transit Authority, Daily Ridership Dashboards.")
ggsave(here("./analysis/output/graphs/STL_Decomp.png"), width = 8, height = 5)

# For forecasting, we need future values of gas prices.
# Use ARIMA to forecast gas prices
```

```r
gas_model <- train |>
  model(ARIMA(gas_price ~ pdq(1, 1, 1) + PDQ(1, 1, 1))) |>
  forecast(new_data = test)
# Update the testing data with the predicted gas prices
test <- test |> mutate(
  gas_price = gas_model$.mean
)

knot1 <- yearmonth("2021 Jan")
knot2 <- yearmonth("2022 Jan")
knot3 <- yearmonth("2023 Jan")
knot4 <- yearmonth("2024 Jan")

# Fit models
my_models <- train |>
  model(
    snaive = SNAIVE(Rail_avg),
    ets    = ETS(Rail_avg ~ trend("Ad") + season("A")),
    reg_control = TSLM(Rail_avg ~ gas_price + season() + trend(knots = c(knot1,knot2,knot3,knot4))),
    arimax = ARIMA(Rail_avg ~ gas_price)
    ) |>
  mutate(ensemble = (snaive+ets+reg_control+arimax)/4)

my_forecasts <- my_models |>
  forecast(new_data = test)

# Create a dataset with the residuals for my models
r_snaive      <- my_models |> select(snaive     ) |> residuals() |> select(Date, .resid) |> rename(snaive      = .resid)
r_ets         <- my_models |> select(ets        ) |> residuals() |> select(Date, .resid) |> rename(ets         = .resid)
r_reg_control <- my_models |> select(reg_control) |> residuals() |> select(Date, .resid) |> rename(reg_control = .resid)
r_arimax      <- my_models |> select(arimax     ) |> residuals() |> select(Date, .resid) |> rename(arimax      = .resid)
r_ensemble    <- my_models |> select(ensemble   ) |> residuals() |> select(Date, .resid) |> rename(ensemble    = .resid)

r_train <- r_snaive |>
  inner_join(r_ets, by = "Date") |>
  inner_join(r_reg_control, by = "Date") |>
  inner_join(r_arimax, by = "Date") |>
  inner_join(r_ensemble, by = "Date")

r_snaive      <- as.data.frame(my_forecasts) |> filter(.model == "snaive"     ) |> select(Date, .mean) |> rename(snaive = .mean)
r_ets         <- as.data.frame(my_forecasts) |> filter(.model == "ets"        ) |> select(Date, .mean) |> rename(ets = .mean)
r_reg_control <- as.data.frame(my_forecasts) |> filter(.model == "reg_control") |> select(Date, .mean) |> rename(reg_control = .mean)
r_arimax      <- as.data.frame(my_forecasts) |> filter(.model == "arimax"     ) |> select(Date, .mean) |> rename(arimax = .mean)
r_ensemble    <- as.data.frame(my_forecasts) |> filter(.model == "ensemble"   ) |> select(Date, .mean) |> rename(ensemble = .mean)

r_test <- test  |>
  inner_join(r_snaive, by = "Date") |>
  inner_join(r_ets, by = "Date") |>
  inner_join(r_reg_control, by = "Date") |>
  inner_join(r_arimax, by = "Date") |>
  inner_join(r_ensemble, by = "Date")

r_test <- r_test |> mutate(
  snaive = Rail_avg - snaive,
  ets = Rail_avg - ets,
  reg_control = Rail_avg - reg_control,
```

```
  arimax = Rail_avg - arimax,
  ensemble = Rail_avg - ensemble
) |> select(Date, Rail_avg, snaive, ets, reg_control, arimax, ensemble)

# append r_train and r_test
residuals <- r_train |> bind_rows(r_test)

# Plot all the residuals over time
ggplot(residuals, aes(x = Date)) +
  geom_line(aes(y = ets, color = "ets")) +
  geom_line(aes(y = reg_control, color = "reg_control")) +
  geom_line(aes(y = arimax, color = "arimax")) +
  geom_line(aes(y = ensemble, color = "ensemble")) +
  geom_hline(yintercept = 0, linetype = "dashed", color = "black") +
  labs(title = "Residuals of Models",
       y = "Residuals",
       x = "",
       caption = "Sources: Washington Metropolitan Area Transit Authority, Daily Ridership Dashboards,\n
       U.S. Energy Information Administration, Petroleum & Other Liquids.") +
  theme_minimal() +
  theme(
    legend.position = "bottom",
    panel.border = element_rect(colour = "black", fill=NA, linewidth=1)
  ) +
  guides(colour = guide_legend(title = ""))
ggsave(here("./analysis/output/graphs/residuals.png"), width = 8, height = 5)


my_models |> select(ensemble) |> gg_tsresiduals() +
  labs(title = "Residuals of Ensemble Model",
       x="")

# Plot the model predictions on the testing data
my_forecasts |>
  autoplot(test, level = NULL) +
  theme_minimal() +
  theme(
    legend.position = "bottom",
    panel.border = element_rect(colour = "black", fill=NA, linewidth=1)
  ) +
  labs(
    y = "Average Daily Boardings (000s)",
    title = "Forecasts of Average Daily Rail Boardings",
    x="",
    caption = "Sources: Washington Metropolitan Area Transit Authority, Daily Ridership Dashboards,\n
    U.S. Energy Information Administration, Petroleum & Other Liquids.") +
  guides(colour = guide_legend(title = ""))
ggsave(here("./analysis/output/graphs/forecasts.png"), width = 8, height = 5)


# IS and OOS accuracy metrics
is_accuracy  <- my_models |> accuracy()
oos_accuracy <- accuracy(my_forecasts, test)
combined_accuracy <- rbind(is_accuracy, oos_accuracy) |> arrange(desc(.type), RMSE)

d <- as.matrix(mutate_if(
```

```
  combined_accuracy, is.numeric, ~round(., 2)
))
stargazer(d, out = here("./analysis/output/graphs/Accuracy Table.tex"), type = "latex")

stargazer(d, type = "text")

# Summarize ridership forecasts vs total actuals
as.data.frame(my_forecasts) |> group_by(.model) |> summarise(avg = mean(.mean), sum=sum(Rail_tot))
```