

# PATH Forecast :: Scenarios

*Planning & Regional Development*

*10/08/2020*

**DRAFT**

## Summary

Outputs using the PATH econometric forecasting model and October employment inputs (moderate and protracted). Ridership and Hudson County population inputs held constant as of February 2020.

One note of significance here is that our vendor, Oxford Economics, is projecting a decrease in *Hudson County population*, seen below. Our previous update (August) incorporated Oxford's expectation for reduced *employment* into the PATH forecast, but those expectations remain largely unchanged between August and October; this forecast incorporates the new expectations for *population*, which puts downward pressure on the weekend projections.

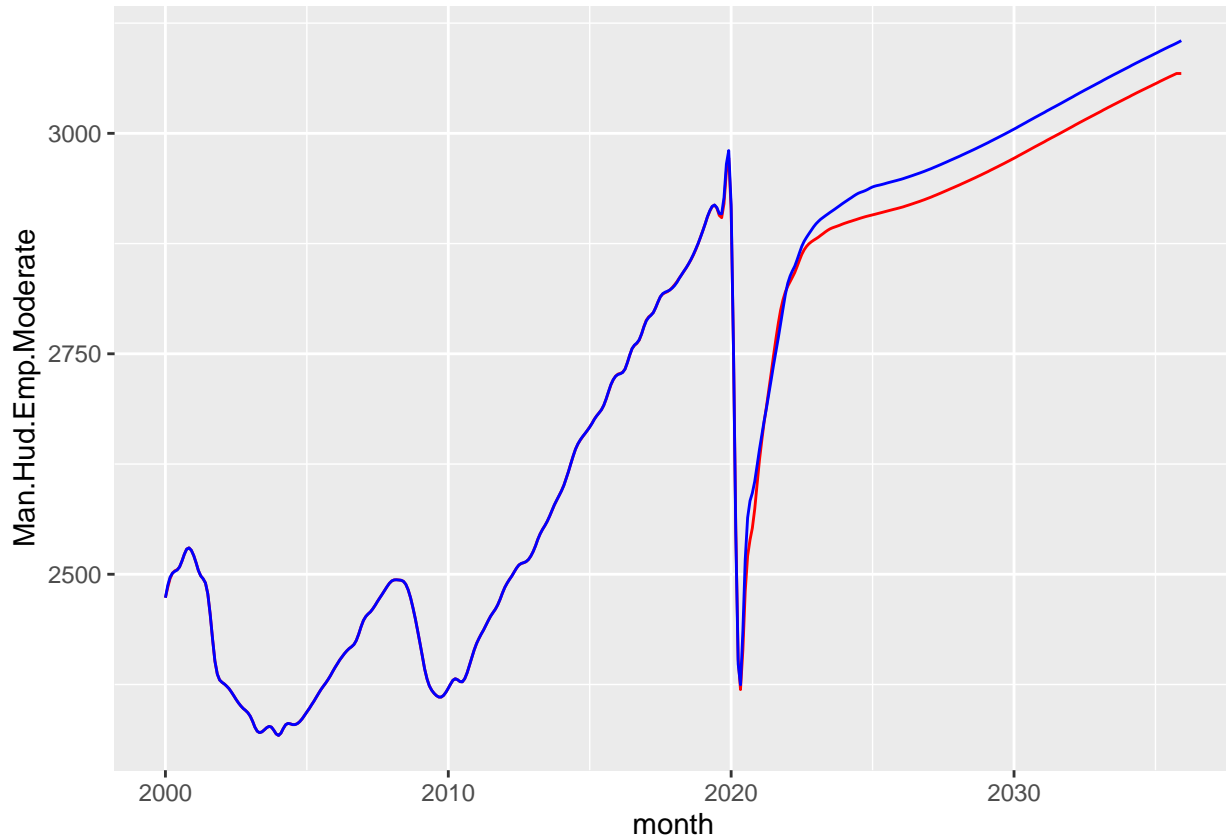
This model iteration only updates the moderate scenario. References here and in the data output files to the protracted scenario and associated data refer to the August run.

```

end = "2020-02-01"
end_and_one = "2020-03-01"
end_cpi = "2020-10-01"

path = read.csv("./input data/PATH_PaxCounts_2000-2009+2010-2020Apr.csv")
path$month = as.Date(paste(path$year, str_pad(path$month, 2,
  pad = "0"), "01", sep = "-"))
path$year = NULL

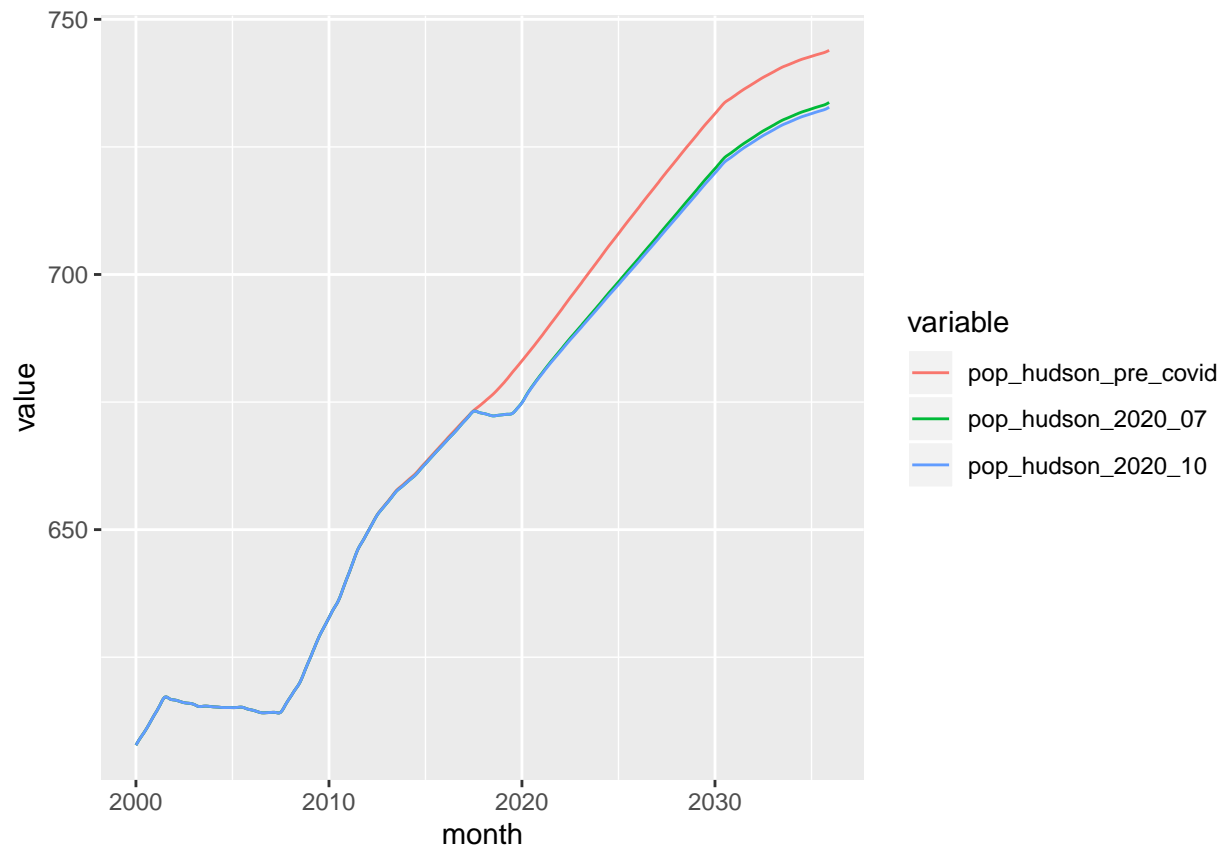
```



```

##      month      variable  value
## 1 2000-01-01 pop_hudson_pre_covid 607.7674
## 2 2000-02-01 pop_hudson_pre_covid 608.3153
## 3 2000-03-01 pop_hudson_pre_covid 608.8197
## 4 2000-04-01 pop_hudson_pre_covid 609.2940
## 5 2000-05-01 pop_hudson_pre_covid 609.7525
## 6 2000-06-01 pop_hudson_pre_covid 610.2143

```



```
path$cpi_base = path[path$month == end_cpi, "cpi_2020_10"]
path$real_farefare = ifelse(path$month <= end, path$fare_nominal *
  path$cpi_base/path$cpi_2020_10, max(path$fare_nominal))
path$cpi_base = NULL
```

```
before = subset(path, path$month <= end & path$month >= "2004-01-01") #before = head(path, 218)
after = subset(path, path$month > end) #after = tail(path, 250)
```

## Models are trained below

### Weekday:

```
fit = arima(ts(before$avg_wkdayholminor_tstile), xreg = oldreg,
  order = c(0, 0, 1), include.mean = T)
```

### Saturday:

```
fitsat = arima(ts(before$avg_satholmajor_tstile), xreg = oldregsat,
  order = c(1, 1, 0))
```

## Sunday:

```
fitsun = arima(ts(before$avg_sun_tstile), xreg = oldregsat, order = c(1,
1, 1))
```

Scenarios:

```
# Scenarios
```

```
fit_pess = arima(ts(before$avg_wkdayholminor_tstile), xreg = oldreg_pess,
order = c(0, 0, 1), include.mean = T)
```

```
pathpredict = predict(fit, n.ahead = forec_horizon, newxreg = newreg) # level=95 #interval = 'predicti
pathpredictsat = predict(fitsat, n.ahead = forec_horizon, newxreg = newregsat)
pathpredictsun = predict(fitsun, n.ahead = forec_horizon, newxreg = newregsat)
pathpredict_pess = predict(fit_pess, n.ahead = forec_horizon,
newxreg = newreg_pess)
```

```
pathpredict_by_month = as.data.frame(cbind(pathpredict$pred,
pathpredictsat$pred, pathpredictsun$pred, pathpredict_pess$pred))
names(pathpredict_by_month) = c("avg_wkdayholminor_tstile", "avg_satholmajor_tstile",
"avg_sun_tstile", "pess_wkdayholminor")
pathpredict_by_month$month = seq(as.Date(end) + extra, as.Date(future),
by = "mon")
```

```
# Annual
```

```
pathpredict_by_month$year = year(pathpredict_by_month$month)
pathpredict_year = summaryBy(sum_wkdayholminor + sum_satholmajor +
sum_sun + sum_wkday_pess ~ year, data = pathpredict_by_month,
FUN = sum)
names(pathpredict_year) = c("year", "base_wkday", "saturday",
"sunday", "pess_wkday")
pathpredict_year$base_total = pathpredict_year$base_wkday + pathpredict_year$saturday +
pathpredict_year$sunday
pathpredict_year$pess_total = pathpredict_year$pess_wkday + pathpredict_year$saturday +
pathpredict_year$sunday
pathpredict_by_month$year = NULL
```

```
resids = as.data.frame(cbind(as.vector(resid(fit)), as.vector(resid(fitsat)),
as.vector(resid(fitsun))))
names(resids) = c("Weekday_residuals", "Saturday_residuals",
"Sunday_residuals")
```

## Output

Save everything as:

```
sheets = list(Data = path, Monthly_Output = pathpredict_by_month,  
              Annual_Output = pathpredict_year, Residuals = resids)  
write_xlsx(sheets, "./output data/Output Scenarios 2020-10 moderate.xlsx") # This exports and names th
```

Monthly output included in Excel file within output folder.

## Modeling statistics and diagnostics

### Weekday

Table 1: Weekday Coefficients

term	estimate	std.error
ma1	0.5885689	0.0472597
intercept	-66282.8096549	24208.0306321
before.man_hud_2020_10	128.8082213	14.1473484
before.dummy_2	6426.3833752	3404.9581196
before.dummy_3	8426.3166080	4546.8194630
before.dummy_4	13129.0593896	4569.6670840
before.dummy_5	15944.6678247	4566.1461631
before.dummy_6	19318.5073280	4563.9967523
before.dummy_7	15364.7767125	4563.2171791
before.dummy_8	9405.2947265	4563.2764879
before.dummy_9	19911.0093160	4563.7788051
before.dummy_10	17389.8455551	4565.3080363
before.dummy_11	10387.4286993	4652.6775874
before.dummy_12	-86.7781006	3487.5670831
before.dum_911_base	-23354.3800842	11597.6582535
before.supersandy	-73984.8135960	10357.4091916
before.real_farefare	-10856.5751517	6454.1838132

Table 2: Weekday Diagnostics

sigma	logLik	AIC	BIC
11163.28	-2073.108	4182.215	4240.943

Table 3: Weekday Additional Diagnostics

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	53.44718	11163.28	8430.106	-0.2850188	3.67054	1.126848	0.281867

## Saturday

Table 4: Saturday Coefficients

term	estimate	std.error
ar1	-0.5584778	0.0605964
before.pop_hudson_2020_10	549.3503137	682.1749042
before.dummy_2	6775.4849976	1793.5748006
before.dummy_3	20283.4143515	1658.3919106
before.dummy_4	22276.2220370	1994.5973374
before.dummy_5	14117.6200361	1988.9668784
before.dummy_6	21120.2451209	2107.8399623
before.dummy_7	17171.6414829	2082.4104247
before.dummy_8	15858.8147034	2120.2042796
before.dummy_9	16956.6872526	2000.5308130
before.dummy_10	20736.7723373	2017.1004418
before.dummy_11	15020.1571936	1702.9567613
before.dummy_12	15649.6273790	1858.2067503
before.dum_911_base	853.3502320	6278.6621335
before.supersandy	-44329.9023363	4174.3251275
before.end_close	-12848.2703808	2122.0648323
before.real_farefare	-24648.7841993	9649.9580436

Table 5: Saturday Diagnostics

sigma	logLik	AIC	BIC
6073.014	-1945.457	3926.915	3985.55

Table 6: Saturday Additional Diagnostics

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	82.43527	6057.28	4730.936	-0.152776	4.462707	0.557604	-0.1646553

## Sunday

Table 7: Sunday Coefficients

term	estimate	std.error
ar1	-0.0431240	0.0876874
ma1	-0.7851897	0.0509257
before.pop_hudson_2020_10	398.5605843	196.3821379
before.dummy_2	5180.2513106	1562.2953613
before.dummy_3	8002.6356434	1572.8426600
before.dummy_4	14040.3429364	1584.6681467
before.dummy_5	18538.2303628	1589.9985177
before.dummy_6	23944.7554272	1593.8202180
before.dummy_7	17476.9404328	1595.5075987
before.dummy_8	15913.4650582	1598.7344349
before.dummy_9	20902.5666777	1595.1316504
before.dummy_10	17466.7852372	1589.3055886
before.dummy_11	14065.8019163	1619.9767252
before.dummy_12	16566.6639674	1611.0089071
before.dum_911_base	-2948.7808810	5024.2948038
before.supersandy	-37635.2456334	3684.4388240
before.end_close	-10194.6303104	1398.0206456
before.real_farefare	-13004.0105321	5568.2164566

Table 8: Sunday Diagnostics

sigma	logLik	AIC	BIC
4875.614	-1903.453	3844.905	3906.798

Table 9: Sunday Additional Diagnostics

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	215.1397	4862.981	3564.332	-0.0489996	4.552661	0.5186372	-0.010297



