PROJECT: FORECASTING THE FUTURE OF MIAMI All Employee & Average Weekly Wage

Time Series & Forecasting Analysis for Business

Abstract

Forecasting is an old but always important field of study for humanity. Forecasting is the study of predicting what possibly can happen in the future, it can be appl-y on daily subject like daily weather to annually prediction of global economics to decade long prediction for disaster. Thus, it is an ever-expanding field of study t-hat can be endlessly improved on. In the study below I am doing a forecasting for each of the monthly change for All Employees: Total Private in Miami-Fort Lauderdale-West Palm Beach, FL (MSA) and Average Weekly Earnings of All Employee-s: Total Private in Miami-Fort Lauderdale-West Palm Beach, FL (MSA)

Zheng, Zhi

Zhizheng0889@floridapoly.edu

1. Introduction

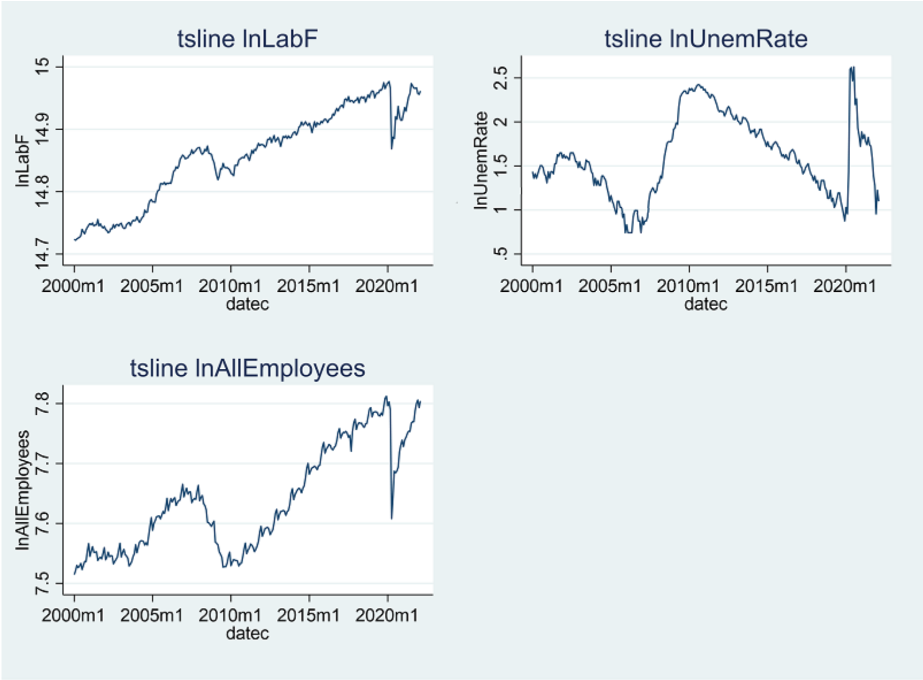
Forecasting is an essential field of study that assist in numerous decision-making situations, from personal matter to global-level strategy. The goal of forecasting is to allow an individual to a global party to see the possible futures based on past data. In this study, I will be forecasting the possible monthly change in the All Employees: Total Private in Miami-Fort Lauderdale-West Palm Beach, FL (MSA) and Average Weekly Earnings of All Employees: Total Private in Miami-Fort Lauderdale-West Palm Beach, FL (MSA). This will allow an insight into the current and future job market for Miami-Lauderdale-West Palm Beach after the pandemic impact known as Covid-19.

2. Data

The data used for this study came from the **Federal Reserve Economic Data** (Fred), a data source with frequently updated US macro and regional economic time series at annual, quarterly, monthly, weekly, and daily frequencies. The variables used in this study are non-seasonally adjusted monthly data of Average Weekly Earnings of All Employees: Total Private, Average Hourly Earnings of All Employees: Total Private, Average Weekly Hours of All Employees: Total Private, All Employees: Total Private, Civilian Labor Force, and Unemployment Rate in Miami-Fort Lauderdale-West Palm Beach, FL (MSA). The study's two variables to be predicted are the All Employees: Total Private in Miami-Fort Lauderdale-West Palm Beach, FL (MSA) and Average Weekly Earnings of All Employees: Total Private in Miami-Fort Lauderdale-West Palm Beach, FL (MSA).

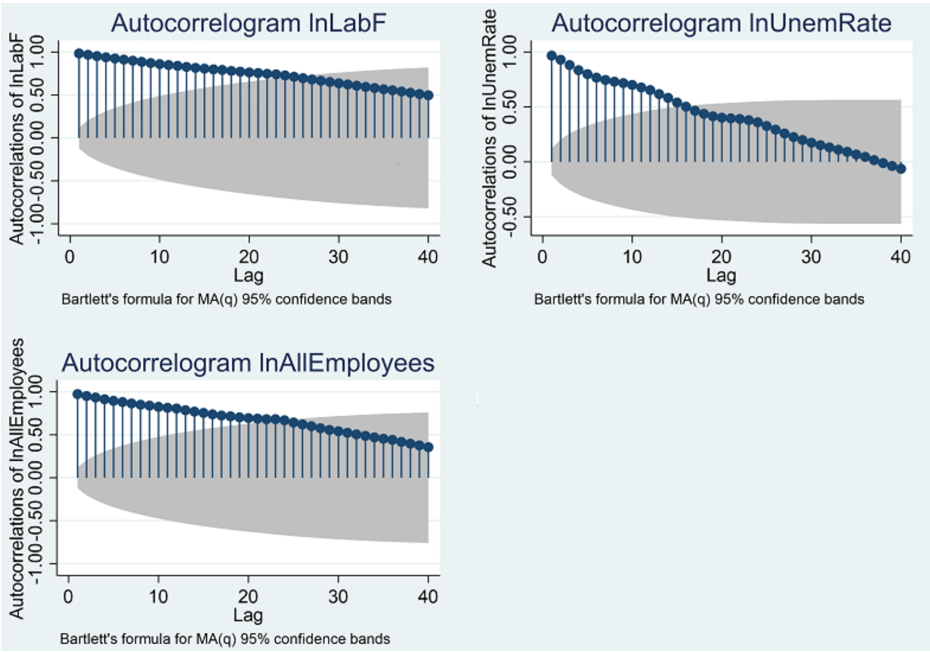
1) Summary statistics: All Employees: Total Private in Miami-Fort Lauderdale-West Palm Beach, FL (MSA)

After testing with some models, the best variables for predicting All Employees are the All Employees, Unemployment Rate, and Labor Force. Then, I generate some tsline graphs, autocorrelograms, and partial autocorrelograms to get a deeper insight into the data. First, start with tsline graphs show the Labor Force and All Employees have a steady nonstationary growth over time for most of the time, and the Unemployment Rate has an ongoing nonstationary downward trend for most of the time.

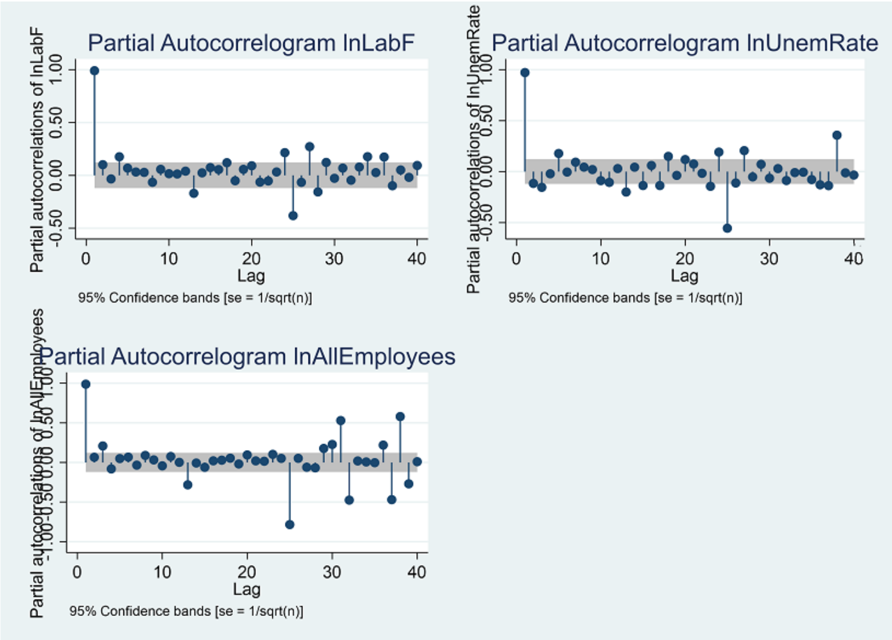


Graphs 1: Tsline Graphs for AllEmployees Predictors

The autocorrelograms shown for lnAllEmployees, lnLabF, and lnUmenRate are not helpful because they have too many lags outside the confidence region on their plots. Thus, I checked the partial autocorrelograms show fewer lags than the autocorrelograms for these variables, assisting me to determine whether using AR (1) model would be appropriate for these variables.



Graphs 2: Autocorrelogram Graphs for AllEmployees Predictors

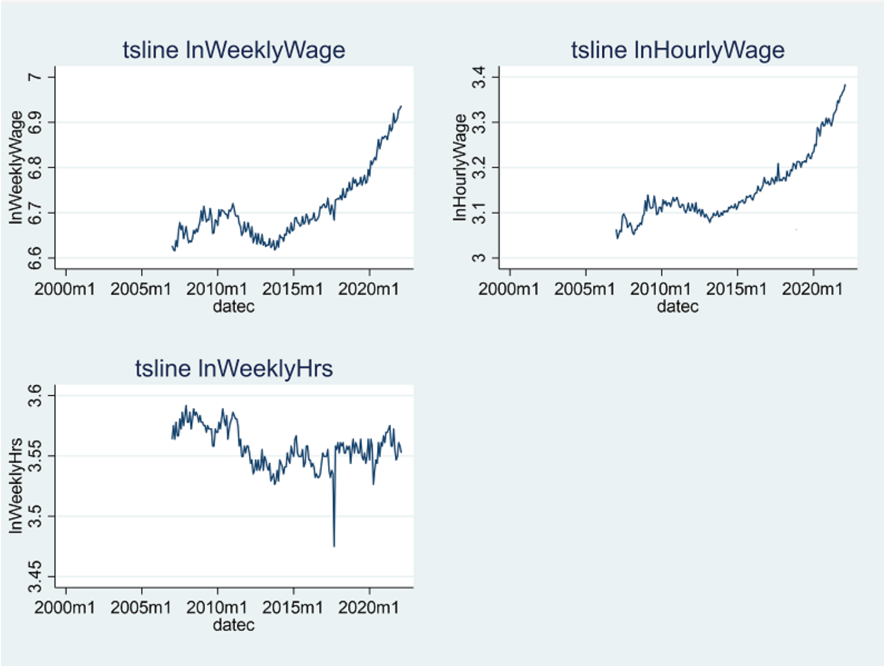


Graphs 3: Partial Autocorrelogram Graphs for AllEmployees Predictors

In conclusion, these insights indicate these three variables are nonstationary & not weakly dependent; hence the variables needed to be first differenced for modeling.

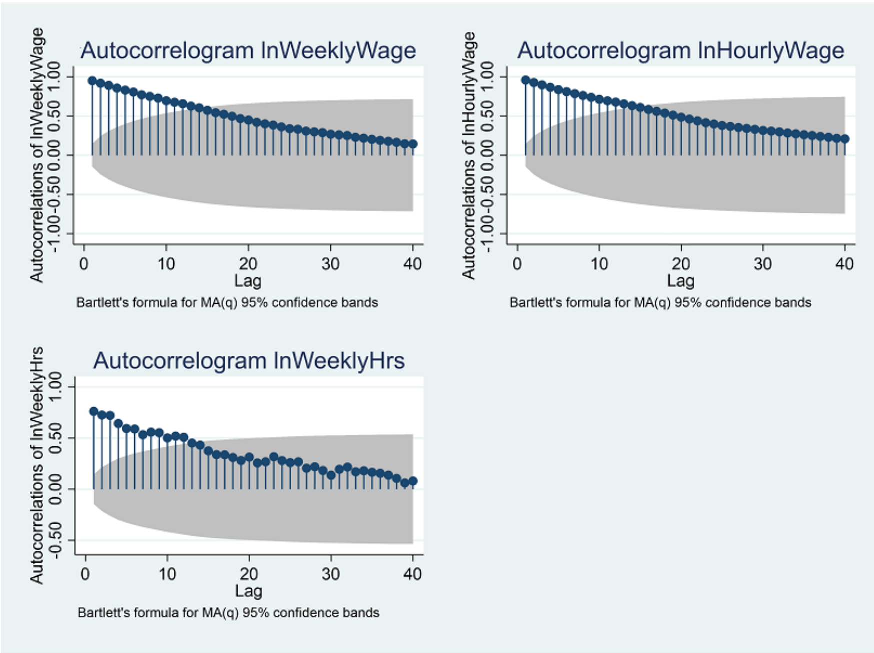
2) Summary statistics: Average Weekly Earnings of All Employees: Total Private in Miami-Fort Lauderdale-West Palm Beach, FL (MSA).

After testing with some models, the best variables for predicting Average Weekly Wage are the Average Weekly Wage, Average Hourly Wage, and Average Weekly Hours. Then, I generate some tsline graphs, autocorrelograms, and partial autocorrelograms to get a deeper insight into the data. First, start with tsline charts show the Average Weekly Wage and Average Hourly Wage have a steady nonstationary growth over time for most of the time. The Average Weekly Hours have a nonsteady nonstationary decrease over time.

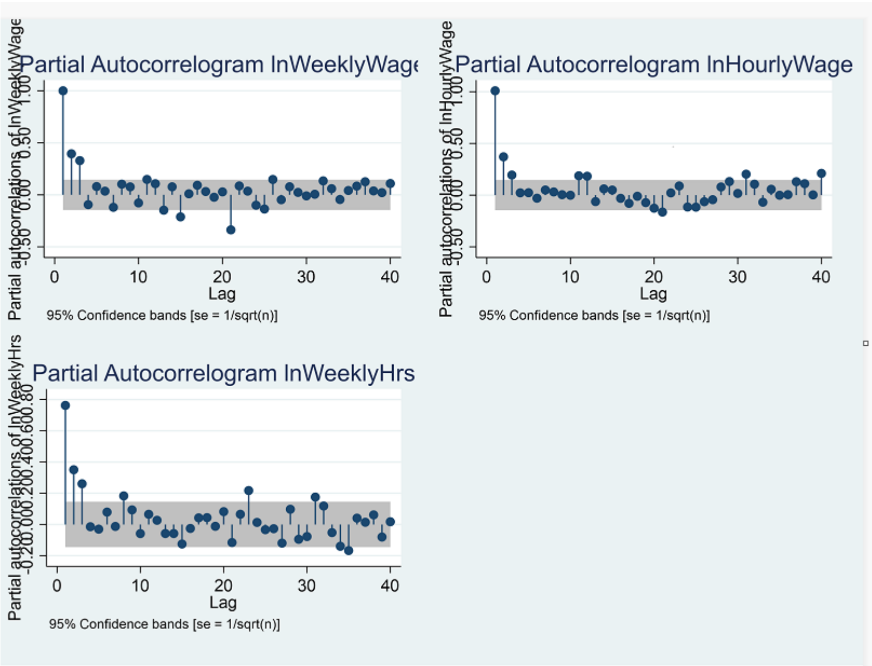


Graphs 4: Tsline Graphs for AverageWeeklyWage Predictors

As the autocorrelograms showed for lnWeeklyWage, lnHourlyWage, and lnWeeklyHrs are not helpful because they have too many lags outside of the confidence region on their plots. Thus, I checked the partial autocorrelograms show fewer lags than the autocorrelograms for these variables, assisting me to determine whether using AR (1) model would be appropriate for these variables.



Graphs 5: Autocorrelogram Graphs for AverageWeeklyWage Predictors



Graphs 6: Partial Autocorrelgram Graphs for AverageWeeklyWage Predictors

In conclusion, these insights indicate these three variables are nonstationary & not weakly dependent; hence the variables needed to be first differenced for modeling.

3. Model estimation and selection

1) Vselect models for forecasting All Employees: Total Private in Miami-Fort Lauderdale-West Palm Beach, FL (MSA)

After testing different models with different predictors, the best variables to use with forecasting All Employees in Miami-Fort Lauderdale-West Palm Beach, FL (MSA) are first differenced & logged All Employees, Labor Force, and Unemployment ratio in Miami-Fort Lauderdale-West Palm Beach, FL (MSA). After running the vselect to estimate and evaluate the models with 12-Lags for each predictor variable, the results are shown below in Table 1:

# Preds	R2ADJ	C	AIC	AICC	BIC
1	.318734	55.40285	-1477.623	-1475.858	-1431.689
2	.3474025	44.39182	-1487.556	-1485.531	-1438.089
3	.402965	22.42448	-1509.13	-1506.825	-1456.129
4	.4136497	19.01953	-1512.764	-1510.16	-1456.23
5	.4199669	17.43182	-1514.575	-1511.651	-1454.507
6	.4264835	15.78493	-1516.507	-1513.246	-1452.906
7	.43111	14.92449	-1517.635	-1514.015	-1450.501
8	.4366408	13.71612	-1519.191	-1515.191	-1448.523
9	.4393245	13.6677	-1519.487	-1515.087	-1445.286
10	.4431539	13.17147	-1520.314	-1515.493	-1442.579
11	.4450533	13.45513	-1520.276	-1515.013	-1439.008
12	.4465508	13.90431	-1520.062	-1514.335	-1435.261
13	.4476672	14.50879	-1519.68	-1513.468	-1431.345
14	.449633	14.78405	-1519.694	-1512.974	-1427.826
15	.4504897	15.49907	-1519.205	-1511.955	-1423.804
16	.4520445	15.94572	-1519.044	-1511.241	-1420.109
17	.4520368	17.00228	-1518.167	-1509.789	-1415.699
18	.4525715	17.84987	-1517.546	-1508.569	-1411.545
19	.453195	18.66531	-1516.972	-1507.372	-1407.437
20	.4539004	19.45159	-1516.441	-1506.194	-1403.372
21	.4545461	20.26301	-1515.887	-1504.97	-1399.286
22	.4532631	21.80623	-1514.446	-1502.833	-1394.31
23	.4515671	23.50048	-1512.82	-1500.487	-1389.151
24	.4503475	25.01033	-1511.421	-1498.342	-1384.219
25	.4485351	26.73717	-1509.757	-1495.906	-1379.021
26	.4463016	28.61406	-1507.908	-1493.26	-1373.639
27	.4443569	30.37651	-1506.201	-1490.729	-1368.399
28	.4423855	32.14207	-1504.49	-1488.167	-1363.154
29	.4399588	34.06756	-1502.582	-1485.382	-1357.713
30	.4373684	36.04428	-1500.61	-1482.505	-1352.208
31	.4347231	38.032	-1498.625	-1479.587	-1346.69
32	.4320584	40.01755	-1496.643	-1476.643	-1341.174
33	.4293448	42.01148	-1494.651	-1473.66	-1335.648
34	.4266012	44.00678	-1492.657	-1470.647	-1330.121
35	.4238303	46.00232	-1490.662	-1467.603	-1324.593
36	.4210262	48	-1488.665	-1464.527	-1319.062

Table 1: Average Weekly Wage Vselect result

Then, I selected models from #Preds 4-10 as the best models because model 4 has the best BIC score and model 10 have the best AIC score. Finally, I ran a regression on these models plus a simple 12-lag AR model to get a table summarization of LOOCV RMSE and other calculations. The result is shown below:

	df	AIC	BIC	LOOCV
Model 12-L~R	24	-907.22397	-831.96481	.01687162
Model 4	16	-957.46606	-907.29328	.02260972
Model 5	17	-960.81984	-907.51127	.02274266
Model 6	18	-960.94383	-904.49946	.02309741
Model 7	19	-964.39234	-904.81217	.02353163
Model 8	20	-965.6378	-902.92183	.02518318
Model 9	21	-958.97441	-893.12264	.02540773
Model 10	22	-961.11562	-892.12805	.02450911

Table 2: Best Models Fit Measure Summarization (All Employees)

The 12-lag AR simple model has the lowest LOOCV RMSE result, model 5 has the best BIC result, and model 8 has the best AIC result. In conclusion, Model 5 is the best fit other than the 12-lag AR model.

2) Rolling Window procedure for final model selection on best models; All Employees: Total Private in Miami-Fort Lauderdale-West Palm Beach, FL (MSA)

After picking the best fit models from the vselect process, then I use the rolling window technique on five different models (Model 12-Lag, and Model 5 to Model 8) based on the best AIC and BIC to make a final model selection for finding the best out of the best fit models with the best prediction window width. The best rolling window rmse result for these models are shown in Table 3 below:

Models	Best Window Width (By Month)	RMSE
Model 12-Lag:	180	.03774197
Model 5:	180	.04004232
Model 6:	180	.04157896
Model 7:	180	.03921726
Model 8:	180	.03858189

Table 3: Rolling Window Best Results(All Employees)

Then, comparing the rmse of each model, the 12-Lag AR only model (reg d.lnAllEmployees l(1/12)d.lnAllEmployees m2 m3 m4 m5 m6 m7 m8 m9 m10 m11 m12) has the best rmse. However, this model is a benchmark model, and I won't use it for forecasting in this study. Thus, the Model 8 (reg d.lnAllEmployees l(1, 2, 7)d.lnAllEmployees l(1, 4, 9)d.lnUnemRate l(5, 6)d.lnLabF m2 m3 m4 m5 m6 m7 m8 m9 m10 m11 m12) have the second-best rmse at the window width of 180 months. In conclusion, model 8 is the best model at the window width at 180 months (15 years).

3) Vselect models for forecasting Average Weekly Earnings of All Employees: Total Private in Miami-Fort Lauderdale-West Palm Beach, FL (MSA)

After testing many different models with different predictors, the best variables to use with forecasting Average Weekly Earnings in Miami-Fort Lauderdale-West Palm Beach, FL (MSA) are first differenced & logged as Average Weekly Earnings, Average Hourly Earnings, and Average Weekly Hours in Miami-Fort Lauderdale-West Palm Beach, FL (MSA). After running the vselect to estimate and evaluate the models with 12-Lags for each predictor variable, the results are shown below in Table 4:

# Preds	R2ADJ	C	AIC	AICC	BIC
1	.3536648	10.20136	-995.6167	-992.8894	-954.928
2	.2775691	29.14068	-975.8933	-972.756	-932.0747
3	.3433141	14.65922	-991.1124	-987.5334	-944.1639
4	.3559384	12.72661	-993.4939	-989.4409	-943.4155
5	.3904361	5.780666	-1001.906	-997.3457	-948.6974
6	.3973962	5.257595	-1002.962	-997.8613	-946.6238
7	.4040225	4.831754	-1003.954	-998.2778	-944.4854
8	.4117794	4.17009	-1005.298	-999.0123	-942.7
9	.4142749	4.715145	-1005.155	-998.223	-939.4267
10	.4134358	6.012603	-1004.058	-996.4446	-935.2006
11	.4114523	7.561366	-1002.641	-994.3081	-930.6538
12	.4107174	8.829029	-1001.592	-992.5012	-926.4745
13	.4103552	10.01291	-1000.658	-990.7705	-922.4103
14	.409835	11.23	-999.6865	-988.9631	-918.3091
15	.408593	12.60126	-998.5172	-986.9172	-914.0099
16	.4080593	13.817	-997.5591	-985.0411	-909.9219
17	.4089908	14.71946	-997.028	-983.5498	-906.261
18	.4085419	15.91625	-996.1112	-981.6295	-902.2143
19	.4062815	17.49155	-994.6868	-979.1574	-897.6599
20	.4041431	19.03456	-993.3083	-976.6861	-893.1515
21	.4021713	20.53665	-991.9881	-974.2269	-888.7014
22	.3993155	22.21409	-990.4299	-971.4826	-884.0134
23	.3965558	23.86329	-988.9118	-968.73	-879.3653
24	.3936142	25.54085	-987.3559	-965.8902	-874.6795
25	.3906392	27.21616	-985.8043	-963.0043	-869.998
26	.388109	28.7939	-984.3892	-960.2031	-865.453
27	.3893845	29.61223	-984.0368	-958.4118	-861.9708
28	.3801133	32.50165	-980.7952	-953.6771	-855.5992
29	.376081	34.34401	-979.0146	-950.3479	-850.6887
30	.3719672	36.18985	-977.2294	-946.9574	-845.7737
31	.3675216	38.0867	-975.3733	-943.4378	-840.7877
32	.3625563	40.06874	-973.3984	-939.7399	-835.6829
33	.3575421	42.04492	-971.4317	-935.989	-830.5862
34	.3523511	44.03889	-969.4401	-932.1508	-825.4647
35	.3472405	46.0022	-967.4913	-928.2913	-820.3861
36	.3418577	48	-965.4944	-924.3179	-815.2592

Table 4: Average Weekly Wage Vselect result

Then, I selected models from #Preds 1-8 as the best models because model 1 has the best BIC score and model 8 have the best AIC score. Finally, I ran a regression on these models plus a simple 12-lag AR model to get a table summarization of LOOCV RMSE and other calculations. The result is shown below:

FIT[9,4]

	df	AIC	BIC	LOOCV
Model 12-L~R	24	-988.20666	-913.08909	.01530392
Model 1	13	-963.15634	-922.39096	.01426116
Model 2	14	-982.07046	-938.16928	.01339526
Model 3	15	-997.41814	-950.38116	.01306212
Model 4	16	-999.34663	-949.17385	.01291191
Model 5	17	-1007.4593	-954.15073	.01263521
Model 6	18	-1007.8937	-951.44932	.01260431
Model 7	19	-1009.6906	-950.11048	.01261056
Model 8	20	-1009.9601	-947.24411	.01255532

Table 2: Best Models Fit Measure Summarization (Average Weekly Wage)

Model 8 has the lowest LOOCV RMSE result and AIC result. Model 5 has the best BIC result. In conclusion, Model 5 is the best fit model because it has the best information criterion scores (BIC) and is the simplest model.

4) Rolling Window procedure for final model selection on best models; Average Weekly Earnings of All Employees: Total Private in Miami-Fort Lauderdale-West Palm Beach, FL (MSA)

After picking the best fit models from the vselect process, then I use the rolling window technique on five different models (Model 12-Lag, and Model 5 to Model 8) based on the best AIC and BIC to make a final model selection for finding the best out of the best fit models with the best prediction window width. The best rolling window rmse result for these models are:

Models	Best Window Width (By Month)	RMSE
Model 12-Lag:	132	.01332579
Model 5:	72	.01285895
Model 6:	72	.01334716
Model 7:	72	.01295601
Model 8:	72	.01278424

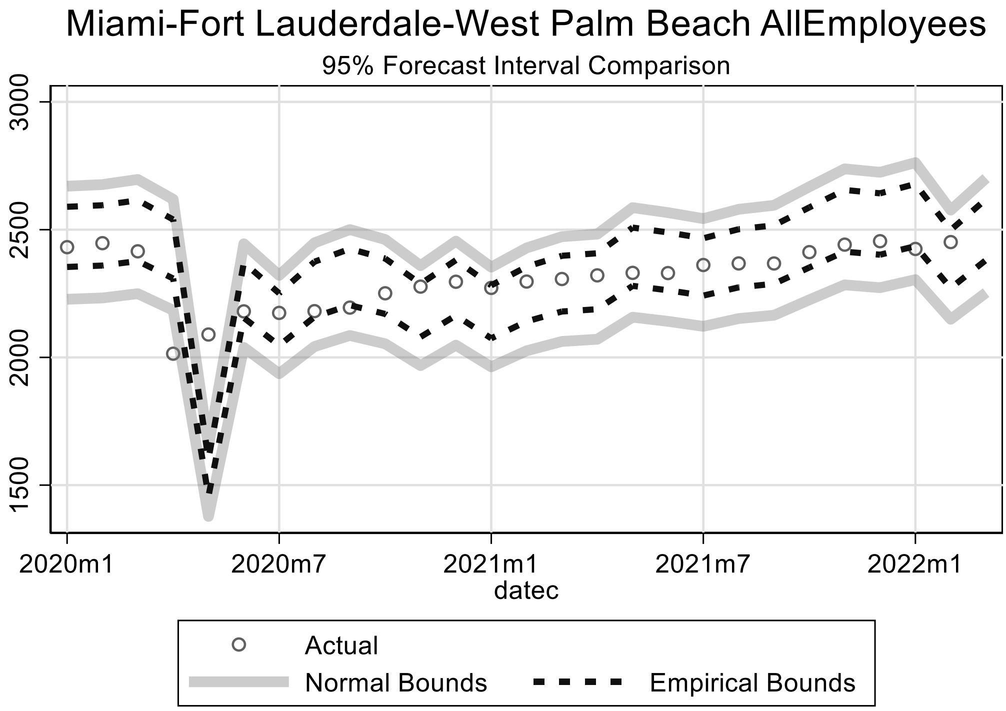
Table 6: Rolling Window Best Results (Average Weekly Wage)

Then, comparing the rmse of each model, the Model 8 (reg d.lnWeeklyWage l(1, 2, 5, 6)d.lnWeeklyWage l(9, 11)d.lnWeeklyHrs l(7, 10)d.lnHourlyWage m2 m3 m4 m5 m6 m7 m8 m9 m10 m11 m12) have the best rmse at the window width of 72 months (6 years).

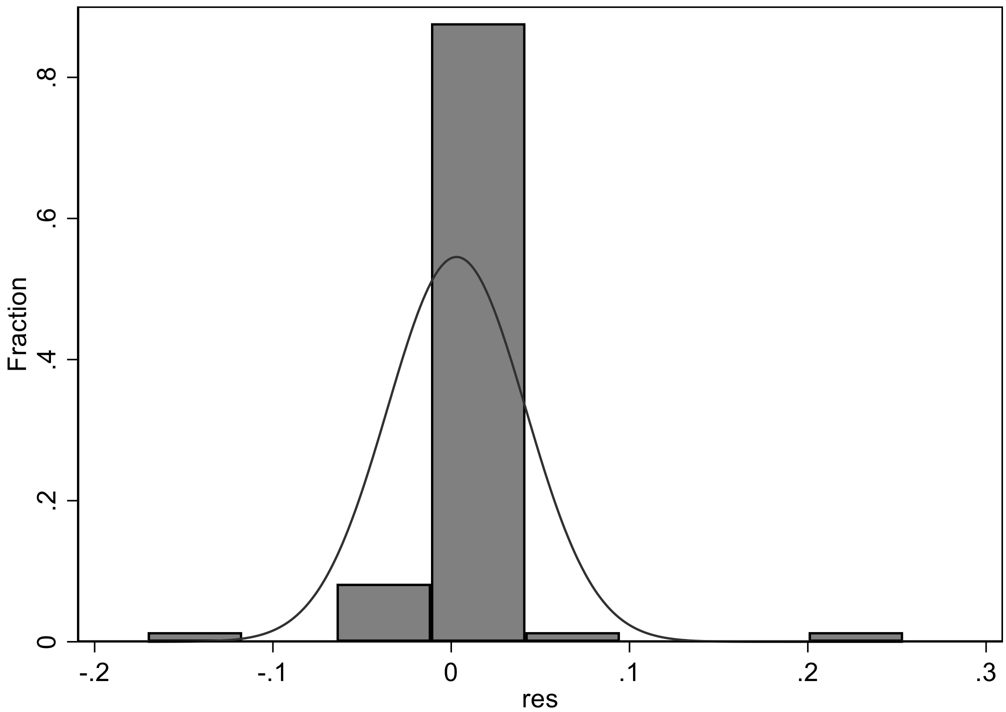
4. Final Results

1) Point and Interval Forecast for AllEmployees

The result from final model selection indicates Model 8 (reg d.lnAllEmployees l(1, 2, 7)d.lnAllEmployees l(1, 4, 9)d.lnUnemRate l(5, 6)d.lnLabF m2 m3 m4 m5 m6 m7 m8 m9 m10 m11 m12) with a window width of 180 months are the best model. I ran the model to generate a point forecast and expand the point forecast into a 95% interval forecast with both an assumed normal approach and an empirical approach. The prediction of assuming a normal forecast has an upper bound of 2701.423, lower bound of 2253.463, and point forecast of 2467.298. The empirical prediction forecast has an upper bound of 2619.767, lower bound of 2381.511, and point forecast of 2475.606. The result shown in the combined interval comparison graph below shows that the empirical approach has a narrower interval than the assumed normal approach. Graphing the distribution of the errors in a histogram shows an uneven spread/distribution of errors indicating the data is not normal. In summary, the empirical approach is better for forecasting the prediction.



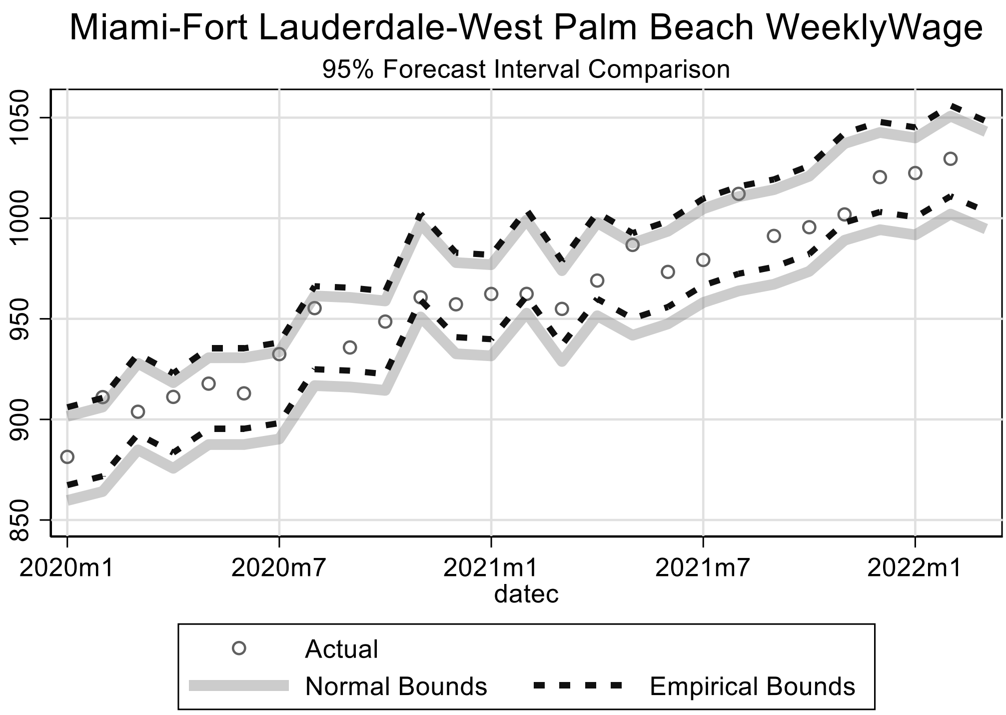
Graph 1: Interval Forecast for AllEmployees



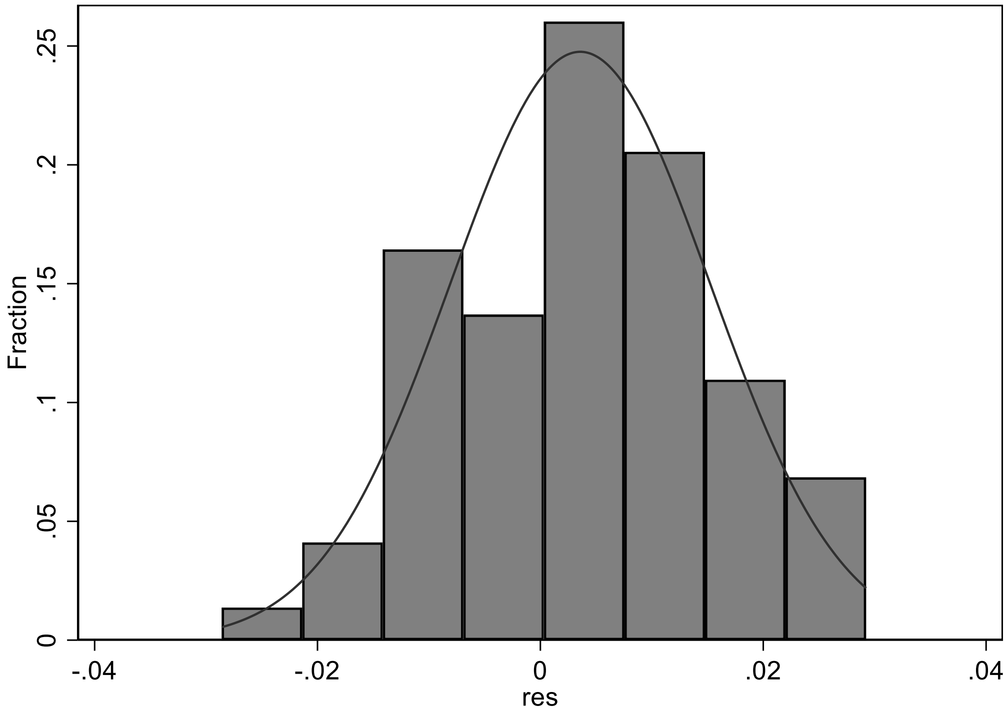
Histogram 1: Spread of error for AllEmployess

2) Point and Interval Forecast for WeeklyWage

The result from the final model selection indicates Model 8 (reg d.lnWeeklyWage l(1, 2, 5, 6)d.lnWeeklyWage l(9, 11)d.lnWeeklyHrs l(7, 10)d.lnHourlyWage m2 m3 m4 m5 m6 m7 m8 m9 m10 m11 m12) have the best rmse at the window width of 72 months (6 years) are the best model. I ran the model to generate a point forecast and expand the point forecast into a 95% interval forecast with both an assumed normal approach and an empirical approach. The prediction of the assumed normal forecast has an upper bound of 1043.01, lower bound of 994.6629, and point forecast of 1018.55. The empirical prediction forecast has an upper bound of 1048.248, a lower bound of 1003.478, and a point forecast of 1022.193. The result shown in the combined interval comparison graph below shows assumed normal approach show the upper bound and lower bound, being lower than the that the empirical approach's upper bound and lower bound. By graphing the distribution of the errors in a histogram offers an uneven spread/distribution of errors indicating the data being not normal. In summary, the empirical approach is better for forecasting the prediction.



Graph 2: Interval Forecast for WeeklyWage



Histogram 2: Spread of error for WeeklyWage

5. Conclusion

This study's time series and forecasting analysis prove that empirical approaches are more accurate than assumed normal approaches. The data used in AllEmployees and WeeklyWage forecasting both have an uneven distribution of errors, which shows data tend not to be normal. Thus, the one step ahead empirical forecast result for AllEmployees is 2619.767 for the upper bound, 2381.511 for the lower bound, and 2475.606 for the point forecast. The one step ahead empirical forecast result for WeeklyWage is 1048.248 for the upper bound, 1003.478 for the lower bound, and 1022.193 for the point forecast.

The study has no outstanding strong points, but it has many limitations and weaknesses. A limitation of the study is not using enough possible predictor variables to test with the predicting variables to see if there are better model(s) due to time constraints. Another limitation of the study is the not having enough observations for predicting the WeeklyWage because there is only observation starting from 2007, which is a bit lower than what I like to have. Finally, the biggest weakness of the study is my rookie skill in time series and forecasting analysis. Thus, I would like to improve and expand on my time series and forecasting analysis skills.