Intermediate Deliverable: Plots

Angel Sarmiento

3/18/2020

Introduction

There are a number of MSA's in Florida that have a pretty significant number of metrics used to describe and analyse them in a macroeconomical sense. MSA's are Metropolitan Statistical Areas, a designation given for areas with largely dense populations that have close economic ties to their surrounding areas. For this investigation, the Miami-West Palm-Fort Lauderdale MSA will be analysed with the final intent of delivering the best model possible for predicting the total nonfarm employment. The different models illustrated in this investigation all involve adjustments and reflections of the time series quality of the data. The best model will be chosen from a host of other generated models that will be compared with a number of comparison statistics. Namely: AIC, BIC, RMSE, and Out-of-sample RMSE.

The relevant predictors include the Economic Conditions Index (ECI): A value that shows the economic growth of an MSA, the average weekly hours worked, the average hourly earnings of employees, the average weekly earnings, all of the good produced, and the total weekly earnings of all employees. The economic conditions index was chosen because it makes intutive sense that the growth of an economy could contribute to employment in that economy. It also makes intuitive sense that there could be some positive relationship between the nonfarm employment and the amount of goods produced.

Analysis of Variables

	nonfarm	eci	avg_week_hrs	avg_hr_earnings
	Min. :1614	Min. :-19.110	Min. :32.44	Min. :20.98
	1st Qu.:1982	1st Qu.: 1.085	1st Qu.:34.62	1st Qu.:22.17
	Median :2218	Median : 2.305	Median :34.96	Median :22.64
	Mean :2195	Mean : 2.147	Mean :35.07	Mean :22.87
	3rd Qu.:2405	3rd Qu.: 3.345	3rd Qu.:35.57	3rd Qu.:23.61
	Max. :2761	Max. : 19.950	Max. :36.29	Max. :25.40
	NA	NA's :5	NA's :204	NA's :204

	avg_week_earnings	all_goods
	Min. :747.0	Min. :160.2
	1st Qu.:776.1	1st Qu.:211.1
	Median :797.7	Median :235.9
	Mean :801.8	Mean :226.6
	3rd Qu.:820.3	3rd Qu.:248.6
	Max. :894.1	Max. :275.8
	NA's :204	NA

Table 1. Summary statistics of all of the variables in the analysis

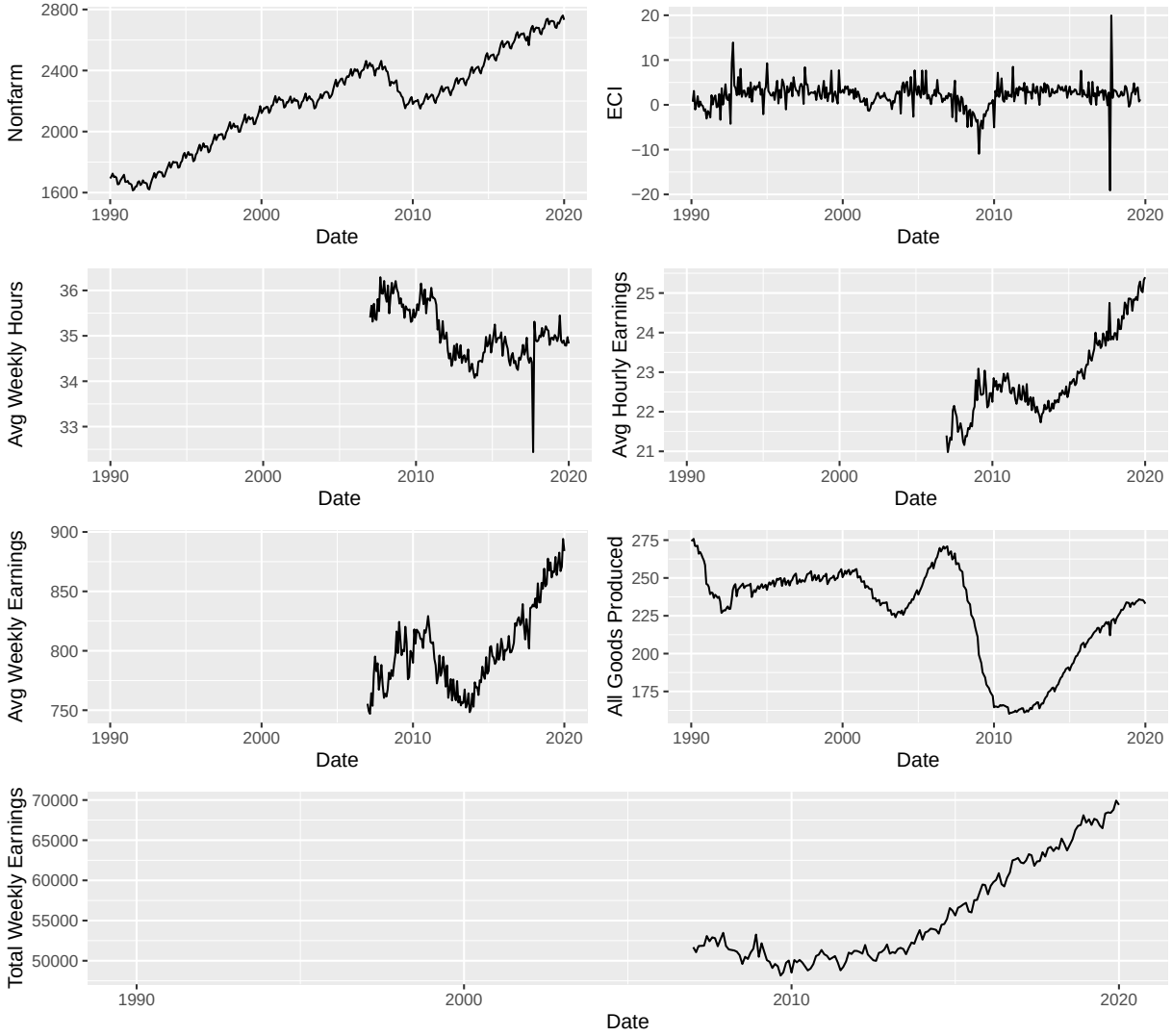

Figure 1. Time Series plots of every variable.

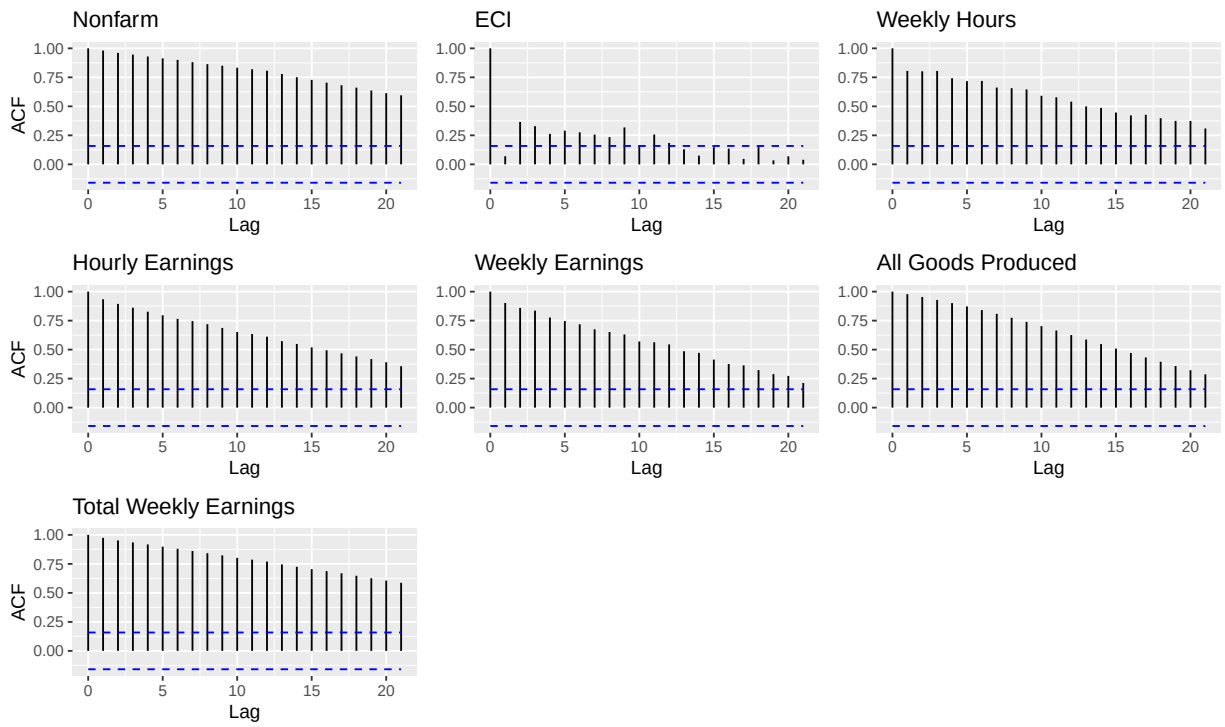

Figure 2. Autocorrelograms of all variables

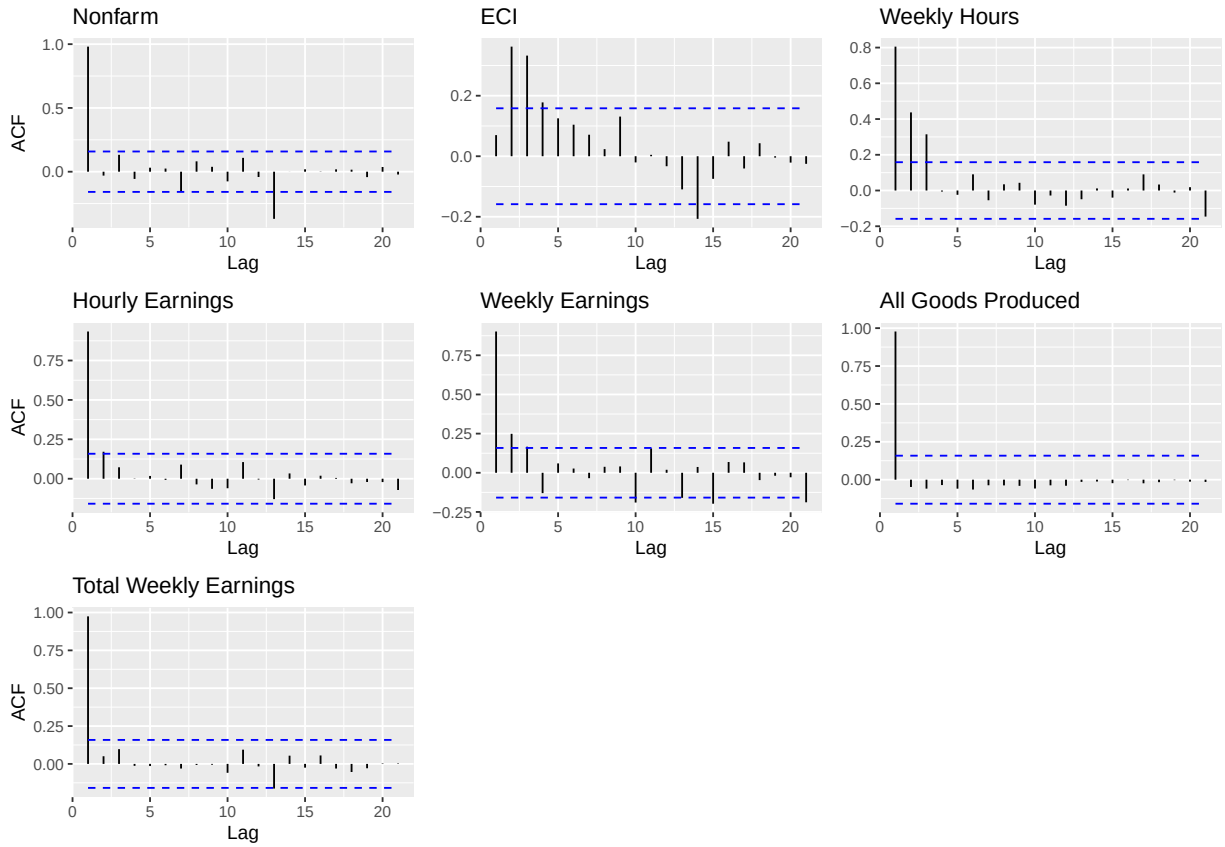

Figure 3. Partial Autocorrelograms of all variables

Appendix A: Code

```
## ----setup, include=FALSE-----------------------------------------------------------------------
knitr::opts_chunk$set(echo = FALSE, warning = FALSE, comments = NA, message = FALSE)


## ---- echo = FALSE, warning = FALSE, comments = NA---------------------------------------------
#library Import
library(caret)
library(tidyverse)

library(fable)
library(feasts)
library(fredr)
library(tsibble)
library(patchwork)
library(kableExtra)
library(ggfortify)

set.seed(23)
data <-  read_csv("dataset.csv")  %>% na.pass()

colnames(data)[2:7] <- c("nonfarm", "eci", "avg_week_hrs", "avg_hr_earnings",
                         "avg_week_earnings", "all_goods")
data["tot_week_earnings"] <- data[5]*data[2]


## ---- warning = FALSE, comments = NA, message = FALSE------------------------------------------
summary(data[2:5]) %>%
  kable(format = "latex") %>%
  row_spec(0,bold=TRUE) %>%
  kable_styling(full_width = FALSE, position = "center", latex_options = c("striped", "HOLD_position"))

summary(data[6:7]) %>%
  kable(format = "latex") %>%
  row_spec(0,bold=TRUE) %>%
  kable_styling(full_width = FALSE, position = "center", latex_options = c("striped", "HOLD_position"))


## -----------------------------------------------------------------------------------------------
dataset <- data %>% as_tsibble()
#creating the months variable
dataset <- dataset %>%
  mutate(Month = lubridate::month(DATE)) %>%
  index_by(Month)


## ---- fig.height = 8, fig.width = 9------------------------------------------------------------
t1 <- dataset %>% autoplot(nonfarm, ts.colour = "red") +
  xlab("Date") +
  ylab("Nonfarm")
t2 <- dataset %>% autoplot(eci) +
  xlab("Date") +
```

```
  ylab("ECI")
t3 <- dataset %>% autoplot(avg_week_hrs) +
  xlab("Date") +
  ylab("Avg Weekly Hours")
t4 <- dataset %>% autoplot(avg_hr_earnings) +
  xlab("Date") +
  ylab("Avg Hourly Earnings")
t5 <- dataset %>% autoplot(avg_week_earnings)+
  xlab("Date") +
  ylab("Avg Weekly Earnings")
t6 <- dataset %>% autoplot(all_goods)+
  xlab("Date") +
  ylab("All Goods Produced")
t7 <- dataset %>% autoplot(tot_week_earnings)+
  xlab("Date") +
  ylab("Total Weekly Earnings")

(t1 + t2) /
(t3 + t4) /
(t5 + t6) /
(t7)


## ---- fig.height = 6, fig.width = 10-------------------------------------------------------------

ac_data <- dataset %>% na.omit()

a1 <- autoplot(acf(ac_data$nonfarm, plot = FALSE), main = "Nonfarm")
a2 <- autoplot(acf(ac_data$eci, plot = FALSE), main = "ECI") + theme(axis.title.y = element_blank())
a3 <- autoplot(acf(ac_data$avg_week_hrs, plot = FALSE), main = "Weekly Hours")  + theme(axis.title.y = e
a4 <- autoplot(acf(ac_data$avg_hr_earnings, plot = FALSE), main = "Hourly Earnings")
a5 <- autoplot(acf(ac_data$avg_week_earnings, plot = FALSE), main = "Weekly Earnings")  + theme(axis.ti
a6 <- autoplot(acf(ac_data$all_goods, plot = FALSE), main = "All Goods Produced ")  + theme(axis.title.y
a7 <- autoplot(acf(ac_data$tot_week_earnings, plot = FALSE), main = "Total Weekly Earnings")


wrap_plots(a1, a2, a3 , a4, a5, a6, a7, ncol = 3, byrow = TRUE)


## ---- fig.height = 7, fig.width = 10-------------------------------------------------------------
p1 <- autoplot(pacf(ac_data$nonfarm, plot = FALSE), main = "Nonfarm")
p2 <- autoplot(pacf(ac_data$eci, plot = FALSE), main = "ECI") + theme(axis.title.y = element_blank())
p3 <- autoplot(pacf(ac_data$avg_week_hrs, plot = FALSE), main = "Weekly Hours")  + theme(axis.title.y =
p4 <- autoplot(pacf(ac_data$avg_hr_earnings, plot = FALSE), main = "Hourly Earnings")
p5 <- autoplot(pacf(ac_data$avg_week_earnings, plot = FALSE), main = "Weekly Earnings")  + theme(axis.ti
p6 <- autoplot(pacf(ac_data$all_goods, plot = FALSE), main = "All Goods Produced ")  + theme(axis.title
p7 <- autoplot(pacf(ac_data$tot_week_earnings, plot = FALSE), main = "Total Weekly Earnings")


wrap_plots(p1, p2, p3 , p4, p5, p6, p7, ncol = 3, byrow = TRUE)
```