Analysis of Unemployment and Production Index Time Series Data

Carl Elgin

Contents

1 Introduction

This paper will consist of first, a thorough univariate analysis of a time series, and second, a multivariate analysis involving a second time series. The analysis will include time series diagnostics, model selection, and forecasting, among other methods. The time series that will be used in the analysis are:

- *Unemployment* refers to the total number of unemployed U.S. males over the age of 20, in thousands.
- *production* refers to the U.S. production index.

Both time series have been downloaded from Data Market: https://datamarket.com/data/list/?q=

1.1 Univariate Analysis

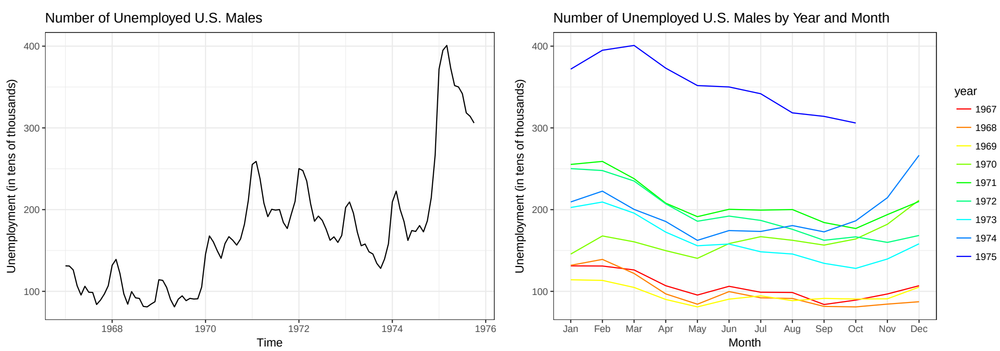

Figure 1: Plot of nnemployment in levels

Figure 1 shows a plot of the time series in linear time scale, as well as a plot of the seasonal pattern in the data. Based on these plots, it is clear that the series is non-stationary and has definite persistency. In order to progress towards stationarity, we will take the first differences of the series. A plot of the series and correlogram in first differences are shown in Figure 5 in the Appendix. Although the series in first differences looks much more stationary, a seasonal pattern is still present, as shown in both plots, as well as persistency. Figure 6 in the appendix shows a plot of the unemployment series after going into both first and seasonal differences as well as autocorrelation(ACF) and partial-autocorrelation(PACF)plots . The series plot appears to be stationary without persistency, and the ACF plot show significant correlations at lags 1 and 2, and the PACF plots show significant correlations at lags 1, 2, and 12.

Testing for a unit root using the augmented Dickey-Fuller test yields a significant p-value of0.012, implying that the series is stationary, which is a prerequisite for the time series models that will be used. and The Box-Ljung test gives a p-value of 0.0006, confirming that although the series is stationary, it is not a white noise. This indicates that by not resembling a random process, the series contains information to be modeled.

1.2 SARIMA Models

The next step in the analysis is to identify potential models and compare how they fit the data. As the series in levels exhibits strong seasonality, seasonal autoregressive moving average(SARIMA) models will be used to model the data. Based on the significant correlations at lags 1 and 2, shown in the ACF plot in Figure 6, it appears that the series follows a moving average process of order 1 or 2. Based on the additional significant correlation at lag 12 shown in the PACF plot in Figure 6, the model should also include a seasonal term.

Table 1: SARIMA model performance comparison.

			Information Criteria		MAE		Residual Correlation	
	Model	HOST	AIC	SIC	h=6	h=12	Chi-Sq	P-Val
1	(011)(011)[12]	(01)(01)	718.512	724.989	168.476	188.000	24.083	0.007
2	(012)(012)[12]	(02)(01)	702.741	713.536	129.570	134.219	2.968	0.982
3	(012)(011)[12]	(02)(01)	701.152	709.787	91.034	124.466	3.208	0.976
4	(112)(111)[12]	(02)(01)	703.232	716.185	91.197	123.707	1.403	0.999
5	(111)(111)[12]	(11)(01)	707.738	718.532	102.121	119.560	8.691	0.562

Note:
MAE is computed out of sample.
HOST: Highest order of significant terms in (pq)(PQ) form

1.3 Model Comparison

Table 1 presents a comparison of 4 different SARIMA models, including parameterization, goodness-of-fit, and residual diagnostics. Both Akaike Information Criterion(AIC) and Schwarz Information Criterion(SIC) are in-sample criteria, used to compare goodness-of-fit, while penalizing for model complexity. Additionally, the models will all be compared using out-of-sample criteria, namely the Mean Absolute Error(MAE) computed at forecast horizons 2 and 6, to asses the predictive performance of each model. Furthermore, the table provides the results of the Box-Ljung test, performed on the residual from each model. In this context, the Box-Ljung test assesses whether the residuals of the model resemble a white noise, indicating that the model is a good fit for the time series. As the null hypothesis of this test suggests that the residuals are a white noise, non-significant values indicate a good fit. Finally, the highest order of significant terms(HOST) is provided for each model, for the purpose of parsimony.

As evident in Table 1, Model 1 is the only model whose residuals are note a white noise, indicating a poor model fit. Models 3 and 4 yield the lowest MAE at forecast horizon 6, and Models 4 and 5 yield the lowest MAE at forecast horizon 12. Given that Models 2 and 3 contain the fewest non-significant terms, the analysis will continue with a comparison of these two models.

1.4 Forecasting Unemployment

Figure 8 in the appendix contains plotted forecasts from Models 2 and 3, including respective 90% and 95% prediction intervals, at forecast horizon 12(one year). Based on the compared model performance in Table 1, it is not surprising that the plotted forecasts look nearly identical. Both forecasts follow a similar profile, predicting a characteristic peak and decline, following the observed seasonal pattern.

Although Model 3 has a lower MAE for forecast horizons 6 and 12, as shown in Table 1, we can use the Diebold-Mariano test to identify whether there is truly any difference in prediction quality between the two models. Table 2 contains results of the Diebold-Mariano test of difference between the MAE of Models 2 and 3 at forecast horizons 6 and 12, and for loss functions of order 1 and 2. As shown in the table, non-significant p-values for each test conclude that the two models have effectively identical predictive performance.

Given the results of the Diebold-Mariano test, Model 3 will be chosen as the final model for parsimony. The equation for this model is formulated as:

$$(1-L)(1-L^{12})Y_t = c + (1-\theta_1 L - \theta_2 L^2)(1-\Theta_1 L^{12})u_t$$

With coefficients θ_1 = 0.3401061, θ_2 = 0.4470895, and Θ_1 = -0.6107956.

Table 2: Diebold Mariano test for differences in forcast accuracy between models 2 and 3.

Test Statistic	Forecast Horizon	Loss Function Power	P-Value
1.026	6	1	0.317
0.920	6	2	0.369
0.541	12	1	0.597
0.326	12	2	0.749

2 Multivariate Analysis

As shown in the plots in Figure 2, the *production* series is also non-stationary, with strong persistency and monthly pattern. Following the same procedure as for the *unemployment* series, we will go into first differences as well as seasonal differences in order to achieve stationarity.

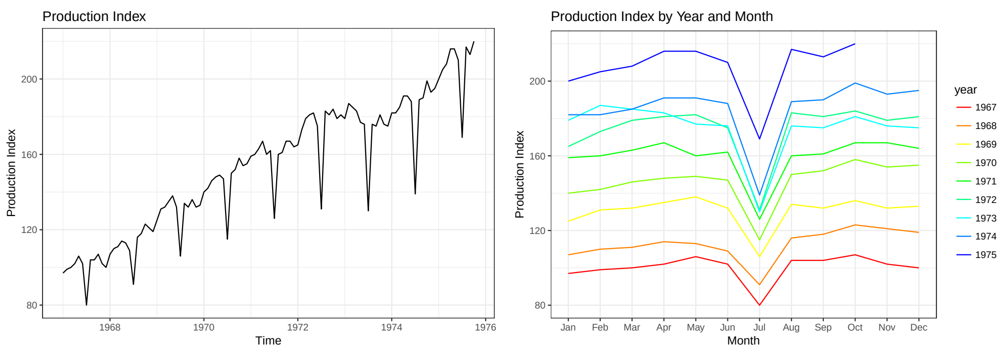

Figure 2: Plot of production index in Levels

Now twice-differenced, the *production* series appears stationary, as shown in Figure 9 in the appendix. Based on significant correlations at lags 1 and 12, as shown in the autocorrelation plots, it likely that a SARIMA model, as used for the *unemployment* series, will be an adequate fit again. Based on a p-value of 0.0001 for a unit root test on this series, we can conclude that the series is now stationary. Additionally, a Box-Ljung test yields a p-value of 0.102, suggesting that the series is not yet a white noise.

2.1 Test for Cointegration

As both *production* and *unemployment* are integrated of order 2, meaning that they are stationary in seasonally adjusted differences, we can test for the presence of a cointegration relation between the two series. Cointegration relations indicate a fundamental link between to variables.This relation takes the form of a linear combination of the variables yielding a stationary series, and represents the long-run relationship between the two variables. For the purpose of this analysis, the Johansen test for cointegration will be used to determine whether a cointegration equation exists, and if so, how many. Based on the SIC value, the *VARselect* function indicates that the optimal number of lags for the relationship is 2. Table 3 provides the results of the Johansen test procedure, using both the trace test and maximum eigenvalue statistics. Based on the results shown in the table, the procedure concludes that there is only one cointegrating equation present according to the 5% critical value.

The resulting cointegration equation is formulated as

$$151.4319532 + EMP_t - 2.087259PROD_t = \delta_t$$

, where δ_t is a stationary time series.

Table 3: Johansen Cointegration Test

	Trace Test			Maximum Eigenvalue		
	10pct	5pct	1pct	10pct	5pct	1pct
r <= 1 |	7.52	9.24	12.97	7.52	9.24	12.97
r = 0 |	17.85	19.96	24.60	13.75	15.67	20.20

Table 4: Granger Causality Test

Res.Df	Df	F	Pr(>F)
Production causes Unemployment			
86	NA	NA	NA
88	-2	4.488	0.014
Unemployment causes Production			
86	NA	NA	NA
88	-2	1.632	0.202

As the *production* and *unemployment* series are cointegrated, it is possible to describe the short-term dynamics of the relationship through the use of an error correction model(ECM). This model characterizes the way in which each series corrects towards the equilibrium relation, based on changes from the other. We can use the ECM to make additional forecasts, as shown in the plots in Figure 11 in the appendix.

2.2 Cross-correlation Plot and Granger Causality Test

The results from the Johansen cointegration test in the previous section indicate that there is a fundamental link between *unemployment* and *production*. For the purpose of further exploring this relationship, we can plot the cross-correlations of the two series, as well as formally test for Granger causality. The cross-correlation plot depicts the correlation between the stationary *unemployment* and *production* series at different lag and lead levels. As shown in the plot in Figure 10 in the appendix, there is a significant correlation at lag 2, which represents the correlation between $\Delta\Delta_{12}PROD_t$ and $\Delta\Delta_{12}EMP_{t+2}$.

Although causality cannot be formally determined, Granger causality refers to whether a process provides incremental predictive power for making predictions about another. Results of Granger causality tests for both directions is provided in Table 4. As the null hypothesis of this test represents no Granger causality, the p-values in Table 4 suggest that *production* Granger causes *unemployment*, but *unemployment* does not Granger cause *production*. In short, this test suggests that using lagged values of *production* will improve the forecasts for *unemployment* but not vice versa.

2.3 Vector Autoregressive Model

Based on the belief of a fundamental link between the *unemployment* and *production* series, we can improve the forecasting ability through the use of dynamic models. In the context of time series analysis, the term dynamic model refers to any model that includes lagged values of the variables under study. Using the previously selected lag of 2, we will continue with the analysis through an estimation of a vector autoregressive(VAR) model, referred to as VAR(2). Using ordinary least squares estimation, the VAR model will provide one predictive equation for each series, with each equation containing 2 lagged values of both series. Table 5 contains the results of the VAR(2) estimation, including estimated coefficients and p-values. As shown in the table, all of the coefficients of lagged values for the predictive *unemployment* equation are significant, apart from $\Delta\Delta_{12}PROD_{t-1}$. This indicates that knowing lagged valued of *production* improves the

Table 5: VAR model coefficients

	Unemployment				Production			
	Estimate	Std. Error	t value	Pr(>|t|)	Estimate	Std. Error	t value	Pr(>|t|)
ddemp.l1	0.2204899	0.0990273	2.2265558	0.0285903	-0.0008516	0.0400854	-0.0212442	0.9831001
ddprod.l1	0.5054051	0.2582356	1.9571472	0.0535739	-0.3018890	0.1045315	-2.8880176	0.0049032
ddemp.l2	0.3447377	0.0985758	3.4971839	0.0007465	0.0673519	0.0399026	1.6879070	0.0950527
ddprod.l2	0.6990380	0.2613754	2.6744596	0.0089577	-0.1842272	0.1058025	-1.7412363	0.0852182
const	0.3152218	1.1378376	0.2770358	0.7824174	0.0858407	0.4605869	0.1863724	0.8525917

forecasting ability for *unemployment*, as concluded in the previous section. Conversely, $\Delta\Delta_{12}PROD_{t-1}$ is the only significant coefficient in the predictive equation for *production*, which supports the previous conclusion of no Granger causality. Using the coefficients estimated from the VAR(2) model, we can simultaneously forecast both $\Delta\Delta_{12}EMP$ and $\Delta\Delta_{12}PROD$, using these equations:

Using the VAR(2) model estimated earlier, we can use the equations to make simultaneous forecasts of the stationary series $\Delta\Delta_{12}EMP$ and $\Delta\Delta_{12}PROD$, as shown in the plots in Figure 12 in the appendix.

Additionally, VAR models are characterized by their impulse response functions(IRF), which capture how an impulse originating at a specific time in one series proceeds through the system. Figure 3 contains plots that describe impulses from *unemployment*(left) and from *production*(right). As shown in the left plot, a unitary impulse in *unemployment* does not have a large effect on the *production* level at any time point. However, from the right plot, it is evident that a unitary impulse in *production* induces a large increase in *unemployment* that lasts until time $t + 3$.

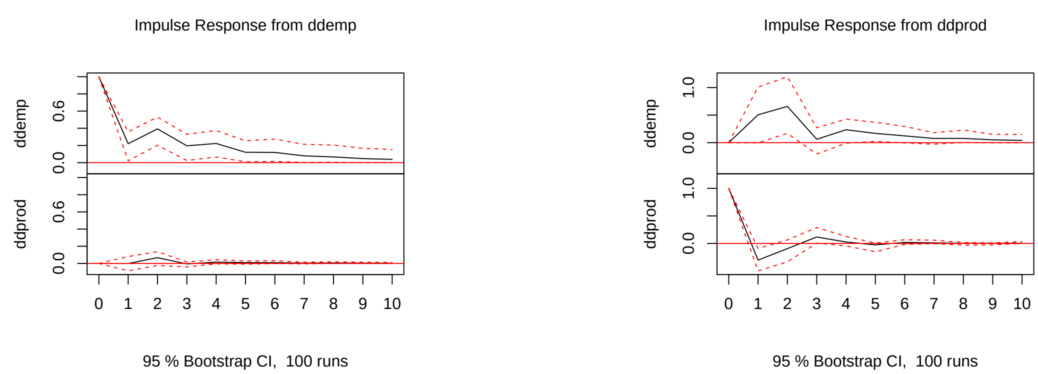

Figure 3: Impulse response function plots

3 Conclusion

In conclusion, we have found that a SARIMA(012)(011) is adequate for modelling the *unemployment* data for the purposes of short term predictions. Additionally, the multivariate analysis concluded that *production* yields incremental predictability, that is, changes in *production* tend to have an effect on *unemployment* within a 1-3 month window.

4 Appendix

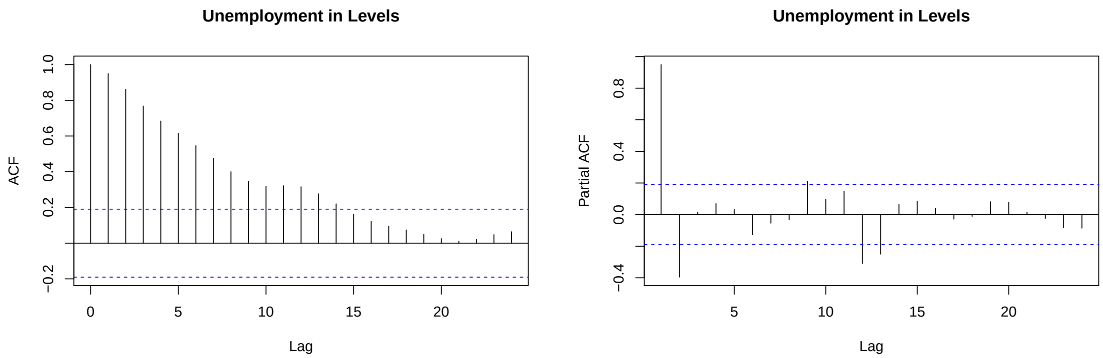

Figure 4: Autocorrelation plots for unemployment in levels.

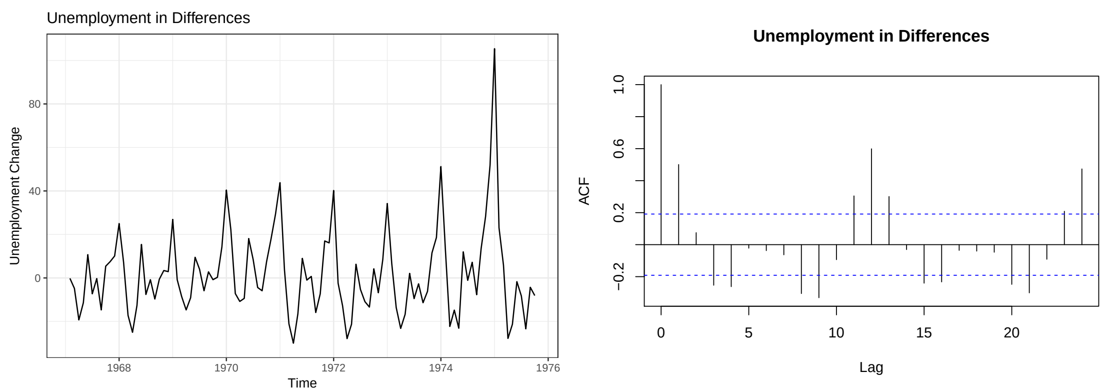

Figure 5: Series plot and correlogram for unemployment in differences.

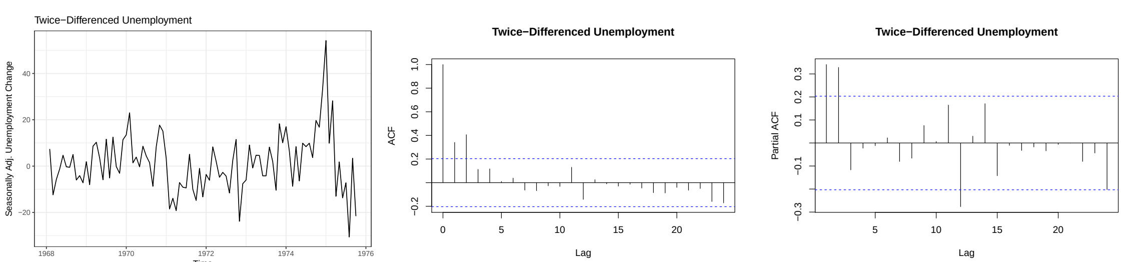

Figure 6: Seasonally adjusted unemployment series in differences.

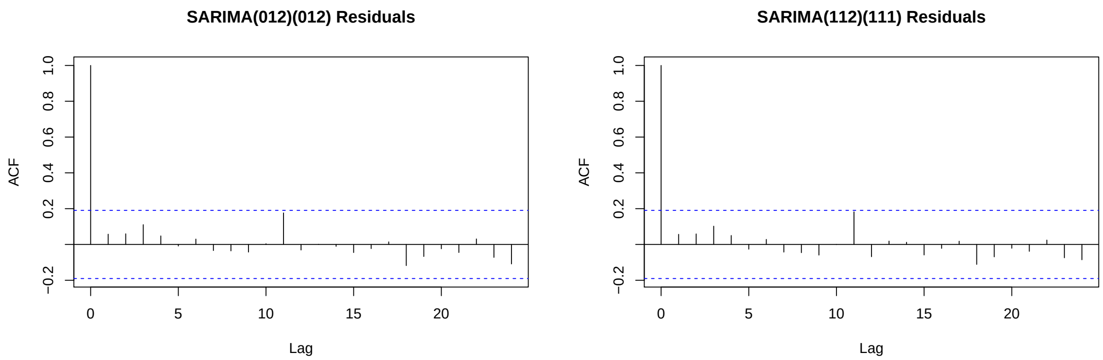

Figure 7: Residual autocorrelation plots of the two SARIMA forecasted models

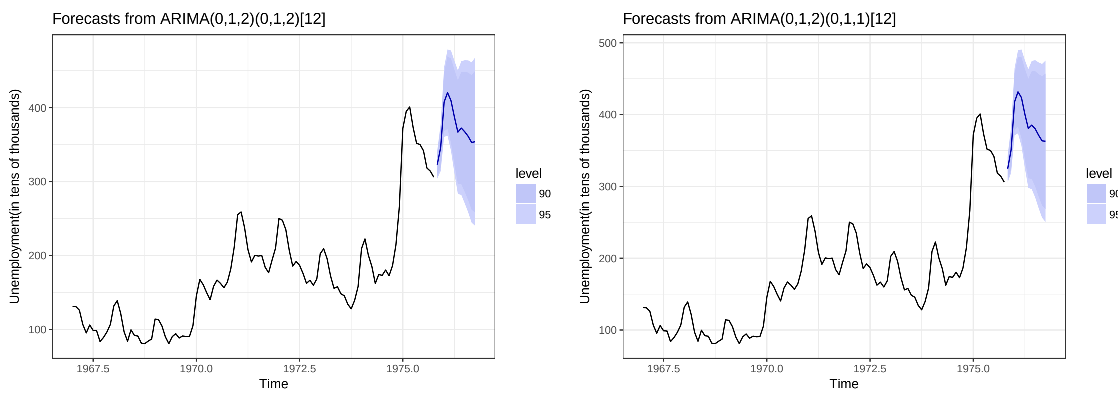

Figure 8: Comparison of forecasted unemployment using two different SARIMA models

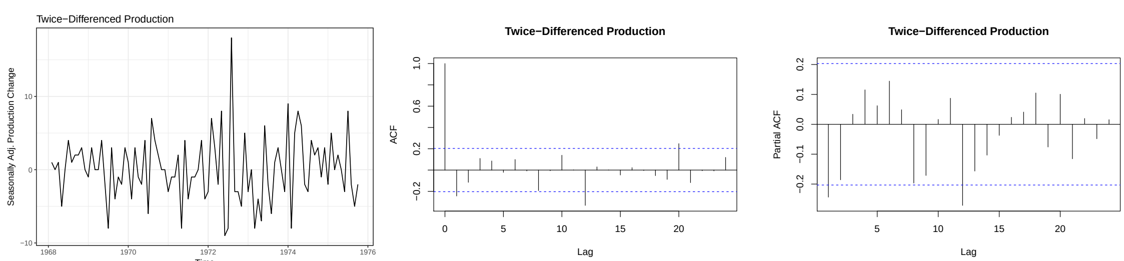

Figure 9: Seasonally adjusted production series in differences.

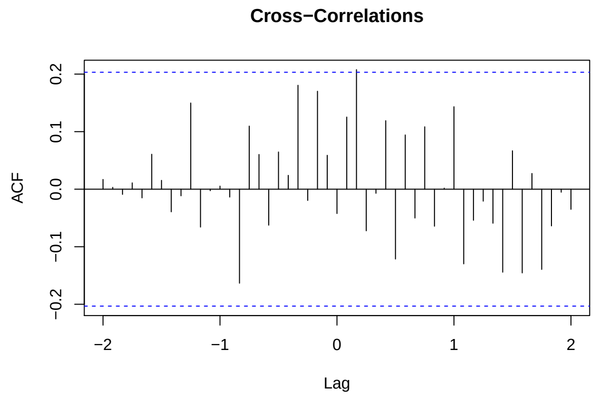

Figure 10: Plot of cross-correlations at respective lags and leads.

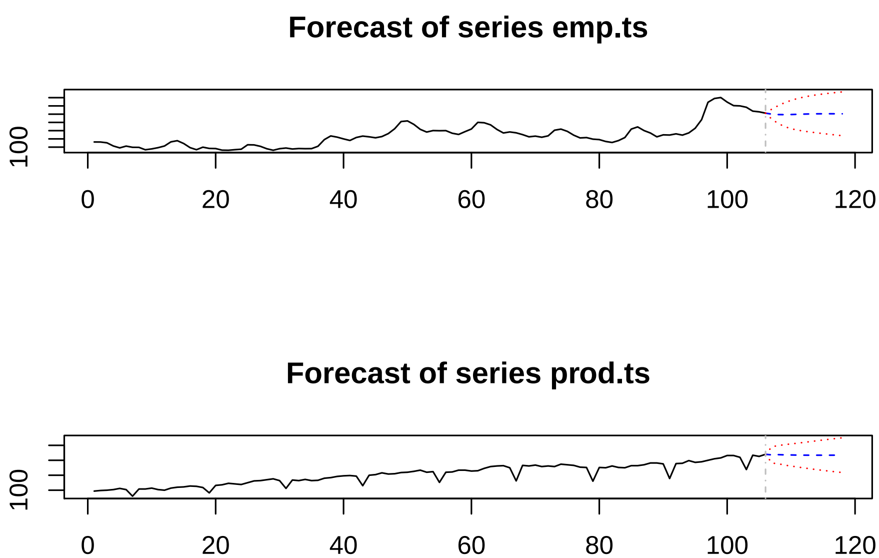

Figure 11: Forecasts for unemployment and production using ECM.

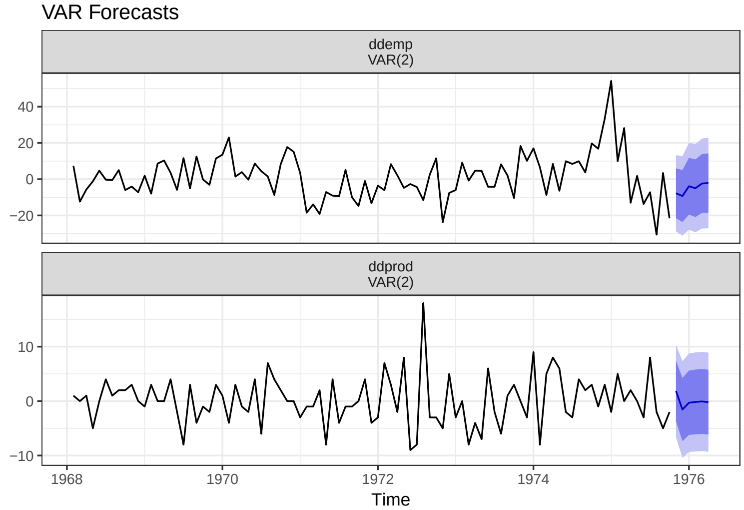

Figure 12: VAR(2) forecast plots

Part I
R Code

```
```{r, include=FALSE}
knitr::opts_chunk$set(echo = FALSE, message=FALSE, warning=FALSE)
library(ggplot2)
library(CADFtest)
library(forecast)
library(vars)
library(kableExtra)
library(magrittr)
library(dplyr)
library(tidyverse)
library(urca)
library(gridExtra)
```
```{r setup, message=FALSE, warning=FALSE}
##Project
prod<-read.csv("prod.csv",header=T)
prod<-prod[-107,]
prod.ts <- ts(prod$General.index.of.industrial.production..monthly
, frequency = 12,start = c(1967,1))
emp.ts <- ts((prod$Monthly.U.S..male..20.years.and.over..unemployment
.figures..10..3..1948.1981/10)
, frequency = 12,start = c(1967,1))
```
```{r level-plot, fig.cap="Plot of nnemployment in levels", out.width="40%"
,fig.show="hold",fig.align='center'}
emp.ts %>% autoplot +labs(x="Time",y="Unemployment (in tens of thousands)"
,title="Number of Unemployed U.S. Males")+theme_bw()
emp.ts %>% ggseasonplot(col=rainbow(12))+labs(y="Unemployment (in tens of thousands)"
,title="Number of Unemployed U.S. Males by Year and Month")+theme_bw()
```
```{r emp-unit_root, message=FALSE, warning=FALSE,error=FALSE}
emp.ts.d.s<- emp.ts %>% diff(lag=12) %>% diff
###Unit Root Test
##NULL:Series is not stationary
max.lag<- emp.ts.d.s %>% length %>% sqrt %>% round
##With drift as no trend is apparent in differences
emp.ts.d.s.urt<- emp.ts.d.s %>% CADFtest( type= "drift"
, criterion= "BIC", max.lag.y=max.lag) %>% summary
###Box-Ljung Test
##NULL:All x autocorrelations are zero, implying a white noise.
emp.ts.d.s.box<-Box.test(emp.ts.d.s, lag = max.lag, type = "Ljung-Box")
```
```{r arima-compare_function, message=FALSE, warning=FALSE,error=FALSE}
arima_compare<-function(model){
  aic <- model$aic
  bic <- model %>% AIC(k = log(64))
  boxt<- model$residuals %>% Box.test( lag = max.lag, type = "Ljung-Box")
  boxchi <- boxt$statistic
  boxp <- boxt$p.value
  list(aic=aic,bic=bic,boxchi=boxchi,boxp=boxp)
}
```
```{r, message=FALSE, warning=FALSE,error=FALSE}
get_pred_error<-function(ts,h,ar_ord,ar_sea){
y<-ts
```
```

```
S=round(0.75*length(y))
error1<-c()
for (i in S:(length(y)-h))
{
  i
mymodel.sub<-arima(y[1:i], order = ar_ord,seasonal=ar_sea)
predict1<-predict(mymodel.sub,n.ahead=h)$pred[h]
error1<-c(error1,y[i+h]-predict1)
}
return(error1)
}
```
```{r generate-sarima_models, message=FALSE, warning=FALSE,error=FALSE}
sar1<- emp.ts%>% arima(order=c(0,1,1), seasonal = list(order = c(0,1,1)))
sar2<- emp.ts%>% arima(order=c(0,1,2), seasonal = list(order = c(0,1,2)))
sar3<- emp.ts%>% arima(order=c(0,1,2), seasonal = list(order = c(0,1,1)))
sar4<- emp.ts%>% arima(order=c(1,1,2), seasonal = list(order = c(1,1,1)))
sar5<- emp.ts%>% arima(order=c(1,1,1), seasonal = list(order = c(1,1,1)))
```
```{r sarima-comparison, tab.cap="SARIMA Model Comparison"
,fig.align="center",warning=FALSE,error=FALSE}
sar1c<- sar1 %>% arima_compare
sar2c<- sar2 %>% arima_compare
sar3c<- sar3 %>% arima_compare
sar4c<- sar4 %>% arima_compare
sar5c<- sar5 %>% arima_compare
tribble( ~" ",~ "Model", ~"HOST", ~ "AIC", ~ "SIC",~"h=6",~"h=12"
, ~ "Chi-Sq", ~ "P-Val",
        "1","(011)(011)[12]","(01)(01)", sar1c$aic, sar1c$bic
        , get_pred_error(emp.ts,6,c(0,1,1),c(0,1,1)) %>%abs%>%mean
        , get_pred_error(emp.ts,12,c(0,1,1),c(0,1,1))%>%abs%>%mean
        , sar1c$boxchi, sar1c$boxp,

        "2","(012)(012)[12]","(02)(01)", sar2c$aic, sar2c$bic
        , get_pred_error(emp.ts,6,c(0,1,2),c(0,1,2))%>%abs%>%mean
        , get_pred_error(emp.ts,12,c(0,1,2),c(0,1,2))%>%abs%>%
        mean, sar2c$boxchi, sar2c$boxp,

        "3","(012)(011)[12]","(02)(01)", sar3c$aic, sar3c$bic
        , get_pred_error(emp.ts,6,c(0,1,2),c(0,1,1))%>%abs%>%mean
        , get_pred_error(emp.ts,12,c(0,1,2),c(0,1,1))%>%abs%>%
        mean, sar3c$boxchi, sar3c$boxp,

        "4","(112)(111)[12]","(02)(01)", sar4c$aic, sar4c$bic
        , get_pred_error(emp.ts,6,c(1,1,2),c(1,1,1))%>%abs%>%mean
        , get_pred_error(emp.ts,12,c(1,1,2),c(1,1,1))%>%abs%>%
        mean, sar4c$boxchi, sar4c$boxp,

        "5","(111)(111)[12]","(11)(01)", sar5c$aic, sar5c$bic
        , get_pred_error(emp.ts,6,c(1,1,1),c(1,1,1))%>%abs%>%mean
        , get_pred_error(emp.ts,12,c(1,1,1),c(1,1,1))%>%abs%>%
        mean, sar5c$boxchi, sar5c$boxp)%>%
    mutate_if(is.numeric, round, 3) %>%
  kable( booktabs=TRUE,caption="SARIMA model performance comparison.") %>%
  kable_styling(bootstrap_options="condensed",full_width=F) %>%
  add_header_above(c(" "," "," ", "Information Criteria"=2
  ,"MAE"=2,"Residual Correlation" = 2)) %>%
   footnote(general = c("MAE is computed out of sample."
   ,"HOST: Highest order of significant terms in (pq)(PQ) form"))
```
##Model Comparison
```
```

```
```{r dm-test,tab.cap="Diebold Mariano test for differences in forcast
accuracy between models 2 and 3.",message=FALSE, warning=FALSE,error=FALSE}
emp_dm1<-dm.test(get_pred_error(emp.ts,6,c(0,1,2),c(0,1,2))
,get_pred_error(emp.ts,6,c(0,1,2),c(0,1,1)),h=6,power=1)%>%
unlist(use.names=FALSE) %>% as.numeric
emp_dm2<-dm.test(get_pred_error(emp.ts,6,c(0,1,2),c(0,1,2))
,get_pred_error(emp.ts,6,c(0,1,2),c(0,1,1)),h=6,power=2)%>%
unlist(use.names=FALSE)%>% as.numeric
emp_dm3<-dm.test(get_pred_error(emp.ts,12,c(0,1,2),c(0,1,2))
,get_pred_error(emp.ts,12,c(0,1,2),c(0,1,1)),h=12,power=1)%>%
unlist(use.names=FALSE)%>% as.numeric
emp_dm4<-dm.test(get_pred_error(emp.ts,12,c(0,1,2),c(0,1,2))
,get_pred_error(emp.ts,12,c(0,1,2),c(0,1,1)),h=12,power=2)%>%
unlist(use.names=FALSE)%>% as.numeric
tribble(~"Test Statistic",~"Forecast Horizon",~"Loss Function Power"
,~"P-Value",
          emp_dm1[1],emp_dm1[2],emp_dm1[3],emp_dm1[5],
          emp_dm2[1],emp_dm2[2],emp_dm2[3],emp_dm2[5],
          emp_dm3[1],emp_dm3[2],emp_dm3[3],emp_dm3[5],
          emp_dm4[1],emp_dm4[2],emp_dm4[3],emp_dm4[5]) %>%
  mutate_if(is.numeric, round, 3) %>%
  kable(booktabs=TRUE,caption="Diebold Mariano test for
  differences in forcast
  accuracy between models 2 and 3.")%>%
  kable_styling(bootstrap_options="condensed",full_width=F)
```
```{r prod-level-plot, fig.cap="Plot of production index in Levels"
, out.width="40%",fig.show="hold",fig.align='center'
, message=FALSE, warning=FALSE,error=FALSE}
prod.ts %>% autoplot +labs(x="Time",y="Production Index"
,title="Production Index")+theme_bw()
prod.ts %>% ggseasonplot(col=rainbow(12))+labs(y="Production Index"
,title="Production Index by Year and Month")+theme_bw()
```
```{r prod-unit_root, message=FALSE, warning=FALSE,error=FALSE}
prod.ts.d.s<- prod.ts %>% diff(lag=12) %>% diff
###Unit Root Test
##NULL:Series is not stationary
max.lag<- prod.ts.d.s %>% length %>% sqrt %>% round
##With drift as no trend is apparent in differences
prod.ts.d.s.urt<- prod.ts.d.s %>% CADFtest( type= "drift", criterion= "BIC"
, max.lag.y=max.lag) %>% summary
###Box-Ljung Test
##NULL:All x autocorrelations are zero, implying a white noise.
prod.ts.d.s.box<-Box.test(prod.ts.d.s, lag = max.lag, type = "Ljung-Box")
```
```{r cointegration-test, message=FALSE, warning=FALSE,error=FALSE}
trace_test<-ca.jo(cbind(emp.ts, prod.ts),type="trace",K=2,ecdet="const"
,spec="transitory")
eigen_test<-ca.jo(cbind(emp.ts, prod.ts),type="eigen",K=2,ecdet="const"
,spec="transitory")

```
```{r johansen-results,tab.cap="Johansen test for cointegration"
,fig.align="right",fig.show="hold", message=FALSE, warning=FALSE,error=FALSE}
cbind(trace_test@cval,eigen_test@cval) %>% kable(caption="Johansen Cointegration Test"
,booktabs=TRUE) %>% kable_styling(bootstrap_options="condensed",full_width=F) %>%
  add_header_above(c(" ", "Trace Test" = 3, "Maximum Eigenvalue"=3))
```
```{r granger-test, tab.cap="Test results for Granger causality."
,fig.show="hold",fig.align="left"}
```
```

```
granger1<-grangertest(prod.ts.d.s,emp.ts.d.s,order=2)
granger2<-grangertest(emp.ts.d.s,prod.ts.d.s,order=2)
rbind(granger1,granger2) %>%   mutate_if(is.numeric, round, 3) %>%
  kable(caption="Granger Causality Test",booktabs=TRUE)%>%
  kable_styling(bootstrap_options="condensed",full_width=F
  ,position="right",font_size=10) %>%
  group_rows("Production causes Unemployment", 1, 2) %>%
  group_rows("Unemployment causes Production", 3, 4)
```
```{r, message=FALSE, warning=FALSE, error=FALSE}
#VARselect(cbind(emp.ts.d.s,prod.ts.d.s))
ddemp<-emp.ts.d.s
ddprod<-prod.ts.d.s
fit_var1<-VAR(cbind(ddemp,ddprod),type="const",p=2)
res<-summary(fit_var1)
emp_var<-res$varresult$ddemp$coefficients[,1] %>% unname() %>% round(3)
prod_var<-res$varresult$ddprod$coefficients[,1] %>% unname() %>% round(3)
```
```{r var-coefficients,tab.cap="VAR(2) model coefficients for Unemployment
and Production"
,fig.align="center", message=FALSE, warning=FALSE,error=FALSE}
cbind(res$varresult$ddemp$coefficients,res$varresult$ddprod$coefficients) %>%
  kable(caption="VAR model coefficients",booktabs=TRUE) %>%
  kable_styling(bootstrap_options="condensed",full_width=F)  %>%
  add_header_above(c(" ", "Unemployment" = 4, "Production"=4))
```
```{r impulse-response-plots,fig.cap="Impulse response function plots"
, out.width="50%",fig.show="hold", message=FALSE, warning=FALSE,error=FALSE}
irf_var1<-irf(fit_var1,ortho=FALSE,boot=TRUE,impulse ="ddemp")
irf_var2<-irf(fit_var1,ortho=FALSE,boot=TRUE,impulse ="ddprod")
plot(irf_var1)
plot(irf_var2)
```
#Appendix
```{r acf-emp-levels,fig.cap="Autocorrelation plots for unemployment in levels."
, out.width="45%",fig.show="hold",fig.align='center'}
emp.ts %>% as.vector %>% acf(lag.max=24,main="Unemployment in Levels")
emp.ts %>% as.vector %>% pacf(lag.max=24,main="Unemployment in Levels")
```
```{r emp-first-differences,fig.cap="Series plot and correlogram for unemployment
in differences."
, out.width="45%",fig.show="hold",fig.align='center'}
emp.ts %>% diff %>% autoplot +labs(x="Time",y="Unemployment Change"
,title="Unemployment in Differences")+theme_bw()
emp.ts %>% diff %>% as.vector %>% acf(lag.max=24,main="Unemployment in Differences")
#emp.ts %>% as.vector %>% pacf(lag.max=24,main="Unemployment in Levels")
```
```{r emp-stationary,fig.cap="Seasonally adjusted unemployment series in differences."
, out.width="30%",fig.show="hold",fig.align='center'}
emp.ts.d.s %>% autoplot +labs(x="Time",y="Seasonally Adj. Unemployment Change"
,title="Twice-Differenced Unemployment")+theme_bw()
emp.ts.d.s %>% as.vector %>% acf(lag.max=24,main="Twice-Differenced Unemployment")
emp.ts.d.s %>% as.vector %>% pacf(lag.max=24,main="Twice-Differenced Unemployment")
```
```{r emp-SARIMA-residuals,fig.cap="Residual autocorrelation plots
of the two SARIMA forecasted models"
, out.width="45%",fig.show="hold",fig.align='center'}
sar2$residuals %>% as.vector %>% acf(lag.max=24,main="SARIMA(012)(012) Residuals")
sar3$residuals %>% as.vector %>% acf(lag.max=24,main="SARIMA(112)(111) Residuals")
```
```{r emp-forecasts,fig.cap="Comparison of forecasted unemployment
```

~~~~
using two different SARIMA models"
, out.width="45%",fig.show="hold",fig.align='center'}
forecast1<- arima(emp.ts,order=c(0,1,2), seasonal = list(order = c(0,1,2))) %>%
forecast(h=12,level=c(90,95))
forecast1 %>% autoplot+labs(x="Time",y="Unemployment(in tens of thousands)")+theme_bw()
forecast2<- arima(emp.ts,order=c(0,1,2), seasonal = list(order = c(0,1,1))) %>%
forecast(h=12,level=c(90,95))
forecast2 %>% autoplot+labs(x="Time",y="Unemployment(in tens of thousands)")+theme_bw()
```
```{r prod-stationary,fig.cap="Seasonally adjusted production series in differences."
, out.width="30%",fig.show="hold",fig.align='center'}
prod.ts.d.s %>% autoplot +labs(x="Time",y="Seasonally Adj. Production Change"
,title="Twice-Differenced Production")+theme_bw()
prod.ts.d.s %>% as.vector %>% acf(lag.max=24,main="Twice-Differenced Production")
prod.ts.d.s %>% as.vector %>% pacf(lag.max=24,main="Twice-Differenced Production")
```
```{r cross-correlation, fig.cap="Plot of cross-correlations at respective lags
and leads."
,fig.align="center",out.width="50%"}
ccf(emp.ts.d.s,prod.ts.d.s,main="Cross-Correlations",lag.max = 24)
```
```{r ecm-forecasts,fig.cap="Forecasts for unemployment and production using ECM."
, out.width="80%",fig.show="hold",fig.align='center'}
fit_var<-vec2var(trace_test,r=1)
predict(fit_var,n.ahead=12) %>% plot()
```
```{r var-forecasts,fig.cap="VAR(2) forecast plots", out.width="60%",fig.show="hold"
,fig.align='center', message=FALSE, warning=FALSE,error=FALSE}
#predict(fit_var1,n.ahead=6)
fit_var1.f<-forecast(fit_var1,h=6)
fit_var1.f %>% autoplot +labs(title="VAR Forecasts")+theme_bw()
```
~~~~