Yukai Tan 928598798； Songhao Li 925322960

1. OLS estimator for AR(1) is biased, but consistent.

Consider $x_{t+1} = \phi_1 x_t + \varepsilon_{t+1}$, $\varepsilon_{t+1} \sim N(0, \sigma^2)$

square loss $L = \sum_{t=1}^{T} (\hat{x}_t - x_t)^2$

$$= \sum_{t=1}^{T} (\hat{\phi}_1 x_{t-1} - x_t)^2$$

$$\frac{\partial L}{\partial \hat{\phi}_1} = 0 \Rightarrow \hat{\phi}_1 = \frac{\sum_{t=1}^{T} x_{t-1} x_t}{\sum_{t=1}^{T} x_{t-1}^2}$$

$$= \phi_1 + \sum_{t=1}^{T} \frac{x_{t-1}}{\sum_{t=1}^{T} x_{t-1}^2} \varepsilon_t$$

ε_t is independent on x_{t-1}, but dependent on $\sum_{t=1}^{T} x_{t-1}^2$.

Thus, $E(\hat{\phi}_1) \neq \phi_1$, so it is biased.

If under the assumption that $|\phi_1| < 1$, which means AR(1) is stationary, by ergodic thm, OLS estimator is consistent.

2. $r_t = 0.01 + 0.1 r_{t-2} + \varepsilon_t$, $\varepsilon_t \sim N(0, 0.02)$

(a) $E(r_t) = 0.01 + 0.1 E(r_{t-2})$

$\Rightarrow 0.9\mu = 0.01 \Rightarrow \mu = \frac{1}{90}$

$Var(r_t) = 0.01 Var(r_{t-2}) + Var(\varepsilon_t)$

$\Rightarrow 0.99 Var(r_t) = 0.02$

$\Rightarrow \quad Var(r_t) = \frac{2}{99}$

(b) $\gamma_1 = Cov(r_t, r_{t-1})$

$= Cov(0.01 + 0.1 r_{t-2} + \varepsilon_t, r_{t-1})$

$= 0$

$= \rho_1$

$$\gamma_2 = Cov(r_t, r_{t-2})$$

$$= Cov(0.01 + 0.1 r_{t-2} + \varepsilon_t, r_{t-2})$$

$$= 0.1\, Cov(r_{t-2}, r_{t-2}) = 0.1 \cdot \frac{2}{99}$$

$$\rho_2 = \gamma_2 / \gamma_0 = 0.1$$

(c) $E_{100}[r_{101}] = E_{100}[0.01 + 0.1 r_{99} + \varepsilon_{101}]$

$$= 0.01 + 0.1 \times 0.02 = 0.012$$

error $= r_{101} - E_{100}[r_{101}] = \varepsilon_{101}$

Std. deviation $= \sqrt{0.02} = 0.1414$

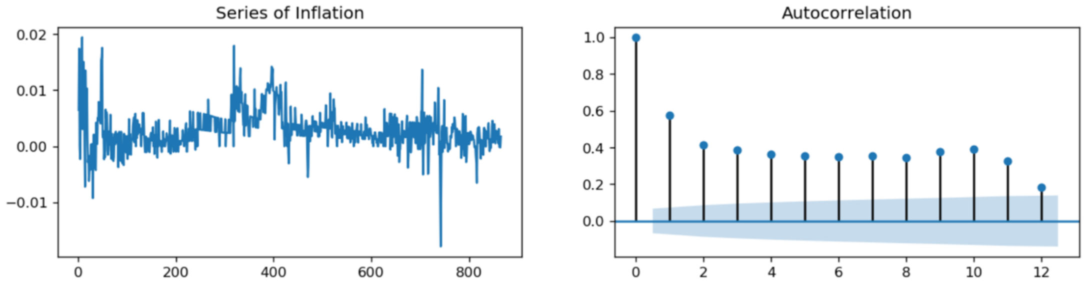
$E_{100}[r_{102}] = E_{100}[0.01 + 0.1 r_{100} + \varepsilon_{102}]$

$$= 0.01 + 0.1 \times 0.01 = 0.011$$

error $= r_{102} - E_{100}[r_{102}] = \varepsilon_{102}$

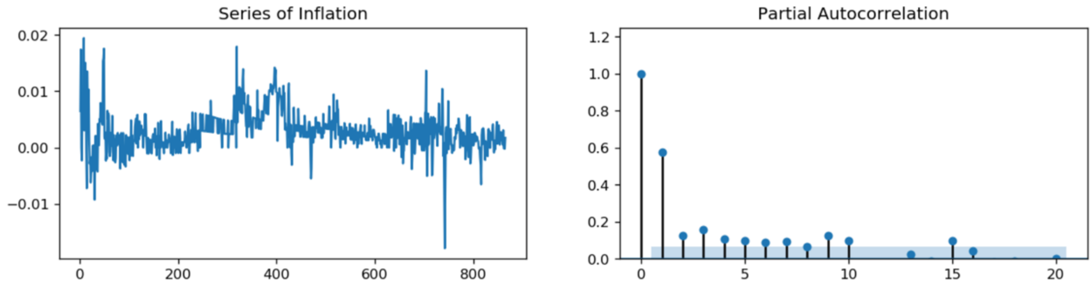
Std deviation $= \sqrt{0.02} = 0.1414$

3. (a)



(b)

Plot the PACF for the inflation series:



From the graph, we see that lag-1 is highly significant, so choose an AR(1) model.

```
                            AutoReg Model Results
==============================================================================
Dep. Variable:                    pai   No. Observations:                  865
Model:                     AutoReg(1)   Log Likelihood                3852.184
Method:               Conditional MLE   S.D. of innovations              0.003
Date:                Mon, 26 Apr 2021   AIC                            -11.748
Time:                        20:43:25   BIC                            -11.731
Sample:                             1   HQIC                           -11.742
                                  865
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
intercept      0.0012      0.000      9.723      0.000       0.001       0.001
pai.L1         0.5755      0.028     20.704      0.000       0.521       0.630
                                    Roots
=============================================================================
                  Real          Imaginary           Modulus         Frequency
-----------------------------------------------------------------------------
AR.1            1.7376           +0.0000j            1.7376            0.0000
-----------------------------------------------------------------------------
```

Using AIC:

```
                            AutoReg Model Results
==============================================================================
Dep. Variable:                    pai   No. Observations:                  865
Model:                    AutoReg(15)   Log Likelihood                3919.851
Method:               Conditional MLE   S.D. of innovations              0.002
Date:                Mon, 26 Apr 2021   AIC                            -12.021
Time:                        20:43:42   BIC                            -11.926
Sample:                            15   HQIC                           -11.985
                                  865
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
intercept      0.0004      0.000      3.191      0.001       0.000       0.001
pai.L1         0.4414      0.034     13.123      0.000       0.376       0.507
pai.L2         0.0763      0.036      2.098      0.036       0.005       0.148
pai.L3         0.0621      0.036      1.710      0.087      -0.009       0.133
pai.L4         0.0153      0.036      0.427      0.669      -0.055       0.085
pai.L5         0.0547      0.035      1.542      0.123      -0.015       0.124
pai.L6         0.0118      0.035      0.335      0.738      -0.057       0.081
pai.L7         0.0575      0.035      1.641      0.101      -0.011       0.126
pai.L8         0.0279      0.034      0.812      0.417      -0.039       0.095
pai.L9         0.0601      0.034      1.752      0.080      -0.007       0.127
pai.L10        0.0766      0.034      2.229      0.026       0.009       0.144
pai.L11        0.0702      0.034      2.039      0.041       0.003       0.138
pai.L12       -0.1683      0.034     -4.897      0.000      -0.236      -0.101
pai.L13       -0.0068      0.035     -0.197      0.844      -0.075       0.061
pai.L14       -0.0350      0.034     -1.025      0.305      -0.102       0.032
pai.L15        0.0948      0.031      3.010      0.003       0.033       0.156
```

Using BIC:

```
                            AutoReg Model Results
==============================================================================
Dep. Variable:                    pai   No. Observations:                  865
Model:                    AutoReg(12)   Log Likelihood                3907.258
Method:               Conditional MLE   S.D. of innovations              0.002
Date:                Mon, 26 Apr 2021   AIC                            -11.966
Time:                        20:55:49   BIC                            -11.888
Sample:                            12   HQIC                           -11.936
                                  865
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
intercept      0.0005      0.000      3.771      0.000       0.000       0.001
pai.L1         0.4287      0.034     12.742      0.000       0.363       0.495
pai.L2         0.0509      0.036      1.396      0.163      -0.021       0.122
pai.L3         0.0472      0.036      1.301      0.193      -0.024       0.118
pai.L4         0.0596      0.036      1.658      0.097      -0.011       0.130
pai.L5         0.0507      0.035      1.438      0.150      -0.018       0.120
pai.L6        -0.0126      0.035     -0.358      0.720      -0.082       0.056
pai.L7         0.1013      0.035      2.878      0.004       0.032       0.170
pai.L8         0.0230      0.035      0.651      0.515      -0.046       0.092
pai.L9         0.0749      0.035      2.132      0.033       0.006       0.144
pai.L10        0.1126      0.035      3.227      0.001       0.044       0.181
pai.L11        0.0484      0.035      1.403      0.161      -0.019       0.116
pai.L12       -0.1759      0.032     -5.519      0.000      -0.238      -0.113
```

Python ar_select_order function selects AR(15) under AIC, and selects AR(12) under BIC.

4.

(a) Economically, the first difference is essentially the difference of current season earnings compared with the past season, while the seasonal difference is compared with the same season last year.

(b) The model is built in codes. The estimated model is:

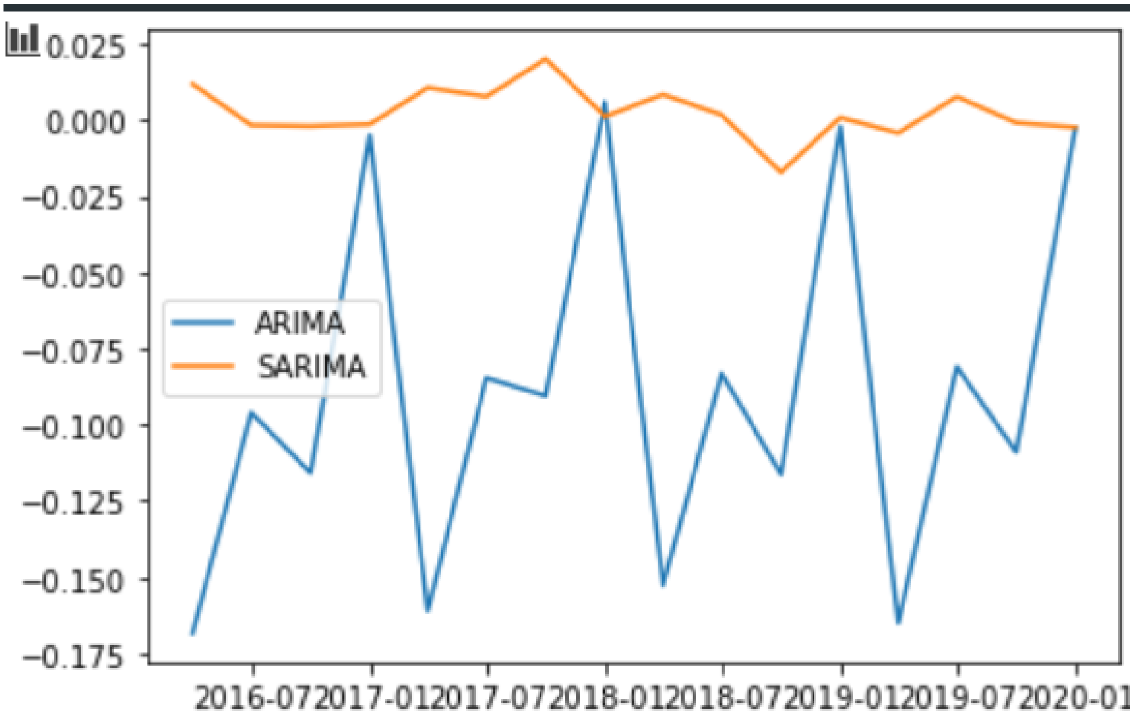
$$x_{t+1} - x_t = 0.0506 + \epsilon_{t+1} - 0.4920\epsilon_t$$

(c) The model is built in codes. The estimated model is:

$$(x_t - x_{t-1}) - (x_{t-4} - x_{t-5}) = (\epsilon_t - 0.1593\epsilon_{t-1}) - 0.4830(\epsilon_{t-4} - 0.1593\epsilon_{t-5})$$

In the model $\theta_1 = 0.1593, \theta_2 = 0.4830$, which means the past earning increase shocks (both the past season and the same season for the last year) is negatively correlated with predicted next season earnings.

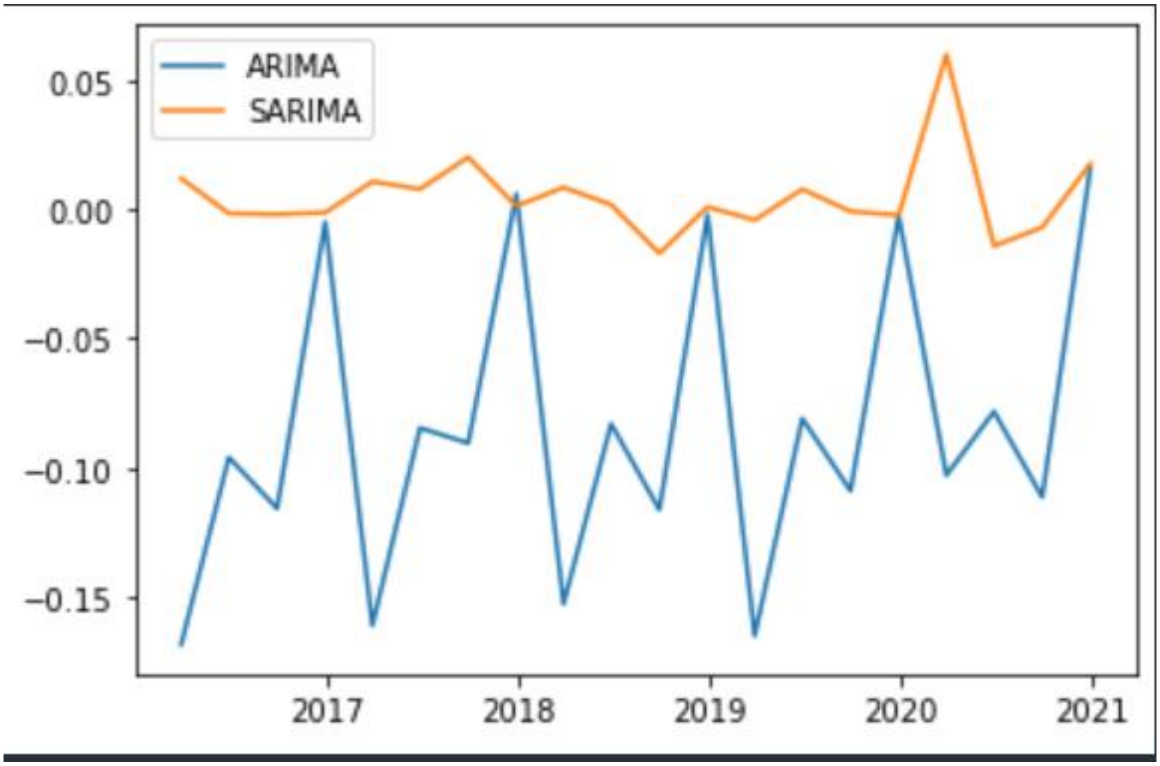
(d) (e)




```
ARIMA MSE: 0.01134832807889095
SARIMA MSE: 7.346810242072138e-05
```

The seasonal arima model (airline model) performs better.

(f)



Both models perform worse in 2020. This is because the COVID pandemic is an exogeneous event outside of the description of the model. To improve the forecast accuracy, we might include macroeconomic predictors such as expected inflation, GDP growth, treasury rate, etc.