Problem 1

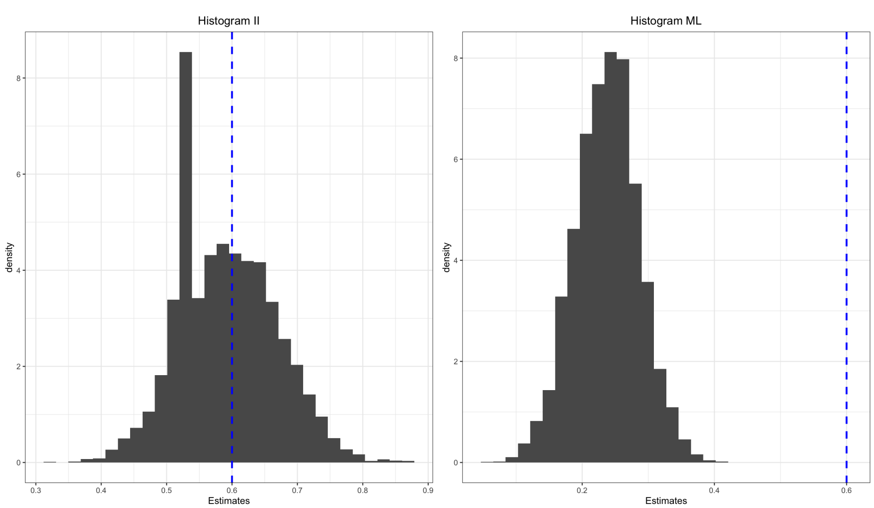
The results of the replication are shown in the following table.

Table 1: Paper replication ML and II

	ML		Indirect Inference	
	Bias	RMSE	Bias	RMSE
Replication	-0.3625	0.3658	-0.0093	0.0758
Original Paper	-0.3619	0.3650	-0.0291	0.0761

As you can see, the results for the ML bias are close to the results found by the authors. On the other hand, the results for the Indirect Inference are better than those the authors found. One possible explanation of the difference we found in the bias of the Indirect Inference estimation is that it seems the results are sensitive to which method of optimization we used.
I also generated a histogram based on the Monte Carlo simulation for both estimation methods.

Figure 1: Histogram Monte Carlo experiment



Problem 2

Part a)

We will first write the model in state space form.
Define the following:

- ΔY_{it} is the first difference of the log of each variable (industry production, personal income less transfer payments, total manufacturing and trade sales, employees on nonagricultural payrolls.
- $\Delta y_{it} = \Delta Y_{it} - \Delta \bar{Y}_i$: deviation from means for each of the variables defined above
- C_t the state variable we will try to recover from data. Similarly, ΔC_t is the first difference of C_t and $\Delta c_t = \Delta C_t - \Delta \bar{C}$

Now, let:

$$a_t = \begin{bmatrix} \Delta c_t \\ \Delta c_{t-1} \\ e_{1t} \\ e_{1,t-1} \\ e_{2t} \\ e_{2,t-1} \\ e_{3t} \\ e_{3,t-1} \\ e_{4t} \\ e_{4,t-1} \end{bmatrix}$$

$$Z = \begin{bmatrix} \gamma_1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \gamma_2 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \gamma_3 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \gamma_4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$$\Delta y_t = \begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \\ \Delta y_{3t} \\ \Delta y_{4t} \end{bmatrix}$$

$$T = \begin{bmatrix} \phi_1 & \phi_2 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \psi_{11} & \psi_{21} & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & \psi_{14} & \psi_{24} \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix}$$

$$v_t = \begin{bmatrix} w_t \\ 0 \\ \epsilon_{1t} \\ 0 \\ \epsilon_{2t} \\ 0 \\ \epsilon_{3t} \\ 0 \\ \epsilon_{4t} \\ 0 \end{bmatrix}$$

where $w_t \sim$ i.i.d $\mathcal{N}(0,1)$ and $\epsilon_{it} \sim$ i.i.d $\mathcal{N}\left(0, \sigma_i^2\right)$
The measurement equation will be given by:

$$\Delta y_t = Z a_t$$

The transition equation will be given by:

$$a_t = T a_{t-1} + v_t$$

The following table shows the estimates of the parameters from the original paper

Table 2: Parameters estimates

	Y_1	Y_2	Y_3	Y_4
γ_i	0.717	0.521	0.47	0.602
ψ_{1i}	-0.04	-0.087	-0.414	0.108
ψ_{2i}	-0.137	0.154	-0.206	0.448
$\sigma_i(\times 10^{-2})$	0.488	0.769	0.735	0.54
	$\phi_1 =$	0.545	$\phi_2 =$	0.032

We use the following equations to recover the state variable a_t, starting from a initial guess a_0^0 and P_0^0:

$$\begin{aligned}
a_t^{t-1} &= Tx_{t-1}^{t-1} \\
P_t^{t-1} &= TP_{t-1}^{t-1}T' + Q \\
K_t &= P_t^{t-1}Z' \left[ZP_t^{t-1}Z'\right]^{-1} \\
a_t^t &= a_t^{t-1} + K_t \left(y_t - Zx_t^{t-1}\right) \\
P_t^t &= \left[\mathrm{I} - K_t Z\right] P_t^{t-1}
\end{aligned}$$

In order to recover C_t, we follow the procedure described in the original paper, that is:

$$C_t^t = C_t^{t-1} + \Delta c_t^t + \delta$$

In order to estimate δ, we use the following:

$$\hat{\delta} = W(1)\Delta\bar{Y}$$

where $W(1)$ is the first row of the following matrix:

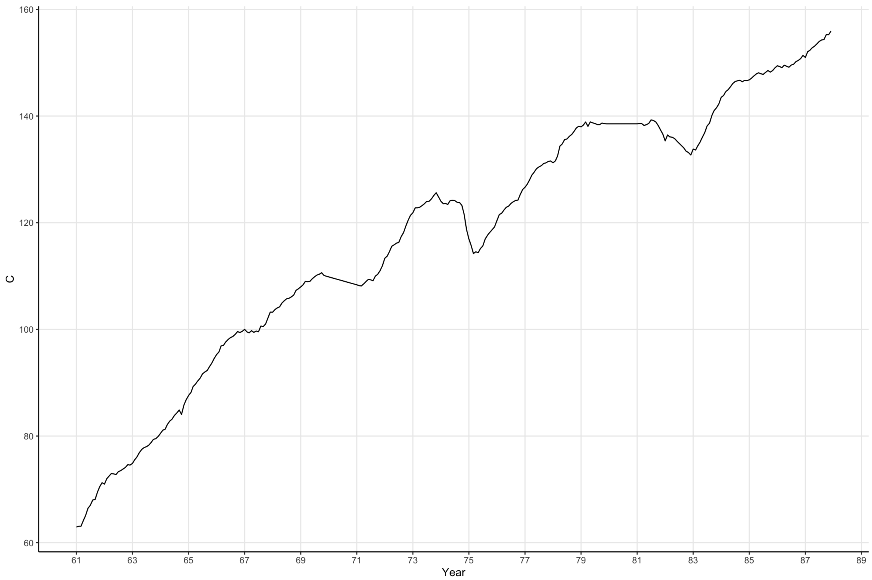
$$\left[\mathrm{I} - (\mathrm{I} - K_T Z)T)\right]^{-1} K_T$$

where K_T is the gain of the Kalman Filter obtained in the last step of the filtering process, and $\Delta\bar{Y}$ is given by:

$$\Delta\bar{Y} = \frac{1}{T}\sum_{t=1}^{T}\begin{bmatrix}\Delta Y_{1t}\\ \Delta Y_{2t}\\ \Delta Y_{3t}\\ \Delta Y_{4t}\end{bmatrix}$$

The following figure shows the plot of C_t. I normalize C_t so that in january 1967 it has a value of 100. As you can see, the behavior of the state variable is similar to the one from the original paper.

Figure 2: Replication of C_t



Part b)

The likelihood function for this model, up to a constant, is given by:

$$\ln L_Y(\Theta) = -\frac{1}{2}\sum_{t=1}^{T} \ln|\Sigma_t(\Theta)| - \frac{1}{2}\sum_{t=1}^{T} \epsilon_t(\Theta)'\Sigma_t(\Theta)^{-1}\epsilon_t(\Theta)$$

where ϵ_t are the innovations, given by:

$$\epsilon_t = y_t - Za_t^{t-1}$$

and Σ_t is the covariance matrix of the innovations, given by:

$$\Sigma_t = ZP_t^{t-1}Z'$$

In order to estimate the model with recent data, I downloaded data from Fred for the four dependent variables in the model.

Table 3: Data Sources

Data	Description	Availability (Start date)
CMRMTSPL	Real Manufacturing and Trade Sales	1967-01-01
INDPRO	Industrial Production	1919-01-01
PAYEMS	Employment in nonagricultural	1939-01-01
W875RX1	Real Personal Income	1959-01-01

Since we don't have data for Real Manufacturing and Trade Sales starting in 1960, we will estimate the model from 1967-2019. The following table reproduces the estimates for the parameters using Maximum Likelihood.

Table 4: Parameters estimates

	Y_1	Y_2	Y_3	Y_4
γ_i	0.0053	0.0017	0.0047	0.0010
ψ_{1i}	-0.0600	-0.1951	-0.4708	0.2208
ψ_{2i}	-0.2575	-0.0035	-0.3183	0.6425
σ_i	0.0049	0.0077	0.0074	0.0054
	$\phi_1 =$	0.3862	$\phi_2 =$	0.1136

Code

First question of this problem set was solved using R. For the second part, the data was previously treated using R and part a) was also solved using R. For part b, I used matlab to solve it. I included all the codes in the folder I shared.