Section 1

Question-1 State what is the objective of the paper and how does panel data help in analyzing the objective of the study in a better manner.

Answer-1]

Introduction- The financial health of the commercial banks depends on the possession of good credit risk management dynamics. Commercial banks may have a keen awareness of the need to identify, measure, monitor and control credit risk as well as to determine that they hold adequate capital against these risks. They need to ensure that they are adequately compensated for risks incurred so that the profitability if Public and Private sector banks in India.

Objective – the study is aimed at investigating the impact of credit risk on profitability of Indian Public and Private sector banks for a sample of pre-covid period i.e., 2017-2019. The research of this study involves conducting the analytical test of panel data. As it's known, the panel data is a combination of cross section data and time series, where the same unit cross section is measured across different times. The application of the panel data analysis in the studies of credit risk on profitability is to overcome the lack of robustness in the modelling.

The analysis employed PooledOLS, Fixed and Random effects models to determine the impact of credit risk on the performance of 33 Public and Private sector banks.

Question-2 Discuss how your analysis is different from the reference paper and link it to the objective of the reference paper. (e.g.) if the reference paper is using net FDI as one of the variables then you can change it to FDI/GDP that is, ratio of net FDI inflow to GDP.

Answer-2] The analysis takes its foundation from the paper, **"Credit Risk and Financial Performance of the Jordanian Commercial Banks: A Panel Data Analysis", dated January 2019 by Ghaith N. Al-Eitan and Tareq Bani-Khalid.** At similar front the objective of both the papers is to analyze the impact of credit risk on the performance of the banks. The base paper tests this for the commercial banks listed in the Amman Stock Exchange for the period 2008-2017. Whereas, this particular paper runs the same analysis for Public and Private banks in India. Major difference between the two lies in the proxies considered for the independent and dependent variables. The base paper uses Return on Assets and Return on Equity as the measure of profitability of banks. On the other hand, the independent variables consist of Credit risk (variable for this is not defined in the paper which could considered a flaw in the model specification), Doubtful loans, non-performing loan, Loan loss to total Loan, size of bank and total deposits. Whereas the present paper defines the variables clearly. The dependent variable used is Return on Assets (ROA) as a proxy for Bank's Profitability and the independent variable used is credit risk which is measured by to proxies namely, Non-performing Assets to Loans ratio (NPALR) and Capital Adequacy ratio (CAR). The units of all the variables are in percentage form.

Question-3 Given the above objective, what are the dependent and independent variables in the model. If you have endogeneity issue then clearly mention the IV. Make sure that you also clearly indicate whether a log transformation or a quadratic specification is important for the analysis. Sometimes you can use lag of one or two independent variables as additional explanatory variables.

Answer-3]

Variable	Formula	Unit	Short Form
Return on Assets	ROA = (Net Income\Total Assets)*100	in %	ROA
Capital Adequacy Ratio	[(Tier I Capital + Tier II Capital) / Risk-Weighted Assets]*100	in %	CAR
Non-Performing Assets to Loan and Advances ratio	(Net Non-Performing Assets/Total Advances)*100	in %	NPALR

Figure 1

No endogeneity issues were found in the data. The unit of all the three variables in the study is in percentage form. The independent variables are ratios, formulas for which are stated in the table (Table 1) above. Taking lags of the independent variables particularly in this model was not found to be relevant and hence it's not employed for the analysis.

Note: detailed description of the variables is given in the next question.

Question-4 Mention the data source(s) and variables used with their proper definition. (e.g.) if you are using net FDI as one of the variables then define the variable as it is mentioned in the data source. Make sure the constant price values or PPP adjusted values are all properly taken into account.

Answer-4]

Data Source –For our analysis, we have borrowed the time series data from the annual accounts of scheduled commercial banks as compiled at Database on Indian Economy, Reserve Bank of India.

Description of variables

Variable	Formula	Description
Return on Assets	ROA = (Net Income\Total Assets)*100	The term return on assets (ROA) refers to a financial ratio that indicates how profitable a company is in relation to its total assets.
Capital Adequacy Ratio	[(Tier I Capital + Tier II Capital) / Risk-Weighted Assets]*	Capital adequacy is the measure of financial strength of the commercial banks. It is also a measure of ability to absorb the financial risk that may be incurred in the commercial bank.
Non -Performing Assets to Loan and Advances ratio	(Net Non-Performing Assets/Total Advances)*100	Non -Performing Assets to Loan and Advances ratio is used to measure the level of the bank's credit risk and quality of outstanding loans.

Figure 2

Question-5 Mention clearly the number of cross-section entities and any specific features they have (e.g., 100 low- and middle-income countries) and the number of years with the years mentioned (e.g., 3 years 2005, 2010 and 2015). Make sure that N>=50 and N>>T.

Answer-5] The number of cross section entities (N) considered are 33. (12 Public Sector Banks and 22 Private sector Banks). The number of years (T) is three (3), 2017, 2018 & 2019. The data was obtained in calendar year form. The total number of observations in this case becomes (N*T) = 99

Section 2

Question-7 Use the xtsum command of STATA or appropriate commands in the other software to present the summary statistics. Discuss the findings from the summary statistics by taking the help of STATA manual or any other resource.

Answer-7] This table shows the descriptive statistics of the variables used in the study of impact of credit risk on profitability of public and private sector banks in India for the period 2017-2019. The average profitability measured by Return on Assets (ROA) in Indian banks is 0.22, which ranges from minimum -4.68 to maximum 4.46 with standard deviation 1.41. Since average ROA is low it indicates that average profitability in Indian public and private banks is low. The minimum and maximum value of Capital Adequacy Ratio are 9.04 and 31.48 with mean 13.95 and standard deviation 3.62. As the average of CAR is good it indicates that banks have moderately adequate capital to face financial risks. Similarly, Non -Performing Assets to Loan and Advances ratio ranges from 0.33 to 16.69 with mean 4.56 and standard deviation 3.91. Since the average NPALR is not very low, it indicates that the policy of public and private banks in India to manage credit and grant loans to customers is not very efficient.

Variables	N	Minimum	Maximum	Mean	Std. Deviation
ROA (Y)	99	-4.68	4.48	0.22	1.41
CAR (X1)	99	9.04	31.48	13.95	3.62
NPALR (X2)	99	0.33	16.69	4.56	3.91

Figure 3

Section 3.1

Question-8 Coefficient Estimates: Based on the objective of your analysis compare the results for the following models with and without time dummies
a. Pooled Model
b. FE model
c. RE model
d. For each of the models in a, b, c above present F-test results on whether year dummies are jointly significant or not.
e. Discuss briefly the overall findings of the results in terms of coefficient significance, signs, magnitudes, differences in results across the models.
f. Interpret the coefficient of the variable that is most relevant to the objective of the study.

Answer-8]

Models Comparison

1. PooledOLS Vs FE

PooledOLS - $ROA_{it} = \beta_1 + \beta_2 CAR_{it} + \beta_3 NPALR_{it} + \varepsilon_{it}$

FE - $ROA_{it} = \alpha_i + \beta_2 CAR_{it} + \beta_3 NPALR_{it} + \varepsilon_{it}$

RE - $ROA_{it} = \beta_1 + \beta_2 CAR_{it} + \beta_3 NPALR_{it} + u_i + \varepsilon_{it}$

Variable Effect	With Dummy									Without Dummy								
	Pooled OLS			Fixed Effect			Random Effect			Pooled OLS			Fixed Effect			Random Effect		
	Coef.	t-stat	p-value	Coef.	t-stat	p-value	Coef.	t-stat	p-value	Coef.	t-stat	p-value	Coef.	t-stat	p-value	Coef.	t-stat	p-value
const	-1.32	-3.747	0.0003	-0.81	-0.87	0.3870	-1.41	-3.02	0.0032	-1.63	-4.27	0.0000	-0.33	-0.33	0.739	-1.66	-3.49	0.0007
CAR(X1)	0.19	8.9	0.0000	0.09	1.56	0.1223	0.19	7.09	0.0000	0.18	8.21	0.0000	0.04	0.56	0.576	0.19	6.72	0.0000
NPALR(X2)	-0.17	-7.6339	0.0000	-0.01	-0.28	0.7793	-0.16	-6.38	0.0000	-0.17	-7.43	0.0000	0.01	0.15	0.8793	-0.16	-6.26	0.0000
Year_Cat 2018	-0.32	-3.01	0.0034	-0.35	-2.32	0.0233	-0.33	-2.11	0.0373	Nil	Nil	Nil	Nil	Nil	Nil	Nil	Nil	Nil
Year_Cat 2019	-0.72	-3.72	0.0003	-0.54	-3.47	0.0009	-0.71	-4.56	0.0000	Nil	Nil	Nil	Nil	Nil	Nil	Nil	Nil	Nil
R-squared	0.78	Nil	Nil	0.17	Nil	Nil	0.70	Nil	Nil	0.73	Nil	Nil	0.01	Nil	Nil	0.68	Nil	Nil
F-stat	79.48	Nil	Nil	3.26	Nil	Nil	55.54	Nil	Nil	128.71	Nil	Nil	0.16	Nil	Nil	Nil	Nil	Nil
Prob >F	0.00	Nil	Nil	0.02	Nil	Nil	0.00	Nil	Nil	0.00	Nil	Nil	0.85	Nil	Nil	Nil	Nil	Nil
F-test (Poolability)	Nil	Nil	Nil	1.97	Nil	0.01	Nil	Nil	Nil	Nil	Nil	Nil	1.99	Nil	0.0099	Nil	Nil	Nil
Hausman Test	Nil	Nil	Nil	Nil	Nil	Nil	14.20	Nil	Nil	Nil	Nil	Nil	Nil	Nil	Nil	22.18	Nil	Nil
No. of Obs.	99.00	Nil	Nil	99.00	Nil	Nil	99.00	Nil	Nil	99.00	Nil	Nil	99.00	Nil	Nil	99.00	Nil	Nil

Figure 4

With time dummies – when dummies are allotted to the 3 time periods for the study, it was found that overall, both Pooled and FE model were significant (measured by F-test). However, the pooled model was found to be explaining more variations in the Return on Assets which are caused by the management of the Credit Risk which is measured by Capital Adequacy Ratio and Non-Performing Assets to Loans ratio. This is confirmed by the higher R squared in the pooled model which is around 77%. Meaning, 77% of the variations in the dependent variable which is ROA in our case are explained by the independent variables [CAR (%) & NPALR(%)], which in fact is very high as compared to the R squared in FE model which is, 17%.

Without time dummies – when dummies are not assigned to the time periods then the fixed effect model is not coming out to be significant. The PooledOLS model is overall and individually significant. Whereas the FE model without time dummies has an extremely small R squared which is almost negligible at 1%. The individual variables are also not coming out to be significant.

2. PooledOLS Vs RE

With time dummies – as compared to the PooledOLS model the Random effects model is also overall significant with the p-value of 0.00<0.05 (at 5% level of significance) and R squared is 0.70, meaning 70% of the variations in the ROA is explained by the fluctuations in the credit risk.

Without time dummies – between RE and pooled model the latter is giving better results in terms of overall and individual significance of the variables. The R squared for RE model without time dummies is around 67% which is lower than the pooled model (R squared 73%)

3. Fixed Effects Vs Random Effects

With time dummies - in the time dummies FE model is performing the worst of the three models in terms of overall and individual significance of the variables. The RE model is still better as the variables are overall and individually significant. The R squared is also higher as compared to the FE model, i.e., around 70%.

Without time dummies- even in the without time dummies models FE model is the worst performing model. This indicates that for the study the entities are not fixed but are random.

d. the time dummies are jointly and individually significant which can be confirmed with the values reflected by the table given above. F-stat values for pooled, FE and RE models are 79.48, 3.26 and 55.54 respectively. Which we can see are very high and even p-values for the time dummies are less than 0.05b which means that time dummies are overall and individually significant at 5% level of significance.

e. Note: Without time dummy results

The variables are statistically significant when we talk about RE and pooled model. But on the contrary the variables are not significant in FE model although their sign remains right. The expected sign for CAR and NPALR is positive and negative respectively.

PooledOLS - $ROA_{it}^* = -1.63^* + 0.18CAR_{it}^* - 0.17NPALR_{it}^* + \varepsilon_{it}$

R Squared = 0.63

In Pooled model a 1% increase in the CAR leads to a 0.18% increase on an average in the Return on assets for a bank. On the other hand, a 1% increase in the NPALR (which is the increase in non-Performing assets or a decrease in loan and advances) leads to a 0.17% decrease in the ROA.

FE - $ROA_{it}^* = \alpha_i + 0.04CAR_{it}^* - 0.01NPALR_{it}^* + \varepsilon_{it}$

R Squared = 0.01

In FE model a 1% increase in the CAR leads to a 0.04% increase on an average in the Return on assets for a bank. On the other hand, a 1% increase in the NPALR (which is the increase in non-Performing assets or a decrease in loan and advances) leads to a 0.01% decrease in the ROA. However, FE model is coming out to be statistically insignificant therefore we may not take these results into consideration.

RE - $ROA_{it}^* = -1.66 + 0.19CAR_{it}^* - 0.16NPALR_{it}^* + u_i + \varepsilon_{it}$

Where,

ε_{it} = is the residual as a whole where the residual is a combination of cross section and time series.

u_i = is the individual residual which is the random characteristic of unit observation the i-thand remains at all times.

R Squared = 0.73

In RE model a 1% increase in the CAR leads to a 0.19% increase on an average in the Return on assets for a bank. On the other hand, a 1% increase in the NPALR (which is the increase in non-Performing assets or a decrease in loan and advances) leads to a 0.16% decrease in the ROA. This model is giving the most appropriate results in terms of the significance, magnitude and sign of the variables. Both the variables are statistically significant at 5% level of significance therefore, CAR and NPALR has a positive and negative effect, respectively on ROA.

f. Non-performing assets to loans ratio can be taken as the most important variable between the two. An increase in non-performing assets should have a negative impact on profitability of a bank or according to the study on ROA theoretically. The RE model is giving the best results in terms of the magnitude at which the ROA is affected by an increase in NPALR.

Section 3.2

Question-9 Tests of Hypotheses: In all the model comparisons below, present the results along with the appropriate test of hypothesis for the best fitting model with and without time dummies.

a. Pooled with FE
b. Pooled with RE
c. FE with RE
d. Comment on the final results by linking it to the objective of the study

Answer-9] Pooled Vs FE – the value of Poolable F-test is considered from the figure 4

With time dummies

H_0: Pooled model is preferred ($p > 0.05$)
H_1: FE model is preferred ($p < 0.05$)

The value of F-stat is 1.97 and the p-value is 0.01

Result: the p-value is less than 0.05 then we reject the null hypothesis of pooled model. Therefore, FE model is preferred.

Without time dummies

H_0: Pooled model is preferred ($p > 0.05$)
H_1: FE model is preferred ($p < 0.05$)

The value of F-stat is 1.99 and the p-value is 0.0099

Result: the p-value is less than 0.05 then we reject the null hypothesis of pooled model. Therefore, FE model is preferred.

Pooled Vs RE – B

With time dummies

H_0: Pooled model is preferred ($p > 0.05$)
H_1: RE model is preferred ($p < 0.05$)

The value of Breusch Pagan- LM stat is 5.326 and the p-value is 0.149

Result: the p-value is greater than 0.05 then we fail to reject the null hypothesis of pooled model. Therefore, Pooled model is preferred.

Without time dummies

H_0: Pooled model is preferred (p>0.05)
H_1: RE model is preferred (p<0.05)

The value of Breusch Pagan- LM stat is 0.55 and the p-value is 0.757

Result: the p-value is greater than 0.05 then we fail to reject the null hypothesis of pooled model. Therefore, Pooled model is preferred.

FE Vs RE – Hausman test is employed, values in figure 4

<u>With time dummies</u>

H_0: RE model is preferred (p>0.05)
H_1: FE model is preferred (p<0.05)

```
hausman test result
14.204300253551235
[15.08627247 11.07049769  9.2363569 ]
Failed to reject the null hypothesis, RE model is preferred
```

<u>Without time dummies</u>

H_0: RE model is preferred (p>0.05)
H_1: FE model is preferred (p<0.05)

Result: ```hausman test result
22.183528478764906
[11.34486673 7.8147279   6.25138863]
Failed to reject the null hypothesis, RE model is preferred```