May Examination Period 2025

ECN225 Econometrics 2 Duration: 2 hours

YOU ARE NOT PERMITTED TO READ THE CONTENTS OF THIS QUESTION PAPER UNTIL INSTRUCTED TO DO SO BY AN INVIGILATOR

Answer ALL questions. The questions carry equal marks. JUSTIFY ALL YOUR ANSWERS. Cross out any answers that you do not wish to be marked. There are tables of critical values from the standard normal, $F_{m,\infty}$, ADF, χ^2_m, and QLR distribution provided on pages 8-9. Also, the cumulative standard normal distribution tables are provided on pages 10-11.

Only nonprogrammable calculators are permitted in this examination. Please state on your answer book the name and type of machine used.

Complete all rough workings in the answer book and cross through any work that is not to be assessed.

Possession of unauthorised material at any time when under examination conditions is an assessment offence and can lead to expulsion from QMUL. Check now to ensure you do not have any notes, mobile phones, smartwatches or unauthorised electronic devices on your person. If you do, raise your hand and give them to an invigilator immediately.

It is also an offence to have any writing of any kind on your person, including on your body. If you are found to have hidden unauthorised material elsewhere, including toilets and cloakrooms it will be treated as being found in your possession. Unauthorised material found on your mobile phone or other electronic device will be considered the same as being in possession of paper notes. A mobile phone that causes a disruption in the exam is also an assessment offence.

EXAM PAPERS MUST NOT BE REMOVED FROM THE EXAM ROOM

Examiner: Dr. Jinu Lee

Question 1 (25 marks)

Suppose that the time series X_t follows a moving average process given by:

$$X_t = 2 + 0.5u_{t-1} + 0.2u_{t-2} + u_t,$$

where the error terms u_t are independent and identically distributed (IID) with $E(u_t) = 0$ and $\text{Var}(u_t) = 3$, i.e. $u_t \sim iid(0, 3)$.

(a) Compute the first three autocovariances of X_t, i.e., γ_1, γ_2 and γ_3. Clearly justify each step of your derivation using expectations and the key properties of moving average (MA) models. Additionally, interpret your results by discussing what these autocovariances reveal about the dependence structure of the process.

(5 marks)

To analyse and forecast time series data, a researcher applies a first-order autoregressive model, AR(1), and estimates it using the Ordinary Least Squares (OLS) method. However, the true underlying process follows a different specification:

$$Y_t = 1 + Y_{t-1} + u_t, \quad \text{(True Model)}.$$

(b) Discuss the potential estimation issues that arise due to this model misspecification, including implications for bias, consistency, and statistical inference.

(5 marks)

A macroeconomic analyst aims to predict changes in inflation rates by using autoregressive models of order p, denoted as AR(p). To identify the most suitable model, the analyst estimates several AR models with lag orders ranging from $p = 1$ to $p = 6$.
Using $T = 172$ months of historical data, the analyst obtains the following results for the logarithm of the sum of squared residuals (SSR) per observation:

p	0	1	2	3	4	5	6
$\ln\left(\frac{SSR(p)}{T}\right)$	1.065	1.007	0.865	0.838	0.837	0.836	0.836

(c) Using the following model selection criteria:

$$AIC(p) = \ln\left(\frac{SSR(p)}{T}\right) + \frac{2(p+1)}{T}, \quad BIC(p) = \ln\left(\frac{SSR(p)}{T}\right) + (p+1)\frac{\ln T}{T},$$

determine the optimal lag order (p) for forecasting changes in inflation rates. Discuss which criterion favors a more parsimonious model and explain the reasoning behind it. Justify the findings by considering the trade-off between model complexity and goodness of fit.

(5 marks)

An economic research institute is analysing the relationship between quarterly real GDP growth and the term spread (the difference between long-term and short-term interest rates) over time to enhance macroeconomic forecasting.

(d) To determine whether the lags of the term spread contain useful predictive information for GDP growth, outline the procedure for conducting an appropriate statistical test in time series analysis. Explain the steps involved in formulating and implementing the test, including the null and alternative hypotheses, estimation method, and decision criteria.

(5 marks)

(e) Based on their empirical and statistical analysis of the relationship, the researchers estimated the following model using quarterly data from 2000:Q1 to 2023:Q4. The estimations pertain to the log difference of real GDP ($\Delta \ln GDP$, billions of currency units):

$$\Delta \widehat{\ln GDP_t} = 0.75 - 0.22\Delta \ln GDP_{t-1} - 0.30\Delta \ln GDP_{t-2} + 0.15\Delta \ln GDP_{t-3} - 0.12\Delta \ln GDP_{t-4} - 0.18TS_{t-1}.$$

Using the recent observations on $\ln GDP_t$ and TS_t (term spread) over the last several quarters, as provided in the table:

	2022:Q3	2022:Q4	2023:Q1	2023:Q2	2023:Q3	2023:Q4
$\ln GDP_t$	4.942	4.950	4.958	4.966	4.974	4.980
TS_t	-1.7	-1.5	-1.3	-1.1	-0.9	-0.8

compute the forecasted GDP for the next quarter, 2024:Q1, i.e.,

$$\widehat{GDP}_{2024:Q1|2023:Q4}.$$

(5 marks)

Question 2 (25 marks)

Suppose the time series X_t and Y_t are defined by the following autoregressive models,

$$Y_t = 0.6Y_{t-1} + 0.4Y_{t-2} + u_t, \quad X_t = 1.5 + 0.85X_{t-1} - 0.25X_{t-2} + u_t.$$

where u_t is an independent and identically distributed (i.i.d.) error term with $E(u_t) = 0$ and Var$(u_t) = \sigma^2$.

(a) Determine whether X_t and Y_t are stationary or nonstationary. **(5 marks)**

(b) A researcher estimates the following Ordinary Least Squares (OLS) regression,

$$Y_t = \beta_0 + \beta_1 X_t + u_t.$$

Evaluate the validity of this regression approach. Discuss any potential econometric issues that may arise due to the properties of X_t and Y_t. If any issues exist, suggest appropriate solutions to address them.

(5 marks)

A financial analyst investigates whether the log exchange rate ($\ln ER_t$) follows a random walk by estimating the following Augmented Dickey-Fuller (ADF) regression,

$$\Delta \widehat{\ln ER_t} = \underset{(0.15)}{0.35} - \underset{(0.032)}{0.087} \ln ER_{t-1} - \underset{(0.055)}{0.21} \Delta \ln ER_{t-1} - \underset{(0.051)}{0.18} \Delta \ln ER_{t-2} + \underset{(0.048)}{0.12} \Delta \ln ER_{t-3}$$

where the number of lags was selected using the Akaike Information Criterion (AIC).

(c) Assess whether the log exchange rate ($\ln ER_t$) contains a unit root. Justify your answer based on the ADF regression results. If further testing is required, explain what additional steps should be taken to determine stationarity.

(5 marks)

A macroeconomist is investigating whether the predictive relationship between consumer sentiment (CS_t) and household consumption growth (CGR_t) has remained stable over time. To analyse this, she estimates an Autoregressive Distributed Lag (ADL) (2,2) model, which includes two lags of both consumption growth and consumer sentiment,

$$CGR_t = \beta_0 + \beta_1 CGR_{t-1} + \beta_2 CGR_{t-2} + \beta_3 CS_{t-1} + \beta_4 CS_{t-2} + u_t.$$

To test for a potential structural break in the predictive relationship, she applies a Quandt Likelihood Ratio (QLR) test with a 15% trimming on all coefficients.

(d) Explain the principles behind the QLR testing method and how it is used to detect instability in the relationship between consumer sentiment and consumption growth. **(5 marks)**

(e) Suppose the resulting QLR test statistic is 4.25. Interpret this result. **(5 marks)**

Question 3 (25 marks)

A research institute is studying job mobility and its relationship with wage growth in a competitive labor market. The hypothesis is that individuals are more likely to switch jobs when their expected wage growth in a new position exceeds the wage increase they would receive if they stayed in their current job along, while also considering broader economic conditions. Using 5,162 observations, the institute estimates the following logit regression, where the dependent variable is 1 if the individual switched jobs within the past year and 0 otherwise:

Variable	*constant*	*female*	*minority*	*married*	*job tenure*	Δ *wage*	*unemployment*
Coef. Estimate	-0.800	0.050	-0.250	-0.150	-0.080	1.200	-0.400
(Standard Error)	(0.250)	(0.110)	(0.190)	(0.150)	(0.055)	(0.320)	(0.450)

where:

- *female*, *minority*, and *married* are binary variables indicating whether the individual is female, from a racial minority group, and married, respectively.
- *job tenure* measures the number of years an individual has been at their current job.
- Δ *wage* represents the expected percentage wage increase if the individual switches jobs.
- *unemployment* is the state-level unemployment rate at the time of the job switch decision.

(a) Interpret the estimated logit model. Comment on the significance and economic meaning of the coefficients. **(5 marks)**

(b) Compute the probability of job switching for an individual who is female, belongs to a racial minority, and is married using the following sample means of the variables:

Variable	*job tenure*	Δ *wage*	*unemployment*
Mean	4.6	0.08	5.2

(5 marks)

(c) Suppose the unemployment rate increases from 5.2% to 7.0%, making job switching riskier. Recalculate the probability of switching jobs as in part (b), and discuss how this change in economic conditions affects job mobility.

(5 marks)

A marketing analyst at an online subscription service is studying customer retention behavior over a one-year period. Each customer is classified into one of four possible subscription statuses after one year:

- Cancelled Subscription (0): Probability p_1
- Downgraded to a Lower Plan (1): Probability p_2
- Retained on the Same Plan (2): Probability p_3
- Upgraded to a Higher Plan (3): Probability $1 - p_1 - p_2 - p_3$

where $0 < p_1, p_2, p_3 < 1$ and $p_1 + p_2 + p_3 < 1$.
The analyst collects a random sample of n = 1000 customers and records the following observations:

Subscription Status	Number of Customers
Cancelled ($X = 0$)	250
Downgraded ($X = 1$)	120
Retained ($X = 2$)	500
Upgraded ($X = 3$)	130

(d) Derive the log-likelihood function for the given sample and explain its components.

(5 marks)

(e) Find the maximum likelihood estimators (MLEs) for p_1, p_2, and p_3, compute their values using the sample data, and interpret the results.

(5 marks)

Question 4 (25 marks)

A research institute is investigating how remote work adoption influences wages over time. The dataset consists of a balanced panel of 600 employees observed from 2019 to 2024 across various industries. The regression model used for analysis is:

$$\ln(Wage_{it}) = \beta_1 + \beta_2 Exp_{it} + \beta_3 Exp_{it}^2 + \beta_4 Hours_{it} + \beta_5 Remote_{it} + \beta_6 Ind_{it} + \beta_7 Urban_{it} + \beta_8 Edu_i + \beta_9 Gender_i + \beta_{10} Race_i + u_{it}$$

where:

- Exp: Years of work experience
- $Hours$: Average weekly working hours
- $Remote$: 1 if the employee works remotely, 0 if fully on-site
- Ind: 1 if working in the tech industry, 0 otherwise
- $Urban$: 1 if living in a metropolitan area, 0 otherwise
- Edu: Years of education (time-invariant)
- $Gender$: 1 if female, 0 if male (time-invariant)
- $Race$: 1 if non-white, 0 if white (time-invariant)

Table 1: Regression Estimates for Log-Salaries

Variable \Regression	Model (1)	Model (2)	Model (3)
Exp	0.0452 (0.0051)	0.110 (0.0048)	0.108 (0.0022)
Exp^2	-0.000580 (0.000090)	-0.000350 (0.000065)	-0.000330 (0.000061)
$Hours$	0.0032 (0.0018)	0.0007 (0.0009)	0.0005 (0.0008)
$Remote$	0.076 (0.028)	0.030 (0.017)	0.032 (0.016)
Ind	0.125 (0.022)	0.047 (0.015)	0.050 (0.014)
$Urban$	0.095 (0.021)	-0.012 (0.016)	-0.015 (0.015)
Edu	0.067 (0.0063)	-	-
$Gender$	-0.200 (0.042)	-	-
$Race$	-0.105 (0.038)	-	-
Individual Effects	No	Yes	Yes
Time Effects	No	No	Yes
$\bar{R}^2$	0.431	0.897	0.902

Table 2: Tests for joint significance of effects (H_0: coefficients on effects are jointly zero).

Model	Test	F statistic	p-value
(2)	*Individual Effects*	38.2	< 0.0001
(3)	*Individual Effects*	36.2	< 0.0001
(3)	*Time Effects*	48.2	< 0.0001

(a) Interpret the coefficient on remote work in Model (1) in Table 1. Does this suggest a wage premium or penalty for remote workers? **(5 marks)**

(b) In Model (2) in Table 1, the variables Edu, $Gender$, and $Race$ are omitted. Why are these time-invariant variables excluded when using fixed effects? **(5 marks)**

(c) Compare Model (1) and Model (2) in Table 1. What differences emerge in the coefficient estimates, particularly for $Remote$ and $Experience$? What might explain these changes? **(5 marks)**

(d) Model (3) in Table 1 introduces time fixed effects. What is the rationale for including them, and how do they alter the regression results compared to Model (2) in Table 1? **(5 marks)**

(e) Which of the three models in Table 1 —(1), (2), or (3)— is the most appropriate for evaluating the impact of remote work on wages? Justify your answer. **(5 marks)**

End of Paper - An Appendix of 4 pages follows

Large Sample Critical Values for the t-statistic from the Standard Normal Distribution			
	Significance level		
	10%	5%	1%
2-Sided Test ($\neq$) Reject if $\|t\|$ is greater than	1.64	1.96	2.58
1-Sided Test ($>$) Reject if t is greater than	1.28	1.64	2.33
1-Sided Test ($<$) Reject if t is less than	-1.28	-1.64	-2.33

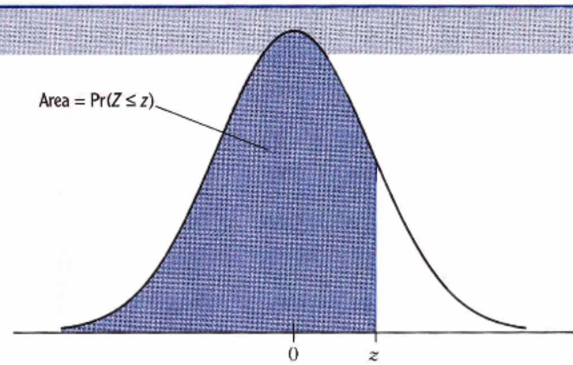
Critical Values for the $F_{m,\infty}$ Distribution			
	Significance level		
Degrees of freedom	10%	5%	1%
1	2.71	3.84	6.63
2	2.30	3.00	4.61
3	2.08	2.60	3.78
4	1.94	2.37	3.32
5	1.85	2.21	3.02
6	1.77	2.10	2.80
7	1.72	2.01	2.64
8	1.67	1.94	2.51
9	1.63	1.88	2.41
10	1.60	1.83	2.32

Large Sample Critical Values of the Augmented Dickey-Fuller Statistic			
	Significance level		
Deterministic regressors	10%	5%	1%
Intercept only	-2.57	-2.86	-3.43
Intercept and time trend	-3.12	-3.41	-3.96

Critical Values for the χ^2 Distribution			
	Significance level		
Degrees of freedom	10%	5%	1%
1	2.71	3.84	6.63
2	4.61	5.99	9.21
3	6.25	7.81	11.34
4	7.78	9.49	13.28
5	9.24	11.07	15.09
6	10.64	12.59	16.81
7	12.02	14.07	18.48
8	13.36	15.51	20.09
9	14.68	16.92	21.67
10	15.99	18.31	23.21

Critical Values of the QLR Statistic with 15% Trimming			
Number of Restrictions (q)	10%	5%	1%
1	7.12	8.68	12.16
2	5.00	5.86	7.78
3	4.09	4.71	6.02
4	3.59	4.09	5.12
5	3.02	3.66	4.53
6	2.84	3.37	4.12
7	2.69	3.15	3.82
8	2.58	2.98	3.57
9	2.48	2.84	3.38
10	2.40	2.71	3.23

TABLE 1 The Cumulative Standard Normal Distribution Function, $\Phi(z) = \Pr(Z \le z)$



	Second Decimal Value of z									
z	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
−2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
−2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
−2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
−2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
−2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
−2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
−2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
−2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
−2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
−2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
−1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
−1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
−1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
−1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
−1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
−1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
−1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
−1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
−1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
−1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
−0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611

continued on next page

TABLE 1 *(continued)*

z	Second Decimal Value of z									
	0	1	2	3	4	5	6	7	8	9
−0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
−0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
−0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
−0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
−0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
−0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
−0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
−0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
−0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986

This table can be used to calculate $\Pr(Z \leq z)$ where Z is a standard normal variable. For example, when $z = 1.17$, this probability is 0.8790, which is the table entry for the row labeled 1.1 and the column labeled 7.

End of Appendix