STOCKHOLMS UNIVERSITET
Statistiska institutionen
Oskar Gustafsson

HOME RE-EXAM, ECONOMETRICS II

2021-08-17

Time for examination: 9.00-14.00, The exam shall be submitted electronically via the department's web site no later than 15.00. The system does not allow submission after deadline. Therefore, start the submission well in advance. The last hour of the exam time is intended for arranging the electronic submission.
All necessary information about submission, anonymous code, extended writing time, etc. can be found in a separate file. If you, despite the instructions, have problems submitting the exam, email the exam to tenta@stat.su.se. However, this is only done in exceptional cases. Exams sent in by email after deadline will not be corrected.

For questions regarding the submission, email to: expedition@stat.su.se. Practical help is only available during the first hour of the exam. For questions regarding the content of the exam, email to: edgar.bueno@stat.su.se. Incoming email questions are answered continuously during the exam. If the course coordinator needs to send out information to all students during the exam, this is done to your registered email address. Therefore, check your email during the exam.

Note: The exam should be written individually. All types of collaborations and/or help from others are strictly forbidden. Suspected cheating is reported to the Disciplinary Board and can lead to suspension
Allowed tools: Pocket calculator, course books and lecture notes.

The exam consists of 4 independent problems. Well motivated and clear solutions are required for full scoring on a problem (multiple choice questions are exceptions). Don't forget to state any necessary assumptions or conditions where needed.
Passing rate: 50% of overall total, which is 100 points. For detailed grading criteria, see the course description. Answers may be given in English or Swedish.

Good luck!

Problem 1. (20 points)
Indicate which alternative that is correct. Answering more than one alternative result in 0 points on the sub-question. No motivation is required.

1. The time series process $y_t = 0.5y_{t-1} + 0.4y_{t-2} + \varepsilon_t$ will have:
   (a) an ACF that show a dampening sinusoidal pattern
   (b) an exponentially decaying PACF
   (c) an exponentially decaying ACF
   (d) an ACF that cuts off after 2 lags.

2. Serial autocorrelation in the error term of the linear regression model will in general lead to (unless accounted for):
   (a) inconsistent OLS estimates
   (b) inconsistent OLS estimates and incorrect standard errors
   (c) efficient parameter estimates
   (d) wrong t-values

3. Witch statement is **not** true about cointegrated processes:
   (a) we can test whether 3 time series are jointly cointegrated
   (b) they all have unit roots
   (c) each time series should be made stationary before fitting a model
   (d) cointegration models (VECM) can be estimated both with OLS and maximum likelihood

4. With panel data methods we cannot account for omitted variables that:
   (a) are heteroscedastic
   (b) vary over time and entity
   (c) are unobserved
   (d) are dummy variables

5. The second difference of a time series can be written as:
   (a) $\Delta^2 y_t = y_t - y_{t-2}$
   (b) $\Delta^2 y_t = y_t - 2y_{t-1} + y_{t-2}$
   (c) $\Delta^2 y_t = B(y_t - y_{t-1})$
   (d) $\Delta^2 y_t = y_t - y_{t-1} - y_{t-2}$

Problem 2. (30 points)

A grocery store at a vacation spot sell cold beverages of a certain brand. The manager of the store buys the product from a local producer and want to be sure to order a reasonable amount of batches from the producer. The manager has to place orders for one quarter at the time and she has access to 15 years of historical data seen in Figure 1.

Table 1: Beverage data - last 2.5 years

t	17.Q3	17.Q4	18.Q1	18.Q2	18.Q3	18.Q4	19.Q1	19.Q2	19.Q3	19.Q4
y	112	85	92	116	121	93	99	122	129	100

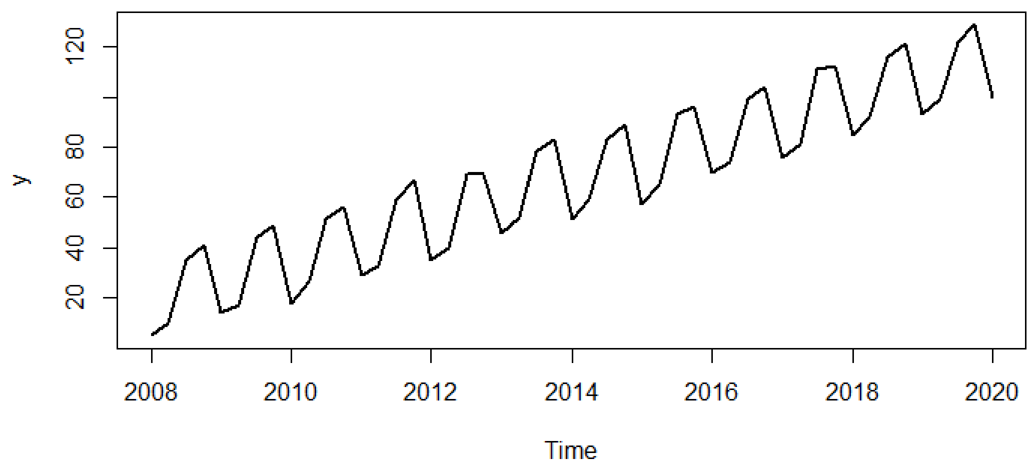

Figure 1: Sales data for the local beverage.

1. Based on Figure 1, would you choose multiplicative- or additive model? Why?
2. Use Holt-Winters seasonal exponential smoothing to smooth the data in Table 1. Use the smoothing constants 0.3, 0.6, 0.8 for the level, trend and seasonal components respectively. For the starting values use: $L_0 = 90$, $b_0 = 2$, and $S_0 = (-12, 12, 15, -15)$ for each quarter $(Q_1 - Q_4)$.
3. Use the model to make forecasts for each quarter of 2020.
4. The true data points where: 105, 135, 135, 110. Compute the *mean absolute deviation*, the *root mean squared error* and the *mean absolute percentage error*.
5. The manager is also interested in having a clear view of the overall sales trend. Make a picture where you plot the real data together with the smoothed data and the smoothed data where the seasonality has been removed, include the forecasts. (You may draw by hand or use your computer to construct the picture).

Problem 3. (30 points) Consider the ARIMA process:

$$y_t = y_{t-1} + 2 + 0.8(y_{t-1} - y_{t-2}) + \varepsilon_t + 0.4\varepsilon_{t-2}, \qquad \text{with } \varepsilon_t \sim N(0,1)$$

1. What are p, d and q for this $ARIMA(p,d,q)$-process?
2. Express the time series in terms of first differences and back-shift operators.
3. Is the first difference stationary and/or invertible?
4. Calculate the mean and variance for the first difference.
5. Assume that we have observed x for 100 periods where the last three observations where:

t	X_t	ε_t
98	13	-0.5
99	12	1
100	14	2.6

   Make point forecasts of x for 1, 2, 3, 4, and 5 steps ahead. If you were to continue to iterate your forecasts, to what will they eventually converge?

Problem 4. (20 points) We are interested in the question of whether a stricter taxation on beer may lead to a lower number of deaths in the traffic. To help us we have a panel data set of 48 US states (excluding Hawaii and Alaska) for the years 1982-1988. Here we are going to make use of 4 variables that are summarized in Table below:

Variable	Description
y	Vehicle fatality rate (10000 thousands)
$BeerTax$	Tax on a case of beer
$income$	Per capita personal income
$unemployment$	unemployment rate

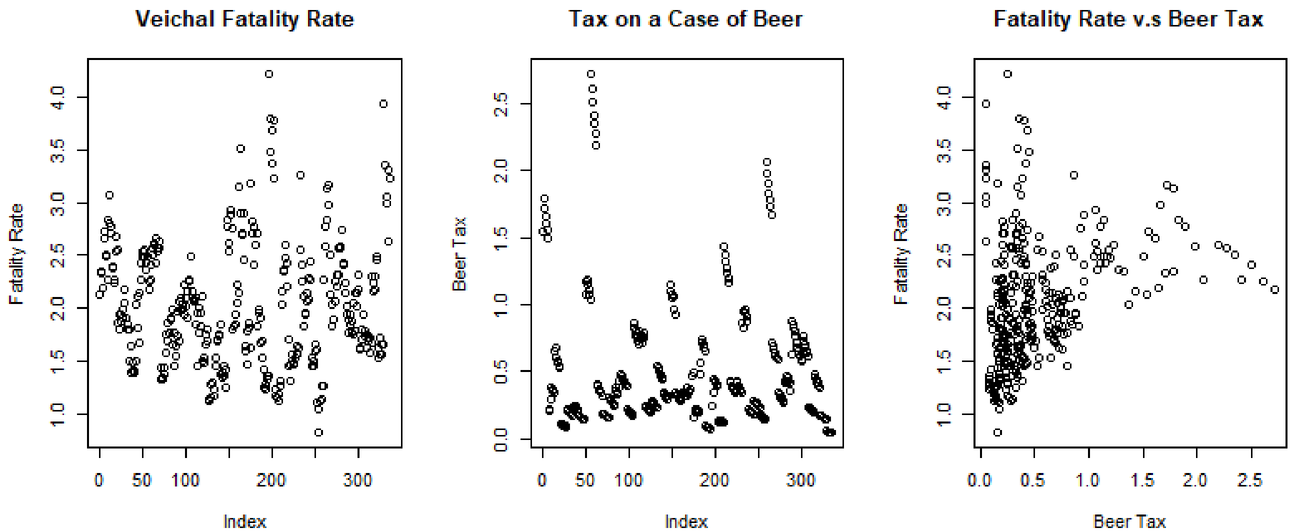

Figure 2: Time series and autocorrelation plots for Problem 1

Our main interest lie in the relation:

$$y = \beta_0 + \beta_1 \times BeerTax + \varepsilon, \quad (1)$$

Questions:

1. Looking at Figure 2, do you see any characteristics that are typical for panel data? Explain.

2. In Table 2, column (1) contain a regression using OLS on the pooled data. Interpret the result, is it what you expected?

3. Calculate the t-statistic of *unemployment* for the OLS regression. How many "stars" should you put on top of it? Do you think this t-statistic is useful?

4. Column (2) and (3) of Table 2 show results for regression models that control for entity- and time-fixed effect. Interpret the results for respective regression. Reason about why they differ.

Table 2:

	Dependent variable:		
	y		
	OLS	*FE*	*FD*
	(1)	(2)	(3)
BeerTax	0.124^{**}	-0.533^{***}	0.123^{**}
	(0.062)	(0.167)	(0.062)
income	-0.0001^{***}	0.0001^{***}	-0.0001^{***}
	(0.00002)	(0.00002)	(0.00002)
unemployment	$-0.027^{?}$	-0.077^{***}	-0.013
	(0.013)	(0.011)	(0.015)
Constant	4.019^{***}		
	(0.302)		
Observations	336	336	336
R^2	0.272	0.321	0.291
Adjusted R^2	0.266	0.185	0.271
Time Fixed Effects	NO	NO	NO
Entity Fixed Effects	NO	YES	NO
Note:			$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$