FIN3006 V21 Midterm

Department of Economics

Examination paper for FIN3006 – Applied Time Series Econometrics

Examination date: April 12, 2021 (09:00) - April 19, 2021 (09:00)

Permitted examination support material: All support material is allowed

Academic contact during examination: Costanza Biavaschi
Phone: 462 39 100

Technical support during examination: Orakel support services
Phone: 73 59 16 00

OTHER INFORMATION

If a question is unclear/vague – make your own assumptions and specify in your answer the premises you have made. Only reach out to academic contact in case of errors or insufficiencies in the question set.

Saving: Answers written in Inspera are automatically saved after you have uploaded your file. If you are working in another program remember to save your answer regularly.

Cheating/Plagiarism: The exam is an individual, independent work. Examination aids are permitted. Inspera uses automatic plagiarism control, but assessments of plagiarism are conducted by the examiner. *Read more about cheating and plagiarism here.*

Citations: We do not require referencing/citations.

Weighting: Unless otherwise stated in the problem set, each question/problem is weighted equally.

Submission: All files must be uploaded before the examination time expires.

- *How to digitize your sketches/calculations*
- *How to create PDF documents*
- *Remove personal information from the file(s) you want to upload*

ABOUT SUBMISSION

- **Your answer will be submitted automatically when the examination time expires and the test closes**, if you have answered at least one question. This will happen even if you do not click “Submit and return to dashboard” on the last page of the question set. You can reopen and edit your answer as long as the test is open. If no questions are answered by the time the examination time expires, your answer will not be submitted.
- **Withdrawing from the exam:** If you wish to submit a blank test/withdraw from the exam, go to the menu in the top right-hand corner and click “Submit blank”. This cannot be undone, even if the test is still open.

- **Accessing your answer post-submission:** You will find your answer in *Archive* when the examination time has expired.

1 Exam text

You are given here the exam text, consisting of one question with several subquestions. Even if you have not used Stata in your course, the command names are self-explanatory and you should be able to quickly figure out what they aim to do.

Midterm Exam S21

When done, upload your answers, in a single PDF file.

Last opp filen her. Maks én fil.

Alle filtyper er tillatt. Maksimal filstørrelse er **50 GB**.

Velg fil for opplasting

Maks poeng: 10

Question 1
Attached

Midterm Exam

FIN 3006

Instructions: This analysis uses Stata. Given that you have a full week to submit your midterm, I assume that even if you are unfamiliar with Stata you should be able to understand the commands. Answer all questions.

"Government agencies periodically release indicators of the level of economic activity in various sectors. However, these releases are typically only available with a reporting lag of several weeks and are often revised a few months later. It would clearly be helpful to have more timely forecasts of these economic indicators. Nowadays there are several sources of data on real-time economic activity available from private sector companies such as Google, MasterCard, Federal Express, UPS, Intuit and many others." (Choi and Varian, 2012, p.2) This in-class test is heavily based on Choi and Varian (2012) and asks you to examine to which extent Google Trends can be used to predict current economic activities. For those unfamiliar with it, Google Trends is a real-time daily or weekly index of the volume of queries that users enter into Google.

In particular, in the following exercises, you will analyze whether initial claims for unemployment benefits can be better forecasted using Google Trends. Initial claims for unemployment benefits measure the number of jobless claims filed by individuals seeking to receive jobless benefits. This number is watched closely by financial analysis and economic institutions because it provides insights into the health of the economy and of the labor market. At the same time, when someone becomes unemployed, it is natural to expect that they will Google searches such as [file for unemployment], [unemployment office], [unemployment benefits], [unemployment claim], [jobs], [resume] and so on. Hence, Choi and Varian argue in their 2012 paper that Google Trends could be useful to predict the number of initial claims, before the official release of this indicator.

The dataset includes weekly observations of initial claims and searches in Google Trends from Jan 10 2004 til July 2 2011, for a total of 391 weeks. On page 4 you find a do-file used to perform the analysis and on page 5 you find the relative log-file.

(a) Provide an asymptotic 95% confidence interval for ACF(1).

(b) Consider the estimated ACF and PACF for initial claims, available in Figure 1 and Figure 2. Identify the process generating this variable, theoretically justifying your answer (Note: these are actual data, so they might be far from the theoretical ACF, PACF you have studied in class. You will be evaluated against your ability of critically look at a problem).

(c) Choi and Varian (2012) continue their analysis by estimating first a simple AR(1) model. Then they estimate an AR(1) in which they include two additional controls: `jobs` and `welfareunemployment`. The first variable - `jobs` - measures search queries on local jobs, while the second - `welfareunemployment` - measures search queries on social services, welfare and unemployment benefits. Consider the information provided in the log files. Test whether `jobs` is a significant predictor of initial claims.

(d) Define and explain the difference between AIC and BIC (SBIC) information criteria. Using these two criteria, what can you conclude from these diagnostic statistics about the appropriate model specification for initial claims?

(e) Derive the three-step-ahead forecast $E_t y_{t+3}$ for the simple AR(1) model estimated. How would your answer change if the estimated parameters of y_{t-1} was closer to zero (but not zero) than to one?

(f) The authors conclude by showing that there is a reduction in predictive power of the model when using Google Trends as additional control variables. An analysis similar to theirs is presented in the do-file from line 29 onwards, and in the log-file from line 29 onwards. Explain which method it is used here to compare model predictions. Would you add any additional steps to assess the forecasting performance of the two models?

(g) Briefly summarize your findings above, emphasizing whether you think that variables based on Google Trends should or should not be used as predictors of current economic activities.

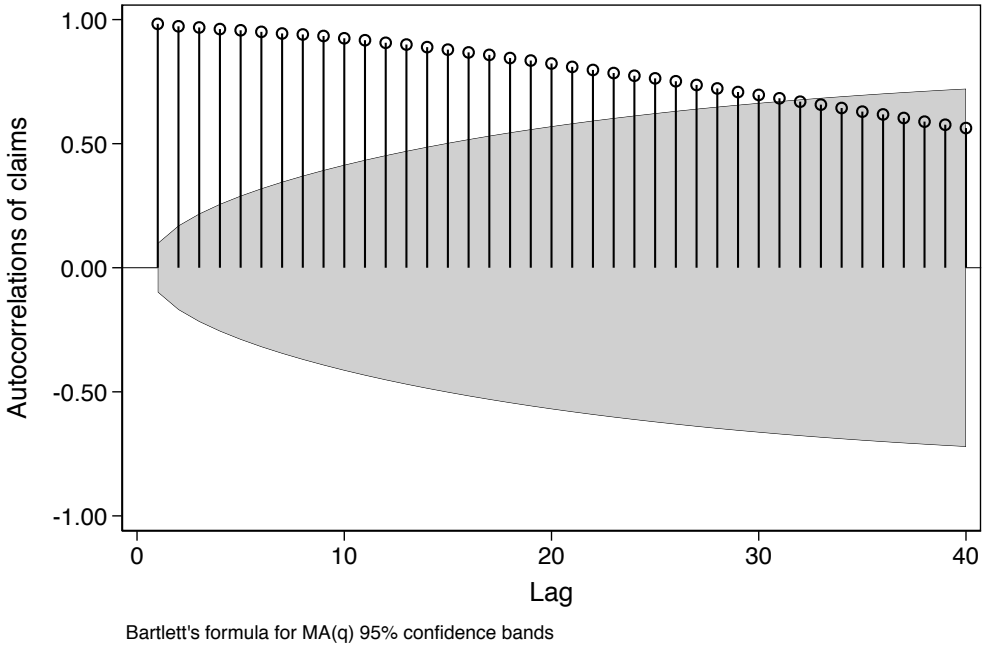

Figure 1: ACF of Initial Claims for Unemployment Benefits (claims)

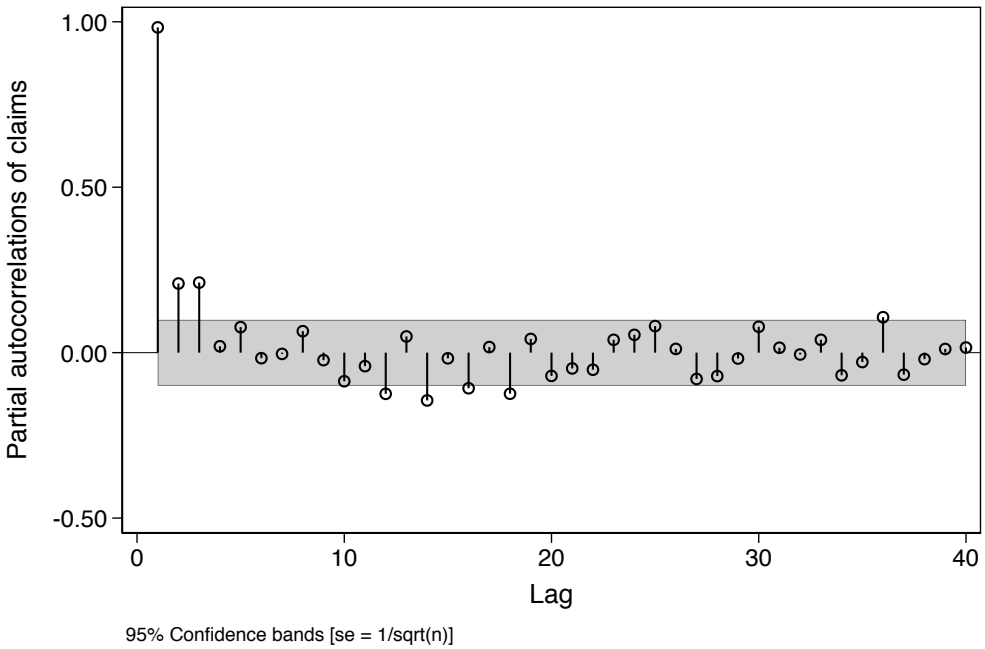

Figure 2: PACF of Initial Claims for Unemployment Benefits (claims)

```
use claims.dta
tsset time
*****************************************************
**  Identification of the process
*****************************************************
ac claims
pac claims
corrgram claims, lags(10)
*****************************************************
**  Estimation
*****************************************************
arima claims, ar(1) nolog
estat ic
arima claims jobs welfareunemployment, ar(1) nolog
estat ic


*****************************************************
**  Forecasting
*****************************************************
quiet: arima claims if t<=200, ar(1) nolog
predict forecast_ar, y t0(201)

quiet: arima claims jobs welfareunemployment if t<=200, ar(1) nolog
predict forecast_ar_google, y t0(201)

gen squared_perror_google = (claims-forecast_ar_google)^2
gen squared_perror_ar = (claims-forecast_ar)^2
su squared_perror_ar if t>=201
local mspe_ar= r(mean)
su squared_perror_google if t>=201
local mspe_google= r(mean)

display (`mspe_google'-`mspe_ar')/`mspe_google'

```

```
 1 . ***********************************************
 2 .
 3 . ** Log-file in-class test 1
 4 .
 5 . ***********************************************
 6 . clear all

 7 . use claims.dta

 8 . tsset time
         time variable:  time, 1 to 391
                 delta:  1 unit

 9 . ***********************************************
10 .
11 . **  Identification of the process
12 .
13 . ***********************************************
14 . ac claims

15 . pac claims

16 . corrgram claims, lags(10)
```

LAG	AC	PAC	Q	Prob>Q	-1 0 1 [Autocorrelation]	-1 0 1 [Partial Autocor]
1	**0.9829**	**0.9831**	**380.62**	**0.0000**		
2	**0.9730**	**0.2091**	**754.59**	**0.0000**		
3	**0.9684**	**0.2119**	**1126**	**0.0000**		
4	**0.9619**	**0.0192**	**1493.3**	**0.0000**		
5	**0.9570**	**0.0772**	**1857.9**	**0.0000**		
6	**0.9507**	**-0.0167**	**2218.6**	**0.0000**		
7	**0.9441**	**-0.0038**	**2575.3**	**0.0000**		
8	**0.9401**	**0.0649**	**2929.9**	**0.0000**		
9	**0.9340**	**-0.0226**	**3280.8**	**0.0000**		
10	**0.9248**	**-0.0864**	**3625.7**	**0.0000**		

```
17 . ***********************************************
18 .
19 . **  Estimation
20 .
21 . ***********************************************
22 . arima claims, ar(1) nolog

ARIMA regression

Sample:  1 - 391                                Number of obs      =        391
                                                Wald chi2(1)       =   12104.56
Log likelihood = -4367.683                      Prob > chi2        =     0.0000
```

claims	Coef.	OPG Std. Err.	z	P>|z|	[95% Conf.	Interval]
claims						
_cons	**394652.8**	**48114.01**	**8.20**	**0.000**	**300351.1**	**488954.5**
ARMA						
ar L1.	**.9813742**	**.0089199**	**110.02**	**0.000**	**.9638915**	**.9988569**
/sigma	**17108.4**	**405.1303**	**42.23**	**0.000**	**16314.36**	**17902.44**

Note: The test of the variance against zero is one sided, and the two-sided
confidence interval is truncated at zero.

```
23 . estat ic
```

Akaike's information criterion and Bayesian information criterion

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	**391**	.	**-4367.683**	**3**	**8741.366**	**8753.272**

Note: N=Obs used in calculating BIC; see **[R] BIC note**.

```
24 . arima claims jobs welfareunemployment, ar(1) nolog
```

ARIMA regression

Sample: **1 - 391** Number of obs = **391**
Wald chi2(**3**) = **11586.71**
Log likelihood = **-4357.718** Prob > chi2 = **0.0000**

claims	Coef.	OPG Std. Err.	z	P>|z|	[95% Conf.	Interval]
claims						
jobs	**679.9971**	**155.6045**	**4.37**	**0.000**	**375.0177**	**984.9764**
welfareunemployment	**15.98935**	**106.7073**	**0.15**	**0.881**	**-193.1531**	**225.1318**
_cons	**398815.5**	**44874.96**	**8.89**	**0.000**	**310862.2**	**486768.8**
ARMA						
ar L1.	**.980761**	**.0091782**	**106.86**	**0.000**	**.962772**	**.9987501**
/sigma	**16680.45**	**410.943**	**40.59**	**0.000**	**15875.01**	**17485.88**

Note: The test of the variance against zero is one sided, and the two-sided
confidence interval is truncated at zero.

```
25 . estat ic
```

Akaike's information criterion and Bayesian information criterion

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	**391**	.	**-4357.718**	**5**	**8725.437**	**8745.281**

Note: N=Obs used in calculating BIC; see **[R] BIC note**.

```
26 .
27 .
```

```
28 . ***********************************************
29 .
30 . **  Forecasting
31 .
32 . ***********************************************
33 . quiet: arima claims if t<=200, ar(1) nolog

34 . predict forecast_ar, y t0(201)
   (200 missing values generated)

35 .
36 . quiet: arima claims jobs welfareunemployment if t<=200, ar(1) nolog

37 . predict forecast_ar_google, y t0(201)
   (200 missing values generated)

38 .
39 . gen squared_perror_google = (claims-forecast_ar_google)^2
   (200 missing values generated)

40 . gen squared_perror_ar = (claims-forecast_ar)^2
   (200 missing values generated)

41 . su squared_perror_ar if t>=201
```

Variable	Obs	Mean	Std. Dev.	Min	Max
squared_pe~r	**191**	**2.52e+09**	**2.90e+09**	**312201.6**	**1.60e+10**

```
42 . local mspe_ar= r(mean)

43 . su squared_perror_google if t>=201
```

Variable	Obs	Mean	Std. Dev.	Min	Max
squared_pe~e	**191**	**2.20e+09**	**2.76e+09**	**829408.6**	**1.47e+10**

```
44 . local mspe_google= r(mean)

45 .
46 . display (`mspe_google'-`mspe_ar')/`mspe_google'
   -.14430604

47 .
48 . log close
```