Solutions to Midterm
Financial Econometrics, Econ 40357
University of Notre Dame
Prof. Mark
Monday 28 September 2020

A **review** of the **rules**: Test is open book, open note, open internet, but not open communication with any other people. Any such communication will be considered cheating. Do not cheat! Submit via email, a pdf of your own work by 10 p.m. tonight. Anything coming in after the deadline will lose points.

1. (10 points) Let r_t be the rate of return on the S&P 500 index. How would you test the hypothesis that r_t is normally distributed?

   `Use the Jarque-Bera test. For sample size` T,`the test statistic is,`

   $$JB = \frac{T}{6}\left(sk_T^2 + \frac{(\kappa_T - 3)^2}{4}\right)$$

   `where` sk_T `is sample skewness,` κ_T `is sample kurtosis of` r_t. `The JB statistic has a` χ_2^2 `distribution under the null hypothesis of normality.`

2. (10 points) Let p_t be the log dividend-adjusted price of the Bankok Chain Hospital Company stock (listed on the Thai stock exchange), where $p_t = p_{t-1} + \epsilon_t$ where ϵ_t is i.i.d. What does stationarity mean in the time-series context, and why is p_t not stationary?

   `Stationarity means the mean and variance of` p_t `are finite (they exist), and the k-th order covarariance` $Cov\left(p_t, p_{t-k}\right)$ `is constant and depends only on` k. `The impulse response to a shock` ϵ_t `should be transitory.`

   p_t `here is not stationary because the variance does not exist. The impulse response here, is permanent.`

3. (10 points) For the model in question 2, what is the optimal predictor (forecast formula) of the 20 period ahead return $p_{t+20} - p_t$?

   $E_t\left(p_{t+20} - p_t\right) = 0$

4. (10 points) Let $x_t = x_{t-1} + u_t$, describe the evolution of the Yoder family farm's tomato crop. The Yoder farm is located in Lakeville IN. Charles runs the regression

   $$p_{t+1} = \beta_0 + \beta_1 x_t + v_{t+1},$$

   where p_t is the stock price of the Bankok Chain Hospital Company, from question 2 above. Charles obtains $\hat{\beta}_1 = 4.35$, $t-$ratio=6.324. Can Charles conclude that the Yoder's tomato output can predict the future price of the Bankok Chain's stock price?

   `No. Charles has encountered the spurious regression problem.` p_t `and` x_t `are independent driftless random walks. The t-ratio will (almost) always be larger than 2.0`

5. (10 points) What is Newey-West and why do we use it?

   `Newey-West is how to compute the standard error of regression coefficients when the error term is conditionally heteroskedastic and serially correlated. We use it in financial econometrics, because financial data are almost always heteroskedastic.`

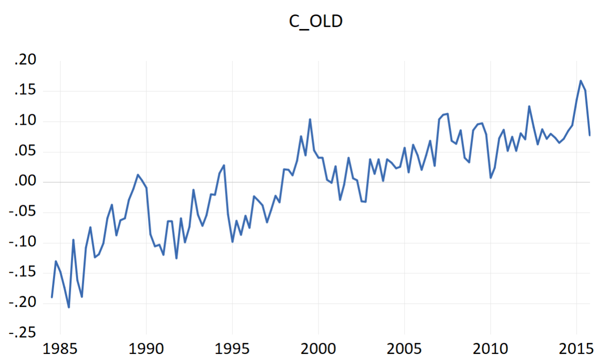
For Questions 6-10, use the Eviews workfile Midterm2020.wf1. c_old is the log of consumption of old households (head of household aged 65 and older). Shock is a monetary policy shock constructed from changes in the price of Federal Funds futures within a 30 minute window of the Federal Reserve's press conference following FOMC meetings. Interpret an increase in shock as a surprise **tightening** of monetary policy–that is, an increase in the Federal Funds interest rate. We want to see hold old people's consumption respond to a monetary policy shock (tightening).

6. (10 points) Why do we want to analyze log consumption instead of consumption (in levels)?

   Consumption tends to grow over time at the rate of real GDP growth. Plots of log consumption will approximately be linear, and the change in log consumption is approximately the growth rate.

7. (10 points) We want to run a VAR using old consumption and shock. Should we use c_old or Δc_old in the VAR? (provide explanation).

   We may want to use Δc_old, because c_old trends up.



   Also, the ADF test with lag selected by AIC or Hannan-Quin has a p-value of 0.1563, which cannot reject the unit root hypothesis. However, the ADF with lag selected by BIC has p-value 0.0246 which does reject the unit root. If you relied on BIC, you would be running the VAR on c_old instead of Δc_old.

8. (10 points) Based on your answer in 7, consider VARs of order 2, 4, 6, and 8. What specification does BIC (Schwarz) choose?

   For c_old and shock,

Lag	BIC
2	-1.404495
4	-1.148021
6	-0.779133
8	-0.554489

   BIC chooses 2 lags.

   For Δc_old and shocks,

Lag	BIC
2	-1.405465
4	-1.134112
6	-0.860888
8	-0.544746

   and BIC again chooses 2 lags.

9. (10 points) Run the VAR(p) with your chosen value of p, and generate the impulse response of log consumption of old people to a monetary policy shock. Ask for 16 periods in the impulse response.

(a) Show a screen shot of the VAR specification

```
If you ran the VAR with consumption in log levels,
```

Var: UNTITLED   Workfile: MIDTERM2020::Midterm\

w | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Impulse | Resids

Vector Autoregression Estimates

ector Autoregression Estimates
ate: 09/28/20   Time: 11:22
ample (adjusted): 1990Q3 2009Q4
cluded observations: 78 after adjustments
tandard errors in ( ) & t-statistics in [ ]

	C_OLD	SHOCK
C_OLD(-1)	0.661110	3.076867
	(0.11224)	(2.54959)
	[5.89032]	[1.20681]
C_OLD(-2)	0.234712	-1.460480
	(0.11478)	(2.60738)
	[2.04487]	[-0.56013]
SHOCK(-1)	-0.005374	0.068722
	(0.00524)	(0.11904)
	[-1.02549]	[0.57731]
SHOCK(-2)	-0.000862	0.100235
	(0.00528)	(0.11989)
	[-0.16340]	[0.83606]
C	0.000904	-0.229409
	(0.00416)	(0.09451)
	[0.21735]	[-2.42732]
-squared	0.718785	0.051463
dj. R-squared	0.703376	-0.000511
um sq. resids	0.077891	40.19385
.E. equation	0.032665	0.742025
-statistic	46.64695	0.990160
og likelihood	158.7798	-84.82041
kaike AIC	-3.943071	2.303087
chwarz SC	-3.792000	2.454159

```
If you ran the VAR with consumption growth rates,
```

Var: UNTITLED   Workfile: MIDTERM2020::Midterm\

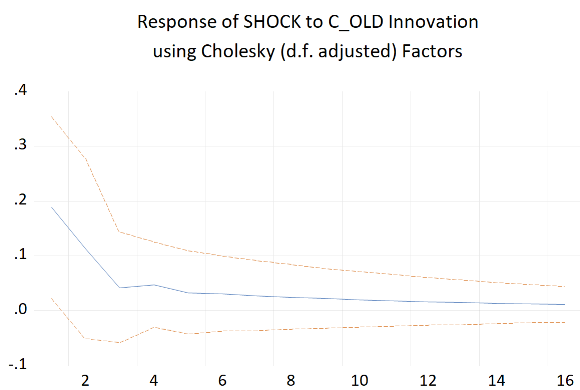
iew | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Impulse | Resids

Vector Autoregression Estimates

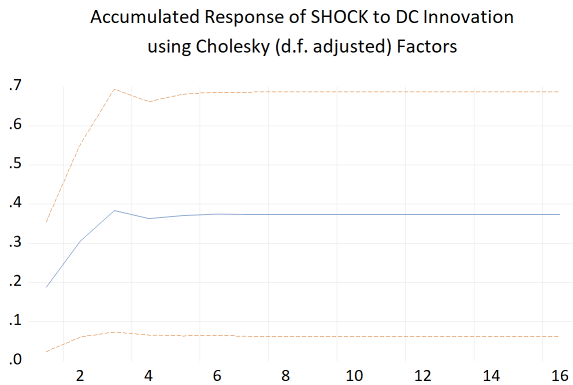
Vector Autoregression Estimates
Date: 09/28/20   Time: 11:24
Sample (adjusted): 1990Q3 2009Q4
Included observations: 78 after adjustments
Standard errors in ( ) & t-statistics in [ ]

	DC	SHOCK
DC(-1)	-0.344277	3.150127
	(0.11307)	(2.56950)
	[-3.04490]	[1.22597]
DC(-2)	-0.170518	2.611210
	(0.11059)	(2.51322)
	[-1.54189]	[1.03899]
SHOCK(-1)	-0.005908	0.077230
	(0.00518)	(0.11764)
	[-1.14130]	[0.65651]
SHOCK(-2)	-0.001865	0.115893
	(0.00519)	(0.11800)
	[-0.35918]	[0.98218]
C	0.000692	-0.225977
	(0.00414)	(0.09409)
	[0.16719]	[-2.40165]
R-squared	0.152822	0.051444
Adj. R-squared	0.106401	-0.000532
Sum sq. resids	0.077829	40.19469
S.E. equation	0.032652	0.742032
F-statistic	3.292110	0.989761
Log likelihood	158.8109	-84.82122
Akaike AIC	-3.943870	2.303108
Schwarz SC	[illegible]	[illegible]

(b) Show a screen shot of the impulse response.
If you ran with consumption in log levels,



And for consumption growth rates, the accumulated response looks like this



10. (10 points) Generate the impulse response for log consumption of old people to a monetary policy shock at horizons 4,8, and 12 using local projections. Compare your results from 7 and 8. Show a screen shot of the local projection at horizon 12.

    `Let` c_t `be log consumption of the old, and` s_t `be the monetary policy shock. If run on consumption, the regression would be`

$$c_{t+p} = \alpha + \beta s_t + \epsilon_{t+p}$$

`For` $p = 4, 8, 12,$ `the slope and t-ratio on` β `are,`

$\hat{\beta}$	t-ratio
0.006051	0.679820
0.005145	0.621921
-0.000668	-0.082250

`The local projection counterpart to the VAR for` Δc `snd` s_t `is`

$\hat{\beta}$	t-ratio
-0.015447	-2.084794
-0.016352	-2.137025
-0.022165	-2.633630

`Comparing the results: There's something fishy about the VARs. How does a surprise increase in the interest rate cause consumption to increase? Same with the local on the log levels. Only the last set of results make sense--that a tightenening of monetary policy causes consumption to fall.`

`So what is wrong with the VAR? Probably an insufficient number of lags.`

11. (10 points extra credit). Consider the VAR(1) for variables y_t and x_t. Writing the system explicitly, we have

$$y_t = ay_{t-1} + bx_{t-1} + \epsilon_{yt}$$
$$x_t = cy_{t-1} + dx_{t-1} + \epsilon_{xt}$$

where $\epsilon_{y,}$ and ϵ_{xt} are zero-meaned, serially uncorrelated shocks but contemporaneously correlated with covarianc σ_{xy} =E$(\epsilon_{yt}\epsilon_{xt}) \neq 0$.What is the optimal predictor (forecasting formula) for y_{t+2}, conditional on information known at t.

Advance the time subscript by two periods on the first equation, then take expectations conditional on information known at time t

$$\begin{aligned} y_{t+2} &= ay_{t+1} + bx_{t+1} + \epsilon_{yt+2} \\ E_t\left(y_{t+2}\right) &= aE_t\left(y_{t+1}\right) + bE_t\left(x_{t+1}\right) \qquad (1) \end{aligned}$$

where

$$E_t\left(y_{t+1}\right) = ay_t + bx_t \qquad (2)$$

$$E_t\left(x_{t+1}\right) = cy_t + dx_t \qquad (3)$$

substitute (2) and (3) into (1) to get

$$E_t\left(y_{t+2}\right) = \left(a^2 + bc\right) y_t + \left(ab + bd\right) x_t$$