

Chapter 11: Regression with Time Series Variables with Several Equations

Financial research often involves regression-type models with more than one equation.

This chapter motivates why such models are used in finance and how they can be analyzed.

Most of the models/ideas in this chapter are straightforward extensions of the ADL and ECM models of Chapter 10.

Topics covered:

- 1. Granger causality**
- 2. Vector Autoregressive (VAR) model.**
- 3. Vector Error Correction Model (VECM)**
- 4. Johansen Test for cointegration**
- 5. Forecasting**
- 6. Informal introduction to variance decompositions**

Granger Causality

In Chapter 3, we said you should be cautious about interpreting correlation and regression results as reflecting causality.

In regression, we label one variable the dependent variable and the others the explanatory variables.

In many cases, because the latter “explained” the former it was reasonable to talk about X “causing” Y.

In house price example, the price of the house was said to be “caused” by the characteristics of the house (e.g. number of bedrooms, number of bathrooms, etc.).

But in many regressions it is not obvious which variable causes which.

e.g. if you have $Y =$ stock prices in country A on $X =$ stock prices in Country B. which causes which?

Granger Causality (cont.)

With time series data we can make slightly stronger statements about causality simply by exploiting the fact that time does not run backward.

If event A happens before event B, then it is possible that A is causing B. However, it is not possible that B is causing A.

These ideas can be investigated through regression models using the notion of Granger causality.

X “Granger causes” Y if past values of X can help explain Y.

If Granger causality holds this does not *guarantee* that X causes Y. But, it suggests that *X might be* causing Y.

Granger Causality (cont.)

Granger causality is only relevant with time series variables.

Work with Granger causality between two variables (X and Y) which are both stationary. Can be extended to many variables.

Non-stationary case, where X and Y have unit roots but are cointegrated, will be mentioned later on.

Of course, in practice you must do Dickey-Fuller tests to see if your variables are stationary or not.

Granger Causality in a Simple ADL Model

Since X and Y are assumed to be stationary, use the ADL model.

Begin with simple ADL(1,1) model:

$$Y_t = \alpha + \phi_1 Y_{t-1} + \beta_1 X_{t-1} + e_t.$$

β_1 is a measure of the influence of X_{t-1} on Y_t .

If $\beta_1=0$ then X does not Granger cause Y.

“if $\beta_1=0$ then past values of X have no explanatory power for Y beyond that provided by past values of Y”.

Granger Causality in a Simple ADL Model (cont.)

Granger causality test uses methods for ADL from Chapter 10.

OLS estimation can be done and the P-value for the coefficient on X_{t-1} examined for significance.

If β_1 is statistically significant (e.g. P-value < .05) then we conclude that X Granger causes Y.

Granger Causality in an ADL Model with p and q Lags

$$Y_t = \alpha + \delta t + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} \\ + \beta_1 X_{t-1} + \dots + \beta_q X_{t-q} + e_t.$$

X Granger causes Y if *any or all of* β_1, \dots, β_q are statistically significant.

Since we are assuming X and Y do not contain unit roots, OLS regression analysis can be used to estimate this model and standard hypothesis testing procedures work.

Granger Causality in an ADL Model with p and q Lags (cont.)

The proper way to do Granger causality testing is to test the hypothesis that $\beta_1=\beta_2=\dots=\beta_q=0$

X Granger causes Y only if the hypothesis is rejected.

Note that the joint test of $\beta_1=\beta_2=\dots=\beta_q=0$ is not exactly the same as q individual tests of $\beta_i=0$ for $i=1,\dots,q$.

Appendix to Chapter 11 describes how joint test can be done using F-test

You should already know F-tests from previous studies, but a quick reminder cannot hurt

Appendix A to Chapter 11: Hypothesis Tests Involving More than One Coefficient

Let us briefly switch back to the standard regression model to remind you of basic theory.

Chapters 5 and 6 introduced F-stat for testing $H_0: \beta_1 = \dots = \beta_k = 0$ in:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + e.$$

t-stats can be used to test $H_0: \beta_j = 0$

What about other cases?

e.g. in the case $k=4$, $H_0: \beta_1 = \beta_3 = 0$

Unrestricted and Restricted models

Most hypotheses you would want to test place restrictions on the model.

E.g. unrestricted regression model is:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + e$$

and you wish to test the hypothesis $H_0: \beta_2 = \beta_4 = 0$, then restricted regression model is:

$$Y = \alpha + \beta_1 X_1 + \beta_3 X_3 + e.$$

Remember: strategy of hypothesis testing is that a test statistic is first calculated and then compared to a critical value.

If test statistic is greater than critical value then reject the hypothesis; otherwise, accept it.

Here the test statistic is F-statistic:

$$f = \frac{(R_U^2 - R_R^2) / J}{R_U^2 / (T - k)},$$

R_U^2 is R^2 from unrestricted regression model

R_R^2 is R^2 's from restricted regression model

J is the number of restrictions

(e.g. J=2 in example since $\beta_2=0$ and $\beta_4=0$ are two restrictions).

T is the number of observations

k is number of explanatory variables in the unrestricted regression (including the intercept)

Most relevant computer packages will calculate the F-statistic automatically if you specify the hypothesis being tested.

Or F-statistic can be obtained by running the unrestricted and restricted regressions and evaluating its formula

Obtaining critical value with which to compare the F-statistic is more problematic (although some software packages will provide a P-value automatically).

Formally, the critical value depends on $T-k$ and J .

Most econometrics or statistics textbooks will contain statistical tables for the F-distribution which will provide the relevant critical values.

Granger Causality in an ADL Model with p and q Lags (cont.)

Remember:

$$Y_t = \alpha + \delta t + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} \\ + \beta_1 X_{t-1} + \dots + \beta_q X_{t-q} + e_t.$$

So Granger causality involves test of \mathbf{H}_0 :
 $\beta_1 = \dots = \beta_q = 0$ in this ADL

Sometimes researchers check for Granger causality simply (albeit imperfectly) using only t-tests.

The P-values for the t-states on individual coefficients can be used to determine whether Granger causality is present.

Using the 5% level of significance, then if *any* of the P-values for the coefficients β_1, \dots, β_q were less than .05, you would conclude that Granger causality is present.

If *none* of the P-values is less than .05 then you would conclude that Granger causality is not present.

Granger Causality in an ADL Model with p and q Lags (cont.)

Warning if you use t-tests:

If any or all of the coefficients β_1, \dots, β_q are significant using t-tests, you may safely conclude that X Granger causes Y.

If none of these coefficients is significant, it is probably the case that X does not Granger cause Y. However, you are more likely to be wrong if you conclude the latter than if you had used the correct joint test of Granger non-causality.

Example: Do Stock Price Movements in Country B Granger Cause Stock Price Movements in Country A?

We have monthly data on logged stock prices for Countries A and B.

Dickey-Fuller and Engle-Granger tests indicate that stock prices in both countries appear to have unit roots, but are not cointegrated.

However, differences of these series are stationary and can be interpreted as stock market returns (exclusive of dividends).

Use differenced variables to investigate whether stock returns in Country A Granger cause those in Country B.

Example: Do Stock Price Movements in Country B Granger Cause Stock Price Movements in Country A? (continued)

ADL Model using Stock Returns in Country A as the Dependent Variable

	Coeff	t Stat	P-value	Lower 95%	Upper 95%
Intercept	-.751	-1.058	.292	-2.156	.654
ΔY_{t-1}	.822	4.850	3.81E-6	.486	1.158
ΔY_{t-2}	-.041	-.222	.825	-.409	.326
ΔY_{t-3}	.142	.762	.448	-.227	.511
ΔY_{t-4}	-.181	-1.035	.303	-.526	.165
ΔX_{t-1}	-.016	-.114	.909	-.299	.267
ΔX_{t-2}	-.118	-.823	.412	-.402	.166
ΔX_{t-3}	-.042	-.292	.771	-.324	.241
ΔX_{t-4}	.038	.266	.791	-.244	.319
Time	.030	2.669	.009	.0077	.052

Example: Do Stock Price Movements in Country B Granger Cause Stock Price Movements in Country A? (continued)

P-values indicate that only the deterministic trend and last period's stock returns in Country A have explanatory power for present stock returns in Country A.

All of the coefficients on the lags of stock returns in Country B are insignificant.

Stock returns in Country B do not seem to Granger cause stock returns in Country A.

Example: Do Stock Returns in Country B Granger Cause Stock Returns in Country A? (continued)

Se found that stock returns in Country B did not Granger cause stock returns in Country A using t-tests.

Here, we will investigate whether these conclusions still hold by carrying out the correct F-tests for Granger causality.

Y = stock returns in Country A

X = stock returns in Country B

Do stock returns in Country B Granger cause stock returns in Country A?

Run following regression:

$$Y_t = \alpha + \delta + \phi_1 Y_{t-1} + \dots + \phi_4 Y_{t-4} + \beta_1 X_{t-1} + \dots + \beta_4 X_{t-4} + e_t.$$

We have T=128 and k=10 (i.e. p=q=4 plus deterministic trend).

OLS estimation of this model yields $R_U^2 = .616$.

The hypothesis that Granger causality does not occur is $H_0: \beta_1 = \dots = \beta_4 = 0$ which involves 4 restrictions; hence J=4.

The restricted regression model is:

$$Y_t = \alpha + \delta t + \phi_1 Y_{t-1} + \dots + \phi_4 Y_{t-4} + e_t.$$

OLS estimation of this model yields $R_R^2 = .613$.

Result: F-statistic is .145.

Since $T-k=118$ and is large, we can compare .145 to a critical value of 2.37.

Since $.145 < 2.37$ we cannot reject the hypothesis at the 5% level of significance.

Accordingly, we accept the hypothesis that stock returns in Country B do not Granger cause stock returns in Country A.

Causality in Both Directions

In many cases, it is not obvious which way causality should run.

Should stock markets in Country A affect markets in Country B or should the reverse hold?

Can check this by running two regressions:

one with Y being the dependent variable and one with X being the dependent variable.

If you have k variables, then run k regressions.

This is an example of how multiple equation models can arise.

Example: Do Stock Price Movements in Country A Granger Cause Stock Price Movements in Country B? (continued)

Before we ran a regression to see whether stock returns in Country B Granger caused stock returns in Country A.

To test whether the causality runs in the opposite direction run the ADL regression with ΔX = stock returns in Country B being dependent variable.

ADL Model using Stock Returns in Country B as the Dependent Variable

	Coeff	t Stat	P-value	Lower 95%	Upper 95%
Intercept	-.609	-.730	.467	-2.262	1.044
ΔX_{t-1}	.053	.312	.755	-.280	.386
ΔX_{t-2}	-.040	-.235	.814	-.374	.294
ΔX_{t-3}	-.058	-.348	.728	-.391	.274
ΔX_{t-4}	.036	.215	.830	-.295	.367
ΔY_{t-1}	.854	4.280	3.83E-5	.459	1.249
ΔY_{t-2}	-.217	-.993	.323	-.649	.215
ΔY_{t-3}	.234	1.067	.288	-.200	.668
ΔY_{t-4}	-.272	-1.323	.188	-.678	.135
Time	.046	3.514	.001	.020	.072

Here we do find evidence that stock returns in Country A Granger cause stock returns in Country B.

In particular, the coefficient on ΔY_{t-1} is highly significant, indicating that last month's stock returns in Country A has strong explanatory power for stock returns in Country B.

Do stock returns in Country A Granger cause stock returns in Country B?

Now test for Granger causality using the F-test

Repeat steps for F-test as above except now dependent variable refers to Country B and the explanatory variable refers to Country A.

We obtain $R_U^2 = .605$ and $R_R^2 = .532$

The F-statistic is 33.412, which is much larger than either the 1% or 5% critical values.

Reject the hypothesis that $\beta_1 = \dots = \beta_4 = 0$ and conclude that stock returns in Country A do Granger cause stock returns in Country B.

Granger Causality with Cointegrated Variables

Testing for Granger causality among cointegrated variables is very similar to the method outlined above.

Remember that, if variables are found to be cointegrate, the an error correction model (ECM) can be used:

$$\Delta Y_t = \varphi + \delta + \lambda e_{t-1} + \gamma_1 \Delta Y_{t-1} + \dots + \gamma_p \Delta Y_{t-p} \\ + \omega_1 \Delta X_{t-1} + \dots + \omega_q \Delta X_{t-q} + \varepsilon_t.$$

Granger Causality with Cointegrated Variables (continued)

Remember: this is almost an ADL model except for the presence of the term λe_{t-1} .

Remember: $e_{t-1} = Y_{t-1} - \alpha - \beta X_{t-1}$, an estimate of which can be obtained by running a regression of Y on X and saving the residuals.

Remember: Many computer packages will estimate ECMs for you or you can estimate them with the 2-step approach described in Chapter 10

Granger Causality with Cointegrated Variables (continued)

X Granger causes Y if past values of X have explanatory power for current values of Y.

In ECM, past values of X appear in the terms $\Delta X_{t-1}, \dots, \Delta X_{t-q}$ and e_{t-1} .

X does not Granger cause Y if $\omega_1 = \dots = \omega_q = \lambda = 0$.

Many computer packages will test this hypothesis for you (and provide a p-value)

Or, if you are using 2-step estimation method, t-statistics and P-values can be used to test for Granger causality in the same way as the stationary case.

F-tests described in Appendix can be used to carry out a formal test of $H_0: \omega_1 = \dots = \omega_q = \lambda = 0$.

Testing whether Y Granger causes X is achieved by reversing roles that X and Y play in ECM.

There is a theorem that implies if X and Y are cointegrated then some form of Granger causality must occur: either X must Granger cause Y or Y must Granger cause X (or both).

Vector Autoregressions

Granger causality often tested using Vector Autoregressions or VARs.

First we will define VARs assuming that all variables are stationary.

With 2 variables, a VAR involves 2 equations:

$$Y_t = \alpha_1 + \delta_1 t + \phi_{11} Y_{t-1} + \dots + \phi_{1p} Y_{t-p} \\ + \beta_{11} X_{t-1} + \dots + \beta_{1q} X_{t-q} + e_{1t},$$

and

$$X_t = \alpha_2 + \delta_2 t + \phi_{21} Y_{t-1} + \dots + \phi_{2p} Y_{t-p} \\ + \beta_{21} X_{t-1} + \dots + \beta_{2q} X_{t-q} + e_{2t}.$$

VARs (continued)

Before, we use first of these equations to test whether X Granger causes Y

The second, whether Y Granger causes X.

Note that now the coefficients have subscripts indicating which equation they are in.

These two equations comprise a VAR.

A VAR is the extension of the autoregressive (AR) model to the case in which there is more than one variable under study.

In general, a k-variable VAR has k equations (one use each variable as the dependent variable)

Each equation uses as its explanatory variables lags of *all the variables under study* (and possibly a deterministic trend).

VARs (continued)

Note that lag lengths, p and q , can be selected using the sequential testing methods discussed in Chapters 8 through 10.

However, common to set $p=q$ and use the same lag length for every variable in every equation.

Result is VAR(p) model.

VARs (continued)

Example: the following VAR(p) has three variables, Y, X and Z:

$$Y_t = \alpha_1 + \delta_1 t + \phi_{11} Y_{t-1} + \dots + \phi_{1p} Y_{t-p} + \beta_{11} X_{t-1} + \dots + \beta_{1p} X_{t-p} + \delta_{11} Z_{t-1} + \dots + \delta_{1p} Z_{t-p} + e_{1t},$$

$$X_t = \alpha_2 + \delta_2 t + \phi_{21} Y_{t-1} + \dots + \phi_{2p} Y_{t-p} + \beta_{21} X_{t-1} + \dots + \beta_{2p} X_{t-p} + \delta_{21} Z_{t-1} + \dots + \delta_{2p} Z_{t-p} + e_{2t},$$

$$Z_t = \alpha_3 + \delta_3 t + \phi_{31} Y_{t-1} + \dots + \phi_{3p} Y_{t-p} + \beta_{31} X_{t-1} + \dots + \beta_{3p} X_{t-p} + \delta_{31} Z_{t-1} + \dots + \delta_{3p} Z_{t-p} + e_{3t}.$$

VARs (continued)

Most relevant computer packages allow you to estimate, test and forecast using VARs automatically.

Alternatively, estimation and testing can be done using OLS methods on one equation at a time.

Remember: we have assumed that all the variables in the VAR(p) are stationary.

Thus, you can obtain estimates of coefficients in each equation using OLS. P-values or t-statistics will then allow you to ascertain whether individual coefficients are significant.

You can also use the material covered in Appendix to carry out F-tests.

Why Use VARs?

Granger causality testing (see above)

Forecasting (see below)

Financial researchers also use VARs in many other contexts.

Examples of financial work which uses VARs:

- **Models with present value relationships often work with VARs using the (log) dividend-price ratio and dividend growth.**
- **Term structure of interest rates (using interest rates of various maturities, interest rate spreads, etc.)**
- **Intertemporal asset allocation (using returns on various risky assets)**
- **The rational valuation formula (using the dividend-price ratio and returns)**
- **The interaction of bond and equity markets (using stock and bond return data)**

Example: What Moves the Stock and Bond Markets?

“What moves the stock and bond markets? A variance decomposition for long-term asset returns” by Campbell and Ammer (Journal of Finance, 1991)

Paper investigates the factors which influenced the stock and bond markets in the long run.

Theoretical model has properties:

- 1. unexpected movements in excess stock returns depend on changes in expectations (i.e. news) about future dividend flows, future excess stock returns and future real interest rates.**
- 2. unexpected movements in excess bond returns depend on changes in expectations (i.e. news) about future inflation, future interest rates and future excess bond returns.**

Example: What Moves the Stock and Bond Markets? (continued)

Which of these various factors is most important in driving the stock and bond markets?

The authors conclude that news about future excess stock returns is most important factor in driving the stock market and news about future inflation is the most important factor in driving the bond market.

Key part of this model (and many similar models) is that the researcher has to distinguish between “expected” and “unexpected” values of variables.

Example: What Moves the Stock and Bond Markets? (continued)

er_t = excess return on stock market at time t .

Consider the investor at time $t-1$ trying to make investment decisions.

At time $t-1$, she will not know exactly what er_t will be, but will have some expectation about what it might be.

Expectation at time $t-1$ of what the excess stock return at time t will be is $E_{t-1}(er_t)$.

Unexpected movements in stock and bond markets are crucial to the underlying financial theory.

These are $er_t - E_{t-1}(er_t)$

General point: expectations such $E_{t-1}(er_t)$ appear in financial models.

Example: What Moves the Stock and Bond Markets? (continued)

VARs are frequently used to model expectations.

Note: the right-hand side of an equation in a VAR only contains variables dated $t-1$ or earlier, it can be thought of as reflecting information available to the investor at time $t-1$.

This reasoning suggests:

Use fitted value from equation where er_t is the dependent variable as an estimate of $E_{t-1}(er_t)$.

Any relevant computer package will allow for calculation of fitted values.

Example: What Moves the Stock and Bond Markets? (continued)

Campbell and Ammer use a VAR involving the following six variables:

- **er is the excess stock return.**
- **r is the real interest rate.**
- **dy is the change in the return on a short-term bond.**
- **s is the yield spread (difference in yields between a 10 year and a two month bond).**
- **dp is the log of the dividend-price ratio.**
- **rb is the relative bill rate (a return on a short term bond relative to the average returns over the last year)**

Example: What Moves the Stock and Bond Markets? (continued)

Monthly data from December 1947 through February 1987

Note: Campbell and Ammer did extensive testing to confirm that all of these variables are stationary.

Remember: before carrying out an analysis using time series data, you must conduct unit root tests.

Remember: if unit roots are present but cointegration does not occur, then the spurious regression problem exists. In this case, you should work with differenced data.

Remember: if unit roots exist and cointegration does occur, then you will have important information that the series are trending together.

Example: What Moves the Stock and Bond Markets? (continued)

Following table presents results from estimation of a VAR(1). There are six variables in our VAR (i.e. er , r , dy , s , dp and rb), so are six equations to estimate.

Each column of table contains results for one equation and lists OLS estimates (P-values in parentheses)

	Dependent Variable					
	er_t	r_t	dy_t	s_t	dp_t	rb_t
Interc.	-1.593 (0.053)	.678 (.354)	.116 (.362)	.066 (.562)	-.007 (.635)	.082 (.516)
er_{t-1}	-.018 (.696)	-.099 (.041)	.013 (.064)	-.004 (.573)	-.043 (.000)	.014 (.042)
r_{t-1}	.033 (.466)	.473 (.000)	-.012 (.089)	.007 (.237)	-.0004 (.608)	-.011 (.104)
dy_{t-1}	-.640 (.056)	.416 (.161)	.067 (.196)	-.045 (.326)	.003 (.585)	.096 (.062)
s_{t-1}	.318 (.173)	.215 (.299)	.075 (.037)	.862 (.000)	.004 (.407)	.100 (.006)
dp_{t-1}	.425 (.012)	-.087 (.561)	-.048 (.066)	.026 (.261)	1.005 (.000)	-.049 (.061)
rb_{t-1}	-.357 (.174)	.064 (.783)	-.011 (.778)	-.017 (.643)	1.56 (.119)	.888 (.000)

Example: What Moves the Stock and Bond Markets? (continued)

Most variables are insignificant (it is often hard to predict financial variables)

Look at significant coefficients (i.e. those with P-value less than .05).

Last month's dividend-price ratio does have significant explanatory power for excess stock returns this month.

Last month's yield spread does have explanatory for the change in short term bond returns.

Example: What Moves the Stock and Bond Markets? (continued)

Could report the results from the VAR as shedding light on the inter-relationships between key financial variables.

However, could use results from VAR as first step in an analysis of what moves stock and bond markets (e.g. fitted values for constructing expected/unexpected values of variables).

For instance: “variance decomposition”

Informal discussion of variance decompositions given in Appendix B.

Campbell and Ammer paper, using variance decompositions to “attribute only 15% of the variance of stock returns to the variance of news about future dividends, and 70% to news about future excess returns”

Lag Length Selection in VARs

How to select p in the VAR(p)?

As with ADL's can use t-stats and F-stats

Can also use an *information criterion*.

We will not provide exact formulae. Most relevant computer packages will calculate several information criteria for VARs

Popular ones are:

Akaike's information criterion (AIC)

Schwarz-Bayes information criterion (SBIC)

Hannan-Quinn information criterion (HQIC).

How to use an information criteria?

**Calculate it for VAR(p) for $p=1, \dots, p_{\max}$
 p_{\max} = maximum possible lag length**

Select the lag length which yields the smallest value for your information criterion.

Example: What Moves the Stock and Bond Markets? (continued)

Information Criteria for VAR(p)

Lag Length	AIC	SBIC	HQIC
p=1	8.121	8.267	8.492
p=2	7.084	7.355	7.774
p=3	7.026	7.424	8.037
p=4	6.934	7.458	8.266

SBIC and HQIC select VAR(2)'s

AIC selects a VAR(4).

This is the kind of conflict which often occurs in empirical practice: one criterion (or hypothesis test) indicates one thing whereas another similar criterion indicates something else.

Forecasting with VARs

The field of forecasting is enormous, so we will do little more than introduce it here.

Most computer packages have forecasting facilities that are simple to use.

Once you have estimated a model (e.g. a VAR or an AR), you can forecast simply by adding an appropriate option to an estimation command.

Rather than deriving theoretical results, we will just provide some intuition and show how it works in practice.

Forecasting with VARs (cont.)

Use data for periods $t=1,\dots,T$ to forecast periods $T+1, T+2$, etc.

Consider VAR(1) with two variables, Y and X :

$$Y_t = \alpha_1 + \delta_1 t + \phi_{11} Y_{t-1} + \beta_{11} X_{t-1} + e_{1t},$$

and

$$X_t = \alpha_2 + \delta_2 t + \phi_{21} Y_{t-1} + \beta_{21} X_{t-1} + e_{2t}.$$

Forecasting with VARs (cont.)

Suppose you want to forecast Y_{T+1}

Using VAR and setting $t=T+1$:

$$Y_{T+1} = \alpha_1 + \delta_1(T+1) + \phi_{11}Y_T + \beta_{11}X_T + e_{1T+1}.$$

If we ignore the error term (which cannot be forecast since it is unpredictable) and replace the coefficients by their estimates we obtain a forecast which we denote as \hat{Y}_{T+1} :

$$\hat{Y}_{T+1} = \hat{\alpha}_1 + \hat{\delta}_1(T+1) + \hat{\phi}_{11}Y_T + \hat{\beta}_{11}X_T.$$

A similar strategy can be used to obtain \hat{X}_{T+1} .

Forecasting with VARs (cont.)

How to forecast Y_{T+2} ?

Before used X_T and Y_T to create \hat{Y}_{T+h}

But \hat{Y}_{T+2} dependz on Y_{T+1} and X_{T+1} .

But since our data only runs until period T , we do not know what Y_{T+1} and X_{T+1} are.

Solution: replace Y_{T+1} and X_{T+1} by \hat{Y}_{T+h} and \hat{X}_{T+h} .

$$\hat{Y}_{T+2} = \hat{\alpha}_1 + \hat{\delta}_1(T+2) + \hat{\phi}_{11}\hat{Y}_{T+1} + \hat{\beta}_{11}\hat{X}_{T+1}.$$

Can use same strategy to produce \hat{X}_{T+h} and same idea to produce \hat{Y}_{T+h} and \hat{X}_{T+h} for any h .

Forecasting with VARs (cont.)

Note: \hat{Y}_{T+h} and \hat{X}_{T+h} are point estimates.

Confidence intervals can also be calculated.

For instance, the Bank of England: “Our forecast of inflation next year is 1.8%. We are 95% confident that it will be between 1.45% and 2.15%”.

Vector Autoregressions with Cointegrated Variables

So far in the VAR discussion, we assumed that all variables are stationary.

If some of the original variables have unit roots and are not cointegrated, then the ones with unit roots should be differenced and the resulting stationary variables should be used in the VAR.

This covers every case except the one where the variables have unit roots and are cointegrated.

In this case, you should use a vector error correction model (VECM).

Like the VAR, the VECM will have one equation for each variable in the model, but each equation will be an error correction model.

VECMs

In the case of two variables, Y and X , which have unit roots and are cointegrated, the VECM is:

$$\begin{aligned}\Delta Y_t &= \varphi_1 + \delta_1 t + \lambda_1 e_{t-1} + \gamma_{11} \Delta Y_{t-1} + \dots \\ &+ \gamma_{1p} \Delta Y_{t-p} + \omega_{11} \Delta X_{t-1} + \dots + \omega_{1q} \Delta X_{t-q} + \varepsilon_{1t}\end{aligned}$$

and

$$\begin{aligned}\Delta X_t &= \varphi_2 + \delta_2 t + \lambda_2 e_{t-1} + \gamma_{21} \Delta Y_{t-1} + \dots \\ &+ \gamma_{2p} \Delta Y_{t-p} + \omega_{21} \Delta X_{t-1} + \dots + \omega_{2q} \Delta X_{t-q} + \varepsilon_{2t}.\end{aligned}$$

where $e_{t-1} = Y_{t-1} - \alpha - \beta X_{t-1}$.

VECMs

Computer packages allow you to estimate/test and forecast VECMs automatically.

Alternatively:

VECM is same as a VAR with differenced variables, except for the term e_{t-1} .

Estimate of e_{t-1} obtained by running an OLS regression of Y on X and saving the residuals.

Plugging in this estimate, can use OLS to estimate ECMs, and P-values and confidence intervals can be obtained.

Forecasting is like with VAR, with added complication that forecasts of the error correction term, e_t , must be calculated.

However, this is simple using OLS estimates of α and β and replacing e_t by residuals.

Johansen Test for Cointegration

Chapter 10 introduced Engle-Granger test for cointegration.

Now that we have introduced VECMs we can discuss (more popular) Johansen test

To explain this test would require a discussion of concepts beyond the scope of this book.

Many software packages do the Johansen test, so you may want to do it in practice.

Accordingly, we offer a brief intuitive description of this test.

Johansen Test (cont.)

More than one cointegrating relationship can exist with several unit root variables

With M variables, can have up to $M-1$ cointegrating relationships (and, thus, up to $M-1$ cointegrating residuals included in the VECM).

Example (from Chapter 10)

Lettau and Ludvigson paper "Consumption, aggregate wealth and expected stock returns"

Uses cointegrated variables (consumption, assets and income)

Dickey-Fuller tests say c , a and y all have unit roots

Financial theory suggests they are cointegrated.

One cointegrating relationship:

$c_t - \alpha - \beta_1 a_t - \beta_2 y_t$ is stationary.

Could have two cointegrating relationships.

e.g. if $c_t - y_t$ and $a_t - y_t$ were both stationary

Back to Johansen Test

Johansen test can be used to test for the number of cointegrating relationships using VECMs.

“number of cointegrating relationships” is referred to as the “cointegrating rank”.

Johansen test statistic is complicated.

However, like any hypothesis test, you can compare the test statistic to a critical value and, if the test statistic is greater than the critical value, you reject the hypothesis being tested.

Many software packages will calculate all these numbers for you.

We will see how this works in an example shortly.

Johansen Test (continued)

With VECMs (which are used in Johansen test) you have to specify the lag length and the deterministic trend term.

Lag length can be selected using information criteria as described above.

With VECMs can put an intercept and/or deterministic trend in the model (as we have done in the equations above – see the terms with coefficients φ and δ on them).

However, can put an intercept and/or deterministic trend in the cointegrating residual

E.g. if $c_t - \alpha - \beta_1 a_t - \beta_2 y_t$ is the cointegrating residual it has an intercept (but no deterministic trend)

Johansen test critical values depend on deterministic terms you use, so you will be asked to specify these before doing the Johansen test.

Example: Consumption, Aggregate Wealth and Expected Stock Returns

"Consumption, aggregate wealth and expected stock returns", Lettau and Ludvigson (Journal of Finance, 2001)

Financial theory says cay variables should be cointegrated and the cointegrating residual should be able to predict excess stock returns.

They then present empirical evidence in favor of their theory.

In "Understanding trend and cycle in asset values: Reevaluating the wealth effect on consumption" (American Economic Review, 2004), using the cay data, they build on this argument using VECMs and present variance decompositions which shed light on their theory.

Example: Consumption, Aggregate Wealth and Expected Stock Returns (continued)

Lettau and Ludvigson's work uses all the tools of this chapter: testing for cointegration, estimation of a VECM and variance decompositions.

We will investigate cointegration issues using U.S. data from 1951Q4 through 2003Q1 on cay variables

Unit root tests say cay variables have unit roots.

Do the Johansen test using a lag length of one and restricting the deterministic term to allow for intercepts only

Example: Consumption, Aggregate Wealth and Expected Stock Returns (continued)

Stata produces following table (other packages similar)

Rank	Trace Statistic	5% Critical value
0	37.27	29.68
1	6.93	15.41
2	.95	3.76

How should you interpret this table?

“Trace Statistic” is the Johansen test statistic

Rank = number of cointegrating relationships

Rank=0 implies cointegration is not present.

Example: Consumption, Aggregate Wealth and Expected Stock Returns (continued)

With Johansen test, hypothesis being tested is a certain cointegrating rank with the alternative hypothesis being that cointegrating rank is greater than hypothesis being tested.

For Rank = 0: Trace Statistic greater than critical value.

Reject the hypotheses that Rank=0 at the 5% level of significance (in favor of the hypothesis that Rank \geq 1)

For Rank =1: Trace statistic is less than the critical value.

Accept the hypothesis that Rank =1 (we are not finding evidence in favor of Rank \geq 2).

Overall: Johansen test finds one cointegrating relationship (same as Lettau and Ludvigson)

Appendix B: Variance Decompositions

To fully understand these would require difficult tools (e.g. matrix algebra).

However, some statistical software packages allow you to calculate variance decompositions in a straightforward manner.

Accordingly, with good software, some intuition and understanding of the financial problem, can do variance decompositions in practice.

Example “What moves the stock and bond markets?”

See body of chapter for reminder of this example.

Key idea in paper:

unexpected movements in excess stock returns should depend on changes in expectations about future dividend flows and future excess stock returns (among other things).

Key question:

which of these factors is most important in driving the stock markets?

Example “What moves the stock and bond markets?”

Campbell and Ammer’s model is much more sophisticated, but a simplified version has:

$$*uer* = *newsd* + *newser*$$

***uer* reflects unexpected movements in expected returns**

***newsd* reflects future news about dividends**

***newser* reflects future news about expected returns.**

Do not worry where these components come from other than to note that they can be calculated using the data and the VAR coefficients.

Example “What moves the stock and bond markets?”

What are relative roles played by *newsd* and *newser* in explaining *uer*?

To answer this calculate the proportion of the variability of *uer* that can be explained by *newsd* (or *newser*)

This is a simple example of a variance decomposition.

Example “What moves the stock and bond markets?”

Remember: variance is a measure of variability.

If variables independent can write:

$$\text{var}(uer) = \text{var}(newsd) + \text{var}(newser).$$

Divide this equation by $\text{var}(uer)$:

$$1 = \frac{\text{var}(newsd)}{\text{var}(uer)} + \frac{\text{var}(newser)}{\text{var}(uer)}.$$

Terms on the right-hand side are variance decompositions.

E.g.: “Proportion of variability in unexpected excess returns that can be explained by news about future dividends is $\frac{\text{var}(newsd)}{\text{var}(uer)}$ ”

These can be calculated with VARs.

"Consumption, aggregate wealth and expected stock returns"

Paper by Lettau and Ludvigson (Journal of Finance, 2001)

Lettau and Ludvigson example using cay data shows another sort of variance decomposition.

Why have huge swings in stock markets not had larger effects on consumption?

They estimate a VECM and calculate a variance decomposition.

Their story: many fluctuations in the stock market were treated by households as being transitory and these did not have large effects on their consumption.

Only permanent changes in wealth affected consumption.

Use “permanent-transitory decomposition”.

Permanent-Transitory Decomposition

Remember (see Chapter 9) that unit root variables are nonstationary.

Cointegrating error is stationary.

Nonstationary: errors have permanent effect

Stationary: error have transitory effect

VECM can be used to obtain these permanent and transitory components

Permanent-Transitory Decomposition (cont.)

A simplified version of Lettau-Ludvigson:

$$a = \textit{permanent} + \textit{transitory} ,$$

a = assets

permanent and *transitory* are the permanent and transitory components

Similar derivation as for Campbell-Ammer:

$$1 = \frac{\text{var}(\textit{permanent})}{\text{var}(a)} + \frac{\text{var}(\textit{transitory})}{\text{var}(a)} ,$$

$\frac{\text{var}(\textit{permanent})}{\text{var}(a)}$ is variance decomposition

“what proportion of the fluctuations in assets can be explained by permanent shocks”

Chapter Summary

- 1. X Granger causes Y if past values of X have explanatory power for Y.**
- 2. If X and Y are stationary, standard statistical methods based on an ADL model can be used to test for Granger causality.**
- 3. If X and Y have unit roots and are cointegrated, an ECM can be used to test for Granger causality.**
- 4. Vector autoregressions, or VARs, have one equation for each variable being studied. Each equation chooses one variable as the dependent variable. The explanatory variables are lags of all the variables under study.**
- 5. VARs are useful for forecasting, testing for Granger causality or, more generally, understanding the relationships between several series.**

- 6. If all the variables in the VAR are stationary, OLS can be used to estimate each equation and standard statistical methods can be employed (e.g. P-values and t-statistics can be used to test for significance of variables).**

- 7. If the variables under study have unit roots and are cointegrated, a variant on the VAR called the Vector error correction model, or VECM, should be used.**

- 8. The Johansen test is a popular test for cointegration included in many software packages.**