Regression with Time Series Variables

Motivation:

Financial researchers often interested in measuring the effect of an explanatory variable or variables on a dependent variable.

Regression modelling goal is complicated when the researcher uses time series data since an explanatory variable may influence a dependent variable with a time lag.

This often necessitates the inclusion of lags of the explanatory variable in the regression (Chapter 8).

Chapter 9 gives reasons why lags of the dependent variable should also be included in the regression.

Many theories in finance imply a regression model with lags of dependent and explanatory variables.

Here we do not derive financial theories to motivate our regression models. However, it is worth to mention a few classic articles in finance and the time series data sets they use.

Example 1:

Campbell and Ahmer "What moves the stock and bond markets? A variance decomposition for long-term asset returns"

American data on excess stock returns, various interest rates, the yield spread (difference between long- and short-term interest rates) and the dividend-price ratio.

Example 2:

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

Data: excess stock returns, asset wealth, labor income and consumption.

Example 3:

Campbell and Shiller, "Stock prices, earnings and expected dividends"

In Chapter 8 did example involving the prediction of long-run stock returns using the dividend-price ratio.

Such regressions grew out papers such as this one.

All of these examples (and many more) work with regressions with lags of dependent and explanatory variables

Autoregressive Distributed Lag (ADL) Model

- **Focus on 2 time series variables: Y_t and X_t.**

- **Can easily extend to many explanatory variables.**

- **Augment AR(p) with lags of explanatory variables produces ADL model**

$$Y_t = \alpha + \delta t + \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} +$$

$$\beta_0 X_t + \beta_1 X_{t-1} + \ldots + \beta_q X_{t-q} + e_t.$$

- **p lags of Y, q lags of X $\Rightarrow$ ADL(p,q).**

The ADL Model

- **Estimation and interpretation of the ADL(p,q) model depends on whether Y and X are stationary or have unit roots.**

- **<u>Before you estimate an ADL model you should test both Y and X for unit roots using the Dickey-Fuller test.</u>**

Time Series Regression when X and Y are Stationary

- **Minimal changes (e.g. OLS fine, testing done in standard way, etc.), except for the interpretation of results.**

- **Lag lengths, p and q can be selected using sequential tests as in Chapters 8 and 9.**

- **It is convenient to rewrite ADL model as:**

$$\Delta Y_t = \alpha + \delta t + \phi Y_{t-1} + \gamma_1 \Delta Y_{t-1} + ... + \gamma_{p-1} \Delta Y_{t-p+1}$$

$$+ \theta X_t + \omega_1 \Delta X_t + ... + \omega_q \Delta X_{t-q+1} + e_{t.}$$

- **This "ADL variant" is equivalent to original one (only written differently).**

The Long Run (Total) Multiplier

- **Effect of a slight change in X on Y in the long run.**

To understand the long run multiplier:

- **Suppose X and Y are in an equilibrium or steady state.**

- **All of a sudden, X changes slightly.**

- **This affects Y, which will change and, in the long run, move to a new equilibrium value.**

- **The difference between the old and new equilibrium values for Y = the long run multiplier effect of X on Y.**

- **Can show that the long-run multiplier is $-\theta/\phi$.**

- **For system to be stable (a concept we will not formally define), we need $\phi < 0$.**

Example: The Effect of Financial Liberalization on Economic Growth

Time series data for 98 quarters for a country

- **Y = the percentage change in GDP**

- **X = the percentage change in total stock market capitalization**

- **Assume Y and X are stationary**

Example: The Effect of Financial Liberalization on Economic Growth (cont.)

- **ADL(2,2) with Deterministic Trend Model**

$$\Delta Y_t = \alpha + \delta t + \rho Y_{t-1} + \gamma_1 \Delta Y_{t-1} + \theta X_t + \omega_1 \Delta X_t + \omega_2 \Delta X_{t-1} + e_t.$$

	Coeff.	Stand. Error	t-Stat	P-val	Lower 95%	Upper 95%
Inter.	-.028	.041	-.685	.495	-.110	.054
Y_{t-1}	-.120	.013	-9.46	4.E-15	-.145	-.095
ΔY_{t-1}	.794	.031	25.628	7.E-43	.733	.856
X_t	.125	.048	2.605	.011	.030	.221
ΔX_t	.838	.044	19.111	3.E-33	.750	.925
ΔX_{t-1}	.002	.022	.103	.918	-.041	.046
time	.001	.001	.984	.328	-.001	.002

Example: The Effect of Financial Liberalization on Economic Growth (cont.)

- **Estimate of long run multiplier:**

$$-\left(.125 \Big/ -.120\right) = 1.042.$$

Remember that the dependent and explanatory variables are % changes:

- **The long run multiplier effect of financial liberalization on GDP growth is 1.042 percent.**

- **If X permanently increases by one percent, the equilibrium value of Y will increase by 1.042 percent.**

Time Series Regression when Y and X Have Unit Roots: Spurious Regression

- **Now assume that Y and X have unit roots.**

- **Consider the standard regression of Y on X:**

$$Y_t = \alpha + \beta X_t + e_t.$$

- **OLS estimation of this regression can yield results <u>which are completely wrong</u>.**

- **These problems carry over to the ADL model.**

The Spurious Regression Problem

- **Even if the true value of β is 0, OLS can yield an estimate, $\hat{\beta}$, which is very different from zero.**

- **Statistical tests (using the t-stat or P-value) may indicate that β is not zero.**

- **If β=0, then the R^2 should be zero. In the present case, the R^2 will often be quite large.**

This is called the *spurious regression problem*.

<u>Practical Implication:</u>

With the one exception that we note below, you should never run a regression of Y on X if the variables have unit roots.

The exception occurs if Y and X are <u>cointegrated</u>.

Cointegration

- **Errors are a linear combination of two unit root series (Y and X).**

$$e = Y - \alpha - \beta X$$

- **If Y and X have unit roots, this usually means e has a unit root. This is the cause of the spurious regression problem.**

- **However, in some cases the unit roots in Y and X "cancel each other out" and e does not have a unit root and the spurious regression problem vanishes. This is referred to as <u>cointegration.</u>**

Intuition about Cointegration

Note: If cointegration does **not** occur then e will have a trend in it.

- If X and Y have unit roots then they have stochastic trends. However, if they are cointegrated, the error does not have such a trend. The error will not get too large and Y and X will not diverge from one another.

- Y and X will **trend together**. Jargon: **common trends** or **co-trending**.

- If we are talking about a financial theory involving an equilibrium concept, e is the **equilibrium error**. If Y and X are cointegrated then the equilibrium error stays small.

- If Y and X are cointegrated then there is an equilibrium relationship between them. If they are not, then no equilibrium relationship exists.

Intuition about Cointegration (cont.)

- It is unlikely that a financial system will ever be in precise equilibrium since there are always shocks and unexpected changes which occur. However, departures from equilibrium should not be too large. Hence, if a financial theory which implies an equilibrium relationship between Y and X is correct, then we should observe Y and X being cointegrated.

Summary: if cointegration is present, then not only do we <u>avoid the spurious regression problem</u>, but we also have <u>important economic information</u> (e.g. that an equilibrium relationship exists or that two series are trending together).

Example: Cointegration Between the Spot and Forward Rates

In futures markets, theories involving investors having rational expectations often imply cointegrating relationships.

Spot rate definition

In foreign exchange markets you can buy any major currency (e.g. the $ or the £) in a regular way (e.g. For instance, at the time I am writing this I could purchase $1.97 for £1). This is *spot rate*.

Forward rate definition

Can also agree an exchange rate now, but carry out the actual trade at some future date (e.g. a deal might have the form "I will guarantee that one year from now, I will give you $2.00 for your £1"). This is the *forward rate*.

Similar contracts (and much more complicated ones) can be written in stock markets and, indeed, such *financial derivatives* play a huge role in modern financial markets.

Many financial theories, involving market efficiency and rational expectations of investors, imply that forward rates should be good predictors of future spot rates.

Empirically, it seems that prices of assets (and an exchange rate is a price of an asset) often have unit roots in them (with returns being stationary).

If we combine the financial theory with this empirical regularity, it turns out that they imply that spot and forward rates should be cointegrated.

See Figure 10.1

Figure 10.1: Spot and Forward Rates
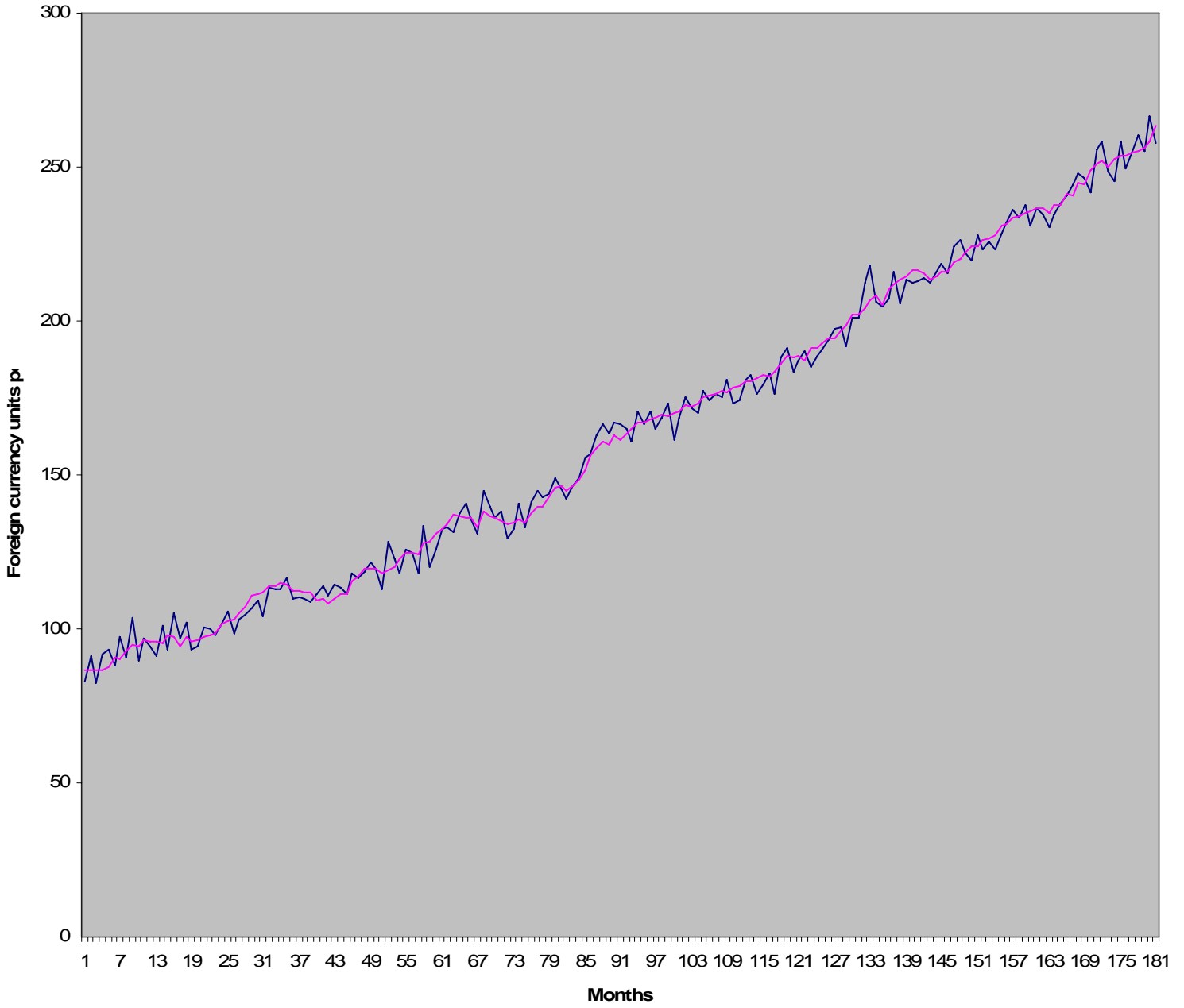
Other examples of financial theories which imply cointegration:

- **Purchasing power parity.**

- **Long term and short term interest rates.**

- **Permanent income hypothesis.**

- **Theories of money demand.**

- **Uncovered interest parity**

- **World business cycles.**

Estimation and Testing with Cointegrated Variables

- **OLS estimation of regression of Y on X is fine <u>if cointegration occurs.</u>**

- **$Y = \alpha + \beta X + e$**
  is called the <u>cointegrating regression.</u>

- **β is the long-run multiplier.**

However, before you run this regression you must make sure the variables are cointegrated, if not you run into the spurious regressions problem.

Testing for Cointegration

- **Many tests available. Here is a simple one called the Engle-Granger test that can be done using regression methods (a better test called the Johansen test described in next chapter)**

- **Intuition: Cointegration occurs if e <u>does not</u> have a unit root.**

- **This suggests we test whether OLS residuals for regression of Y on X have a unit root using Dickey-Fuller test.**

- **This is called the Engle-Granger test**

The Engle-Granger Test

1) Run the regression of Y on X and save the residuals.

2) Carry out a unit root test on the residuals. If the unit root hypothesis is rejected then conclude that Y and X are cointegrated. However, if the unit root is accepted then conclude cointegration does not occur.

Note: When you do Dickey-Fuller test on residuals, do not include a deterministic trend. Inclusion of deterministic trend $\Rightarrow$ errors exhibit trend behaviour.

Don't forget: Before you do this test, you must test for unit roots in Y and X individually. Only if you fail to reject the unit root hypothesis for both X and Y should you even consider thinking about cointegration.

Example: Cointegration Between Spot and Forward

- **Y = spot rate**

- **X = forward rate**

- **Cointegrating regression:**

$$\hat{Y}_t = 0.774 + .996\, X_t.$$

- **Save residuals, u.**

Example: Cointegration Between Spot and Forward Rates (cont.)

- **The Dickey-Fuller strategy suggests we should regress Δu_t on u_{t-1}.**

	Coeff.	St. Error	t-Stat	P-val.	Lower 95%	Upper 95%
Inter.	.024	.292	.083	.934	-.552	.600
u_{t-1}	-1.085	.075	-14.5	6.E-32	-1.233	-.938

- **Compare t-stat for the coefficient on u_{t-1} to a critical value of –2.89.**

- **Since –14.5 is more negative –2.89 we reject the unit root hypothesis and conclude that the residuals do not have a unit root. Hence, Y and X are cointegrated.**

Note: Dickey-Fuller test here is based on an AR(1) model for the residuals. Should have worked with AR(p) and used sequential testing strategy to choose p.

Example: Cointegration Between Spot and Forward Rates (cont.)

- Since X and Y are cointegrated we do not need to worry about the spurious regression problem.

- The coefficient on the forward rate is .996 which is very close to the value of 1 predicted by financial theory

- The long run multiplier is .996.

The Error Correction Model (ECM)

Previous material allows us to:

- **Test whether Y and X are cointegrated.**

- **Obtain an estimate of long run multiplier.**

- **What if you are interested in short run properties (e.g. adjustment to equilibrium)?**

Granger Representation Theorem

If Y and X are cointegrated, then the relationship between the two can be expressed as an ECM.

Simple Example of an ECM

$$\Delta Y_t = \varphi + \lambda e_{t-1} + \omega_0 \Delta X_t + \varepsilon_t,$$

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

Assume $\lambda < 0$ (a stability condition we will motivate later).

ECM Intuition

Dependent variable: ΔY

Explanatory variables:

1. $\Delta X \Rightarrow Y$ changes if X changes.

2. $e_{t-1} \Rightarrow Y$ changes if equilibrium error was made last period.

Note: e_{t-1} is called the *error correction term*.

How Does "Error Correction" Happen?

Assume $\Delta X = 0$ to focus on error correction term.

- If $e_{t-1} > 0$ then Y_{t-1} is above equilibrium.

- ECM says that this will cause ΔY_t to be negative (since $\lambda < 0$ and, hence, $\lambda e_{t-1} < 0$).

- Y starts falling towards equilibrium.

- If $e_{t-1} < 0$ opposite occurs and Y starts rising towards equilibrium.

- Including e_{t-1} as explanatory variable allows equilibrium error to be "corrected".

Note: This reasoning requires $\lambda < 0$. If $\lambda > 0$, equilibrium errors are magnified.

Estimation of ECMs

Many computer packages allow for automatic estimation of ECMs using (slightly) better methods than that described here.

Step 1: Run a regression of Y on X and save the residuals.

Step 2: Run a regression of ΔY on ΔX and the residuals from Step 1 lagged one period.

Remember: before carrying out this two-step estimation procedure for the ECM, you must verify that Y and X have unit roots and are cointegrated.

General form of an ECM

- **Just as in ADL model, may need extra lags of dependent and explanatory variables.**

$$\Delta Y_t = \varphi + \delta t + \lambda e_{t-1} + \gamma_1 \Delta Y_{t-1} + \dots + \gamma_p \Delta Y_{t-p}$$
$$+ \omega_0 \Delta X_t + \dots + \omega_q \Delta X_{t-q} + \varepsilon_t.$$

- **Do not put in extra lags of e.**

- **Use 2 step estimation method described on previous overhead.**

- **Number of lags can be selected using the sequential testing strategy we used for ADL model.**

Example: Cointegration Between Spot and Forward Rates (continued)

- **Before we ran regression of Y on X and saved the residuals, u_t.**

- **Use residuals to estimate the ECM:**

$$\Delta Y_t = \varphi + \lambda u_{t-1} + \omega_0 \Delta X_t + \varepsilon_t.$$

	Coeff.	St. Error	t-Stat	P-val.	Lower 95%	Upper 95%
Inter.	-.023	.342	-.068	.946	-.700	.654
u_{t-1}	-1.085	.075	-14.46	9.E-32	-1.233	-.937
ΔX_t	1.044	.182	5.737	4.E-08	.685	1.403

- **λ and ω_0 summarise short run properties.**

- **$\hat{\lambda} = -1.085 \Rightarrow$ positive errors tend to cause ΔY to be negative and hence Y to fall. In particular, an equilibrium error of 1 unit tends to cause the spot rate to fall by 1.085 unit in the next period, *ceteris paribus*.**

Time Series Regression when Y and X Have Unit Roots but are NOT Cointegrated

- *Do not run a regression of Y on X (spurious regression problem).*

- **The presence of such characteristics suggests that you should rethink your basic model and include other explanatory variables.**

- **Instead of working with Y and X themselves, for example, you could difference them. (Remember that if Y and X have one unit root, then ΔY and ΔX should be stationary).**

- **In this case, you could work with ΔY and ΔX and estimate the ADL model using the techniques described at the beginning of this chapter.**

Time Series Regression when Y and X Have Unit Roots but are NOT Cointegrated (cont.)

- **Use an ADL using** *changes of the variables:*

$$\Delta Y_t = \alpha + \delta t + \phi_1 \Delta Y_{t-1} + ... + \phi_p \Delta Y_{t-p}$$
$$+ \beta_0 \Delta X_t + ... + \beta_q \Delta X_{t-q} + e_t.$$

- **OLS estimation and testing can be done as for the ADL model.**

- **Be careful in how you interpret results.**

Interpretation

Example

- **Suppose Y = exchange rates and X = interest rates.**

- **If Y and X are cointegrated, or if both are stationary, we obtain an estimate of the long run effect of a small change in interest rates on exchange rates.**

- **If Y and X are neither stationary nor cointegrated, we obtain an estimate of the long run effect of <u>a small change in the change of interest rates</u> on <u>the change in exchange rates</u>.**

Chapter Summary

1. If all variables are stationary, then an ADL(p,q) model can be estimated using OLS. Statistical techniques are all standard.

2. A variant on the ADL model is often used to avoid potential multicollinearity problems and provide a straightforward estimate of the long run multiplier.

3. If all variables are nonstationary, great care must be taken in the analysis due to the spurious regression problem.

4. If all variables are nonstationary but the regression error is stationary, then cointegration occurs.

5. If cointegration is present, the spurious regression problem does not occur.

6. Cointegration is an attractive concept for financial researchers since it implies that an equilibrium relationship exists.

7. Cointegration can be tested using the Engle-Granger test. This test is a Dickey-Fuller test on the residuals from the cointegrating regression.

8. If the variables are cointegrated then an error correction model can be used. This model captures short run behavior in a way that the cointegrating regression cannot.

9. If the variables have unit roots but are not cointegrated, you should not work with them directly. Rather you should difference them and estimate an ADL model using the differenced variables.