WARWICK
Business School

Empirical Finance
Lecture 4: When OLS is no longer valid: examples of alternative estimators used in empirical finance.

Module Leader: Dr Stuart Fraser

stuart.fraser@wbs.ac.uk

Room D1.18 (Social Studies)

The University of Warwick

Introduction

- Last time…
  - Estimating and testing CAPM and 3-factor model.
  - Need to carry out misspecification testing (test assumptions of the CLRM).
- Today…
- Consequences of heteroscedasticity and autocorrelation for OLS estimators.
  - Look at a solution widely used in empirical finance: Newey-West HAC var-cov matrix.
- Consequences of endogenous regressors (correlated with the error term).
  - Alternative estimators used in empirical finance when regressors are endogenous: IV/GMM

Consequences of heteroscedasticity and autocorrelation

In the presence of heteroscedasticity or autocorrelation OLS point estimators remain unbiased and consistent (see e.g., Gujarati Chps 11+12; Brooks Chp 4)

> An estimator is consistent if its sampling distribution 'collapses' on the true parameter value as $T \rightarrow \infty$

However the standard formula for the variance-covariance matrix...

$$\text{var}\left(\hat{\beta}\right) = \sigma^2 \left(X'X\right)^{-1}$$

...is no longer correct.

Therefore whilst OLS point estimators are unbiased (and consistent) inferences based on the above formula (*t*-, *F*-tests and confidence intervals) are invalid.

Consequences of heteroscedasticity and autocorrelation

Under het. and/or auto. the correct formula for $\operatorname{var}(\hat{\beta})$ is:

$$\operatorname{var}(\hat{\beta}) = (X'X)^{-1} X'\Omega X (X'X)^{-1}$$
$$\Omega = E(\varepsilon\varepsilon')$$

Ω is the variance-covariance matrix of the error terms. If the errors are homoscedastic and uncorrelated then this matrix is diagonal

$\Omega = \sigma^2 I$ see Appendix 1

In that case

$$\operatorname{var}(\hat{\beta}) = \sigma^2 (X'X)^{-1}$$

Therefore if a <u>consistent</u> estimator of $\operatorname{var}(\hat{\beta})$ can be found then we can…

- Use OLS point estimators (which are unbiased and consistent)
- Combined with a consistent estimator of $\operatorname{var}(\hat{\beta})$

…yielding an estimator which is consistent <u>and</u> gives valid inferences.

Principle underlying the use of <u>OLS point estimates</u> with inferences based on a <u>Newey-West HAC var-cov matrix</u>.

Newey-West heteroscedasticity and autocorrelation consistent (HAC) variance-covariance matrix

$$\widehat{\text{var}}\left(\hat{\beta}\right) = \left(X'X\right)^{-1} T\hat{V}_T \left(X'X\right)^{-1}$$

where

$$\hat{V}_T = T^{-1}\sum_{t=1}^{T}\hat{\varepsilon}_t^2 x_t x_t' + T^{-1}\sum_{j=1}^{q} w_j \sum_{t=j+1}^{T}\left(x_t\hat{\varepsilon}_t\hat{\varepsilon}_{t-j}x_{t-j}' + x_{t-j}\hat{\varepsilon}_{t-j}\hat{\varepsilon}_t x_t'\right)$$

This first term corrects the var-cov matrix for heteroscedasticity:

σ_t^2 is estimated using $\hat{\varepsilon}_t^2$

(on its own this term is known as a White var-cov matrix)

The second term corrects the var-cov matrix for autocorrelation:

$E\left(\varepsilon_t\varepsilon_{t-j}\right)$ is estimated using $\hat{\varepsilon}_t\hat{\varepsilon}_{t-j}$

This estimator truncates the autocorrelations at lags q<T.
It also gives a diminishing weight w to autocorrelations as the lag increases.

Bartlett weights are commonly used ⇒

$$w_j = 1 - j/q$$
$$w_0 = 1, w_q = 0$$

$T\hat{V}_T$ is a consistent estimator of $X'\Omega X$

Application of Newey-West HAC variance-covariance matrix

So to re-cap:

- OLS point estimates with inferences based on Newey-West HAC standard errors (rather than OLS standard errors) is one solution to the problem of het. and/or auto.
- However this estimator is not BLUE – there is an estimator with a smaller variance (more efficient): Generalized Least Squares (GLS).

Nonetheless Newey-West HAC var-cov matrices are widely used in empirical finance.

A common instance in which they are used is where the holding period for returns is greater than the sampling frequency of the data.

⇒Overlapping data problem (see Verbeek Chp 4.11.3 for an illustration of this problem in the FX market and see below).

Overlapping data problem

Suppose we have a sample of daily data for returns but our model is for *m*-period holding returns (*m* > one day)

$$r_t^m = p_t - p_{t-m} = (p_t - p_{t-1}) + (p_{t-1} - p_{t-2}) + \ldots + (p_{t+1-m} - p_{t-m})$$

$$= r_t + r_{t-1} + \ldots + r_{t+1-m}$$

log prices

Even if the one-period returns are independent the m-period returns for different periods consist of 'overlapping' one-period returns ⇒ the m-period returns are correlated:

$$r_t^m = r_t + r_{t-1} + \ldots + r_{t+1-m}$$

$$r_{t-1}^m = r_{t-1} + \ldots + r_{t+1-m} + r_{t-m}$$

$$\ldots$$

$$r_{t+1-m}^m = r_{t+1-m} + r_{t-m} + \ldots$$

$$\operatorname{var}(r_t^m) = m \operatorname{var}(r_t)$$

$$\operatorname{cov}(r_t^m, r_{t-1}^m) = (m-1)\operatorname{var}(r_t)$$

$$\ldots$$

$$\operatorname{cov}(r_t^m, r_{t+1-m}^m) = \operatorname{var}(r_t)$$

$$\operatorname{cov}(r_t^m, r_{t-m}^m) = 0$$

The autocorrelations decay to zero after m-1 lags. This is indicative of an MA(m-1) process.

Any static model involving the m period returns will therefore have an autocorrelated [MA(m-1)] error term ⇒ OLS inferences are invalid.

Use of Newey West standard errors is an appropriate remedy in this case.

Method of Moments Estimation (MME)

OLS as a MME

$$y = X\beta + \varepsilon$$

The CLRM requires the following *population* moment conditions

$$E(X'\varepsilon) = 0$$

$A2: E(\varepsilon|X) = 0$

$\Rightarrow E(X'\varepsilon) = 0$

X is linearly independent of the error term

This is a k×1 vector.
⇒There are k moment conditions which the OLS estimator must satisfy.

These moment conditions imply the values of X are determined outside of the model:
X is exogenous

The MME finds $\hat{\beta}$ by solving the *sample* moment conditions

$$\frac{1}{T}X'\hat{\varepsilon} = \frac{1}{T}X'\left(y - X\hat{\beta}\right) = 0$$

There are *k* sample moment conditions and *k* unknown parameters ⇒possible to find a unique solution for $\hat{\beta}$.

OLS as a MME

$$\frac{1}{T} X'\left(y - X\hat{\beta}\right) = 0$$

$$\Rightarrow X'y = X'X\hat{\beta}$$

$$\Rightarrow \hat{\beta} = \left(X'X\right)^{-1} X'y$$

The MM estimator for the CLRM is <u>identical</u> to the OLS estimator.

Properties of MME

MME is a general approach to estimation which imposes population moment conditions (required by the statistical model) to hold exactly in the sample.

These moment conditions are then solved for the unknown parameters in the model (example above).

MME has 3 attractive features:

1. It makes no distributional assumptions.
2. It is a consistent estimator.
3. It is a *very* general technique (e.g., applicable to non-linear models).

Endogenous regressors

In many instances in economics/finance there is a two way or simultaneous relationship between X and y.

$\Rightarrow$both X and y are determined <u>inside</u> the model.

$\Rightarrow$X is <u>endogenous</u>.

Endogeneity is common due to the <u>non-experimental</u> nature of economic/finance data $\Rightarrow$

$$\boxed{E(X'\varepsilon) \neq 0}$$

In that case OLS/MME estimation (assuming $E(X'\varepsilon) = 0$) is invalid. The estimator is biased (and inconsistent)

$$\begin{aligned}\hat{\beta} &= (X'X)^{-1} X'y \\ &= (X'X)^{-1} X'(X\beta + \varepsilon) \\ &= \beta + (X'X)^{-1} X'\varepsilon\end{aligned} \quad \Rightarrow \quad \begin{aligned}E(\hat{\beta}) &= \beta + (X'X)^{-1} E(X'\varepsilon) \\ &\neq \beta \quad \text{unless } E(X'\varepsilon) = 0.\end{aligned}$$

Examples of endogeneity in finance: testing CIP and UIP (Cuthberston & Nitzsche Chps 24.3/4, 25.1/2)

Covered Interest Parity

$$\Rightarrow \frac{F_t^h}{S_t} = \frac{1+r_t}{1+r_t^*}$$

Where F^h (h-period forward exchange rate), S (spot exchange rate) are denominated in terms of the domestic currency price of a unit of foreign exchange.

r (domestic interest rate), r^* (foreign interest rate) (interest rate on h-period T-Bills).

CIP implies you <u>can't</u> earn abnormal profits from…

1. Borrowing £x at the rate r.
2. Converting £→$ at the spot rate S: $\pounds x \rightarrow \$\frac{x}{S}$
3. Investing in h-period US bonds at the rate r*: $\$\frac{x}{S}\left(1+r^{*}\right)$
4. Simultaneously switching $'s back to £'s at the h-period forward rate F^h: $\pounds\frac{F^{h}}{S}x\left(1+r^{*}\right)$
5. Given riskless arbitrage (the investors receipts are covered in the forward market) would expect under EMH that:

$$\frac{F^{h}}{S}x\left(1+r^{*}\right)=x\left(1+r\right)$$

$$\Rightarrow\frac{F^{h}}{S}=\frac{1+r}{1+r^{*}}$$

Uncovered Interest Parity

Similar idea to CIP but with the key difference that investors are willing to take a bet on what the exchange rate will be at the time of converting \$'s back to £.

$$\frac{E_t\left(S_{t+h}\right)}{S_t} = \frac{1+r_t}{1+r_t^*}$$

The investment in \$'s is risky because the £ receipts are not covered (uncovered) in the forward market (contrast CIP).

⇒UIP will only hold if the market is dominated by risk neutral speculators.

Testing CIP and UIP

If we want to test these hypotheses it is not clear which variable is the dependent variable and which is the explanatory variable.

Both sides of the equation will adjust to deviations from equilibrium.

For example: In CIP, suppose a shock causes the forward rate to depreciate (F/S rises). As a result, the demand for foreign assets will rise (p*rises, r* falls) which drives the market back to CIP equilibrium.

In other words the variables (forward rate premium and relative interest rates) are endogenous.

Testing equations for CIP and UIP

<u>CIP equation</u>

h-period forward rate: h≥1.

$$\left(f^{h}-s\right)_t=\alpha+\beta\left(r-r^{*}\right)_t+\gamma x_t+\varepsilon_t$$

h-period interest rates

<u>UIP equation</u>

$$\Delta s_{t+h}=\alpha+\beta\left(r-r^{*}\right)_t+\gamma x_t+\varepsilon_t$$

where $f=\log(F), s=\log(S)$ and $\log(1+r)\cong r$ (x represents other variables).

For each relationship the null hypothesis is

$$H_0:\alpha=0,\beta=1,\gamma=0$$

To repeat, OLS is <u>invalid</u> for testing these hypotheses due to the <u>endogeneity</u> of the regressors.

Instrumental Variable Estimator (IVE)

In the case of endogenous regressors the model is:

$$y = X\beta + \varepsilon$$
$$E(X'\varepsilon) \neq 0$$

→ OLS/MME invalid

Z is a T×m matrix
(Recall X is a T×k matrix).

But suppose we can find a set of m variables Z that are correlated with X but not ε.

In that case the Z are Instrumental Variables (IV).

The IVs must satisfy:

$E(Z'\varepsilon) = 0$	(Z **uncorrelated** with the error term)	$IV1$
$E(Z'X) \neq 0$	(Z **correlated** with/informative about X)	$IV2$

IVE (just/exactly identified model)

Given the above moment conditions and assuming the model is just/exactly identified (m=k)…

⇒One instrument for each endogenous regressor

…then we can solve the *k* sample moment restrictions to find the IV estimator of β:

$$\frac{1}{T}Z'\hat{\varepsilon} = \frac{1}{T}Z'\left(y - X\hat{\beta}_{IV}\right) = 0$$

$$\Rightarrow Z'y = Z'X\hat{\beta}_{IV}$$

$$\Rightarrow \hat{\beta}_{IV} = \left(Z'X\right)^{-1}Z'y$$

This IVE is another example of a MME.

The estimator is consistent if IV1 and IV2 hold.

Note that if m<k the model is under-identified ⇒ it is not possible to estimate β.

YOU NEED AT LEAST ONE INSTRUMENT FOR EACH ENDOGENOUS REGRESSOR FOR IV TO WORK.

IVE (over-identified model): Two-stage least squares (2SLS)

If m>k there is no unique solution for β based on the previous IVE/MME.

We have to add a preliminary step to the previous estimator

STAGE 1: Regress the endogenous variables on the instruments

$$x_i = Z\pi_i + v_i, \quad i = 1,\dots,k$$

Obtain the fitted values $\hat{x}_i$. Form a $T \times k$ matrix of the fitted values:

$$\hat{X} = (\hat{x}_1,\dots,\hat{x}_k)$$

This stage purges the regressors of endogeneity. (The fitted values are linear combinations of the Instruments.)

STAGE 2: Solve the k sample moment restrictions

$$\frac{1}{T}\hat{X}'\hat{\varepsilon} = \frac{1}{T}\hat{X}'\left(y - X\hat{\beta}_{IV}\right) = 0$$

$$\Rightarrow \hat{X}'y = \hat{X}'X\hat{\beta}_{IV}$$

$$\Rightarrow \hat{\beta}_{IV} = \left(\hat{X}'X\right)^{-1}\hat{X}'y$$

Stage 2 is equivalent to regressing y on the stage 1 fitted values (instead of X).

This approach is also valid if the model is just identified.

If the model is just-identified the 2SLS estimator is identical to the IVE/MME on the previous slide.

2SLS: CIP relationship between UK-US (60 day/3 month forward premium) (see also Appendix 2)

Dependent Variable: LOG(UK_FRATE_USD_GBP)				
-LOG(UK_SPOT_RATE_USD_GBP)				
Method: Two-Stage Least Squares				
Date: 02/05/06 Time: 21:12				
Sample (adjusted): 5/10/2001 9/30/2005				
Included observations: 1147 after adjustments				
Instrument list: LOG(UK_GDP)-LOG(UK_GDP(-1))				
LOG(US_GDP) -LOG(US_GDP(-1)) UK_TBILLS(-1) US_TBILLS(-1)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.00721	0.000423	-17.0571	0
(UK_TBILLS-US_TBILLS)/	0.947822	0.016892	56.11167	0
R-squared	0.734942	Mean dependent va		0.01585
Adjusted R-squared	0.73471	S.D. dependent var		0.006551
S.E. of regression	0.003374	Sum squared resid		0.013037
Durbin-Watson stat	1.458438	Second-stage SSR		0.013336

$$H_0 : \alpha = 0, \beta = 1$$

Test Statis	Value	df	Probability
F-statistic	3630.205	(2, 1145)	0
Chi-square	7260.41	2	0

Generalized Method of Moments (GMM) estimator

MME works when m=k – if m>k the model is over-identified (more equations than unknowns).

One solution would be to drop instruments – but this would reduce the efficiency of the estimator.

Instead GMM chooses estimates of β such that the m sample moments are as close as possible to zero.

This is done by minimizing a quadratic form:

Choose $\hat{\beta}$ to minimize

$$\left(\frac{1}{\mathrm{T}}\hat{\varepsilon}'Z\right)W_T\left(\frac{1}{T}Z'\hat{\varepsilon}\right)$$

W is an m×m weighting matrix. It tells how much weight to attach to each of the sample moment conditions.

Sample moments with a low variance should receive more weight than those with a large variance (because they're more informative about the β's).

This suggests using the inverse of the var-cov matrix of the sample moments as a weighting matrix

IV as a GMM estimator

The GMM estimator is given by

$$\hat{\beta}_{GMM} = \left(X'ZW_T Z'X\right)^{-1} X'ZW_T Z'y$$

If we assume homoscedasticity and no autocorrelation then

$$E(\varepsilon\varepsilon') = \sigma^2 I$$

$$\Rightarrow W_T = \left[\frac{1}{T} Z'E(\varepsilon\varepsilon')Z\right]^{-1} = \left(\sigma^2 \frac{1}{T} Z'Z\right)^{-1}$$

The weighting matrix is the inverse of the var-cov of the sample moments.

In that case 2SLS and GMM are identical

$$\hat{\beta}_{GMM} = \left(X'Z(Z'Z)^{-1}Z'X\right)^{-1} X'Z(Z'Z)^{-1}Z'y$$

$$= \left(\hat{X}'X\right)^{-1} \hat{X}'y = \hat{\beta}_{IV}$$

$$X = Z\pi + v \quad \text{(2SLS Stage 1)}$$

$$\hat{\pi} = (Z'Z)^{-1} Z'X$$

$$\Rightarrow \hat{X} = Z\hat{\pi} = Z(Z'Z)^{-1} Z'X$$

$$\Rightarrow \hat{X}' = X'Z(Z'Z)^{-1} Z'$$

IV as a GMM estimator

More generally we can allow for autocorrelation and/or heteroscedasticity in the model.

In that case the weighting matrix is given by

$$W_T = \left(\frac{1}{T} Z'\Omega Z \right)^{-1}$$

We can estimate the var-cov of the sample moments…

$$\frac{1}{T} Z'\Omega Z$$

…using a Newey-West HAC estimator (see slide 5).

GMM estimation of the UIP relationship

The CIP relationship tested previously was for a 3-month forward premium.

The corresponding test of UIP is therefore for the 3 month holding period return on sterling.

But the data are sampled daily:

⇒overlapping data problem.

⇒autocorrelated distrubances (see above).

Also there is an endogeneity problem with the regressor

⇒so we need to use an IV/GMM estimator

Therefore we'll use an IV/GMM estimator with Newey-West estimated weights.

GMM with Newey-West HAC Var-Cov matrix: Testing UIP between the UK and US (60 day/3 month holding period returns)

Dependent Variable: LOG(UK_SPOT_RATE_USD_GBP(60))				
-LOG(UK_SPOT_RATE_USD_GBP)				
Method: Generalized Method of Moments				
Date: 01/31/07 Time: 21:29				
Sample (adjusted): 5/10/2001 9/30/2005				
Included observations: 1147 after adjustments				
Kernel: Bartlett, Bandwidth: Fixed (6), No prewhitening				
Simultaneous weighting matrix & coefficient iteration				
Convergence achieved after: 5 weight matrices, 6 total coef iterations				
Instrument list: LOG(UK_GDP)-LOG(UK_GDP(-1))				
LOG(US_GDP)-LOG(US_GDP(-1)) UK_TBILLS(-1) US_TBILLS(-1)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.010532	0.010875	0.968495	0.333
(UK_TBILLS-US_TBILLS)/	-0.91841	0.454022	-2.02284	0.0433
R-squared	0.029879	Mean dependent va		-0.01115
Adjusted R-squared	0.029031	S.D. dependent var		0.038325
S.E. of regression	0.037764	Sum squared resid		1.632947
Durbin-Watson stat	0.039079	J-statistic		0.011141

$$H_0 : \alpha = 0, \beta = 1$$

Test Statis	Value	df	Probability
F-statistic	87.37413	(2, 1145)	0
Chi-square	174.7483	2	0

Conclusions

1. Newey-West HAC var-cov matrices are used widely in empirical finance to correct OLS inferences for het. and/or auto.

   It's a particularly useful approach in the context of overlapping data problems.
2. Endogenous variables are common in economics/finance (the data are non-experimental)

   Use an IV/GMM estimator <u>not</u> OLS when the explanatory variables are endogenous.
3. Weight the sample moments appropriately if there is heteroscedasticity and/or autocorrelation.

   Estimate the weights with Newey-West's HAC estimator.

References

Cuthbertson and Nitzsche (2004) Quantitative financial economics: stocks, bonds and foreign exchange, Wiley: Chichester. Chps 24.3/4, 25.1/2** (CIP and UIP)

Verbeek (2004) A Guide to Modern Econometrics, 2nd Edition, Wiley: Chichester. Chp 5** (for more on IV/GMM).

**Key reading

Appendix 1: Variance-covariance matrix of the error term

$$\Omega = E(\varepsilon\varepsilon') = \begin{pmatrix} E(\varepsilon_1^2) & E(\varepsilon_1\varepsilon_2) & \dots & E(\varepsilon_1\varepsilon_T) \\ E(\varepsilon_2\varepsilon_1) & E(\varepsilon_2^2) & & E(\varepsilon_2\varepsilon_T) \\ \dots & & \dots & \dots \\ E(\varepsilon_T\varepsilon_1) & E(\varepsilon_T\varepsilon_2) & \dots & E(\varepsilon_T^2) \end{pmatrix}$$

The terms on the diagonal are the variances of the error terms $\Rightarrow \sigma_t^2$

The terms off the diagonal are the covariances between errors in different periods (auto-covariances).

Under the assumptions homoscedasticity and no autocorrelation then:

$$\Omega = \sigma^2 I$$

A diagonal matrix with σ^2 on the diagonal and zeroes everywhere else

Appendix 2: Interpreting the CIP results

CIP is rejected ($H_0 : \alpha = 0, \beta = 1$ is rejected)

The model suggests an abnormal return of -0.8%

$$
\begin{aligned}
f - s &= r - r^* + \hat{\alpha} + (\hat{\beta} - 1)(r - r^*) \\
&= r - r^* - 0.007 + (0.948 - 1)(0.043 - 0.020) \\
&= r - r^* - 0.008
\end{aligned}
$$

(the sample average UK and US interest rates are 4.3% and 2.0% respectively).

Accordingly

$$\frac{F}{S} = \frac{1 + r}{1 + r^*} e^{-0.008}$$

The return for a UK investor (investing £x in the US) is therefore

$$£ \frac{F}{S} x(1 + r^*) = £ e^{-0.008} x(1 + r)$$

Interpreting the CIP results

In other words the UK investor makes a *loss*

$$\pounds e^{-0.008} x(1+r) - \pounds x(1+r) < 0$$

So if x=£1m and r=4.3% the implied loss is about £8,311 (annualized gross).

However for a US investor (investing $x in the UK) the return is:

$$\$ \frac{S}{F} x(1+r) = \$ e^{0.008} x(1+r^*)$$

So this investor gains $\$e^{0.008}x(1+r^*) - x(1+r^*) > 0$

So if x=$1m and r*=2.0% the implied risk free profit is about $8,193 (annualized gross).

But transactions costs of 0.8% or more would wipe out any abnormal returns.