Tutorial, Part 13

(G)ARCH models

Example

AutoRegressive Conditional Heteroscedastic (ARCH) models

Heteroscedasticity: $Var(u_t) \neq \sigma^2$

→ the disturbance variance is modeled as:

$$\sigma_t^2 = h_t = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \dots + \alpha_q u_{t-q}^2$$

→ VARIANCE ALWAYS NON-NEGATIVE

$\alpha_i \geq 0 \quad \forall i = 0, 1, 2, \dots, q$

or, by analogy with ARMA models:
Generalized ARCH → GARCH (p, q) models

$$h_t = \sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \dots + \alpha_q u_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \beta_2 \sigma_{t-2}^2 + \dots + \beta_p \sigma_{t-p}^2$$

In most cases GARCH (1, 1) is sufficient.

a) ensure stationarity

b) constant term = mean

LM test for ARCH effects of order 1:

$u_t^2 = \gamma_0 + \gamma_1 u_{t-1}^2 \Longrightarrow LM = T \cdot R^2 = 191{,}26$

c) reur → ARMA (0,1)
GARCH (1,1)

$$reur_t = y_t = \mu + u_t + \theta_1 u_{t-1}$$

$$h_t = \sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta \sigma_{t-1}^2$$

(MLE)

Test: $H_0: \alpha_1 = \beta = 0$

d) Asymmetry of news-response

plot conditional variance predictions at different values of e_t (u_t) and s_t^2 (σ_t^2) → (static)

GJR model:

additional threshold term

STATA: tarch

$$reur_t = y_t = \mu + u_t + \theta u_{t-1}$$

$$h_t = \sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta \sigma_{t-1}^2 + \gamma u_{t-1}^2 \cdot I_{t-1}$$

where $I_{t-1} = \begin{cases} 1 & ; u_{t-1} < 0 \\ 0 & ; \text{otherwise} \end{cases}$

Exponential GARCH → EGARCH model

$$\log(\sigma_t^2) = f(\log(\sigma_{t-1}^2), u_{t-1}, \sigma_{t-1}^2)$$

STATA: earch, egarch

e) GARCH-in-mean → GARCH-M model

$$reur_t = y_t = \mu + \delta \sigma_{t-1} + u_t + \theta_1 u_{t-1}$$

$$h_t = \sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta \sigma_{t-1}^2$$

δ : feedback from cond. variance to cond. mean
if $\delta > 0$ → risk premium
(not sig.) STATA: archm

f) static vs. dynamic (out-of-sample) forecast

g) VAR model → simplest identified option: 1 lag

Diagonal half-vectorization → DVECH model
assumes no volatility spillovers →
no covariances, only variances

Initial restrictions (as a starting point for easier optimization):
- all ARCH terms are the same,
- all GARCH terms are the same.

→ estimation without constraints based on this result

STATA:
mgarch
↳ MULTIVARIATE

h)

Constant Conditional Correlation → CCC model

direct correlation (as opposed to covariance)
modeling → assumption:
correlations between disturbances u_t (or variable values y_t) are fixed through time
→
conditional variances (diagonal elements) are identical to GARCH,
off-diagonal elements h_{ijt} $(i \neq j)$ are defined via correlation coefficients ρ_{ij}:

$$h_{ijt} = \rho_{ij} \sqrt{h_{iit} \, h_{jjt}}$$

i) volatility = variance forecasts

STATA: tsappend (= generate new observations)

Tutorial, Part 14

Discrete choice models, MLE

Example 1

we will only deal with dichotomous dep. vars.

a) dependent variable: binary (dichotomous)
explanatory variables: numeric, binary

b) relative frequencies, conditional relative frequencies

c) Linear probability model

P (butcher = 1) = f(expense-chicken, safe-butcher, supermarket, age, hhsize)
significance of the partial regression coefficients?

t-test : $H_0: \beta_3$ (supermarket) = 0
t-test : $H_0: \beta_4$ (age) = 0 ⇒ $b_4 = 0{,}00417$: With an additional year of age the probability of buying chicken meat at the butcher's increases on average by 0,417 perc-points, ceteris paribus.

d) · Weird values (outside the [0,1] interval) for predicted probability.
· residuals not normally distributed
· heteroscedasticity
· questionable R^2
· linear relationship?

y^* - the unobserved (underlying, latent) process
y - the observed final decision based on y^* ⇒ transformation
$P(y=1/x) = \Psi(y^*)$. (Ψ - the "link" function) → MLE

e) MLE (Maximum Likelihood Estimation)

- the „opposite" approach
- We don't know exactly which distribution generated the data and try to find it by considering the parameters as functions of the observed (= fixed) values.
- However, there are 2 assumptions:
  - we know the shape of the distribution
  - observations are i.i.d.

The Likelihood function: $L = \prod_{i=1}^{n} L_i$

↳ joint probability distribution = product of probability density function values for every observation

⇓

maximize the L → solve for the parameters

↳ easier to manipulate $\log L$ (same max.)

---

*

$$P(y=1 \mid x) = \Psi(y^*)$$

CDF

if the distribution function Ψ is:

- standard normal (ϕ) → PROBIT (NORMIT) MODEL
- logistic (Λ) → LOGIT MODEL

PROBIT and LOGIT models are non-linear → contributions of partial regression coefficients are not constant: only direction (= sign) and stat. significance matter, no interpretation of values

→ interpretation of MARGINAL EFFECTS (mf_{x_j}):
- usually calculated at mean values of all explanatory vars (= at the centroid = multi-dimensional mean)
- same sign as the reg. coeffs. (as distr. f. is strictly increasing)

PROBIT MODEL: $P(y=1|x) = \phi(y^*)$

Interpretation of marginal effect (e.g. $mf_{age} = 0.0042$):
With an additional year of age the probability of buying chicken meat at the butcher's increases on average by 0.42 perc. points, ceteris paribus, estimated at mean values of expl. vars.

f) LOGIT MODEL: $P(y=1|x) = \Lambda(y^*)$
↳ easier to interpret in terms of odds (ratio)

g) - $\text{coeffs}_{LOGIT} \approx 1.7 \ \text{coeffs}_{PROBIT}$
- very similar marginal effects

h) $\text{hit ratio} = \frac{\text{correctly classified cases}}{\text{all cases}}$

i) lower hit ratio → not surprising because of less expl. variables

Example 2

a) Yes, the signs make sense. (but no interpretation!)

b) test statistic for joint statistical significance of all partial regression coefficients:

$H_0: \beta_j = 0 \quad \forall j = 1, 2, \dots, k$

just the constant, without expl. vars.

$$LR = -2(\ln L_{INITIAL} - \ln L_{MODEL}) = -2(-27.525553 - (-16.782923)) = 21.485$$

$$df = k_{MODEL} - k_{INITIAL} = \#\text{ of expl. vars}$$

$$\text{Pseudo } R^2 = 1 - \frac{\ln L_{MODEL}}{\ln L_{INITIAL}} = 1 - \frac{-16.78}{-27.53} = 0.39$$

c) Marginal effects (PROBIT) for the centroid (at means):

- continuous x_j: $\frac{\partial p(x)}{\partial x_j} = \varphi(xb) \cdot b_j$

AT MEANS

PROBABILITY DENSITY FUNCTION (PDF)

↳ Stata: normalden()

- discrete x_j (e.g. dummy):

$$\Delta p(x) = \phi(b_0 + \dots + b_j \cdot 1) - \phi(b_0 + \dots + b_j \cdot \cancel{0})$$

AT MEANS

(CUMULATIVE) DISTRIBUTION FUNCTION (CDF)

↳ Stata: normal()

d) $b_{LOGIT} \approx 1.7\, b_{PROBIT}$, same signs

e) Calculated similarly as for PROBIT, but logistic distribution used instead of std. normal:

$\varphi \to \lambda$ (Stata: logisticden()); $\phi \to \Lambda$ (Stata: logistic())

↳ similar results as for PROBIT

Interpretation:

Change by 1 (unit) in logs is not a small change:
↑ by 1 (unit) in log Y = ↑ by 172% in Y
Reversed:
↑ by 1 % in Y = ↑ by 0.00995 (unit) in log Y
⇒ if multiplied by 100 to interpret in perc. points:
$0.00995 * 100 = 0.995 \approx 1$

$mf_{log\text{-}income} = 1.0348$ → If a person's income increases by 1%, then on average P(abroad=1) increases by 1.0348 perc-points, ceteris paribus, at means.

$mf_{age} = -0.0244$ → With an additional year of age the probability of spending holidays abroad is on average lower by 2.44 perc. points, at means, ceteris paribus.

$mf_{pet} = -0.5135$ ⇒ If a person has a pet the probability of spending holidays abroad is on average lower by 51.35 perc. points compared to a person without a pet, estimated at mean values, ceteris paribus.

f) The ODDS: $\Omega = \frac{P(Y=1)}{P(Y=0)}$

How many times more likely is it for $Y=1$ to happen compared to $Y=0$.

$$\Omega = e^{X\beta} = e^{Y^*} \quad / \log$$

$$\log \Omega = X\beta \ln e = X\beta = Y^*$$

LOGIT

Odds ratio for an expl. var. X_j (for odds of $Y=1$):

$$OR_{x_j} = \frac{\Omega(X_j; x_j+1)}{\Omega(X_j; x_j)} = \frac{e^{(xb_j\,(x_j+1)b_j)}}{e^{xb}} = \frac{\cancel{e^{xb}}\,e^{b_j}}{\cancel{e^{xb}}} = e^{b_j}$$

Interpretation:

OR_log_income = 70.1505 --> 70.1505 ^ ln1.01 = 1.043 -->
If income increases by 1% then the odds of spending holidays abroad increase on average by a factor 1.043 (by 4.3%), ceteris paribus.

$$OR_{age} = e^{b_{age}} = e^{b_2} = e^{-0.1004} = 0.904 < 1 \longrightarrow$$

With an additional year of age the odds of spending holidays abroad decrease by a factor 0.904 on average, ceteris paribus. (the P(abroad=1) decreases, too)

$$OR_{pet} = e^{b_{pet}} = e^{b_3} = e^{-2.278} = 0.102 \longrightarrow$$

If a person has a pet the odds of spending holidays abroad are on average lower by a factor 0.102 compared to a person without a pet, ceteris paribus.