Ch.16 Which Method

- The main content of this chapter

1.ML vs. GMM

2.Why ML is often ignored

3.OLS VS. GLS cross-sectional regressions

4.Additional examples of trading off efficiency for robustness

5. Addressing model misspecification

6. Auxiliary model

7. Statistical philosophy

8. summary

Advantages of GMM

- Linear Factor Model: Limited Advantage
- Main applications:
  - Nonlinear models;
  - Conditioning information;
  - Circumvent inevitable model misspecification and focus on the interesting issues.

Which method

- Trade off between
  - statistical efficiency
  - Economic interpretability of the results

1.ML vs. GMM

ML: ·consistency

$$p\lim(\hat{\theta}_T) = \theta$$

asymptotic normally

$$\hat{\theta}_T \overset{a}{\sim} N(\theta, I^{-1}(\theta))$$

asymptotic efficiency

$$\sqrt{n}(\hat{\theta}_T - \theta) \rightarrow N(0, V)$$

- Invariance

If $\hat{\theta}$ is the MLE of θ , $g(\hat{\theta})$ is the MLE of $g(\theta)$

- E(score)=0,var(score)=$I(\theta)$

- ML vs. GMM: A bad issue

1).ML is a special case of GMM. (ch.14)

The key issue: the choice of moments

2).GMM is more flexible than ML.

3).GMM doesn’t have to pair with the discount factor model ;while ML doesn’t have to do with the expected return-beta formulation.

4)ML requires an auxiliary statistical model;while GMM does not require such model.

2.Why ML is often ignored

1)ML plus the assumption of normal i.i.d , disturbances leads to time-series or cross-sectional regressions, empirical procedures that are close to the economic content of the model..

2)But asset returns are not normally distributed or i.i.d. (fatter tails, heteroskedastic, autocorrelated etc.)

3)ML specifies a GLS cross-sectional regressions, but we use OLS cross-sectional regressions instead,distrusting the GLS weighting matrix.

4)the questions about the null hypotheses

ML does not necessarily produce robust or easily interpretable estimates. It provides efficient estimates.

Time series regressions are almost universally run with constant, while ML does not.

- 3.OLS vs, GLS cross-sectional regressions

1) In the simple environment the choice between OLS vs GLS cross-sectional regressions is not important; but in more complex environment this is not the case.

2) Fama and French (1997) report important correlations between betas and pricing errors in a time-series test of a three-factor model on industry portfolios. OLS has no this capacity.

3) GLS and second-stage GMM gain their asymptotic efficiency when the covariance and spectral density matrices have converge to their population values.

Danger: Overfitting and model misspecification.

4) Kandel and Stambaugh (1995) and Roll and Ross (1995) think that GLS cross-sectional regressions also is model misspecification's result.

$$E(R^e) = \alpha + \lambda\beta$$

$$E(AR^e) = A\alpha + \lambda A\beta$$

$$\lambda = (\beta^{'}\beta)^{-1}\beta^{'}E(R^e)$$

$$\lambda = (\beta^{'}A^{'}A\beta)^{-1}\beta^{'}A^{'}AE(R^e)$$

$$\lambda = (\beta' \Sigma^{-1} \beta)^{-1} \beta' \Sigma^{-1} E(R^e)$$

$$\lambda = (\beta' A' (A \Sigma A')^{-1} A \beta)^{-1} \beta' A' (A \Sigma A')^{-1} A E(R^e)$$

$$= (\beta' \Sigma^{-1} \beta)^{-1} \beta' \Sigma^{-1} E(R^e)$$

How to choose between OLS and GLS

- Depends on what you want:
- Prove the model wrong, GLS helps you focus on the most informative portfolios for proving.
- However, many models are wrong, but still perform well for some portfolios.
- Compromise: OLS with GLS.

- 4.Additional examples of trading off efficiency for robustness

1)Low-frequency time series model: Minimize one step ahead forecast error variance.

$$\text{AR}(1): y_t = \rho y_{t-1} + \varepsilon_t$$

$$\text{ML: } \min E(\varepsilon_t^2)$$

- One may be interested in the long-run behavior of the short rate of interest, thus we need to know the sum of autocorrelations or moving average coefficient, in which case ML can make very bad predictions.

2) Lucas' money demand estimate.

$$m_t = a + by_t + \varepsilon_t$$

B=1, but the error term is strongly serially correlated. Researchers suggest to use:

$$m_t - m_{t-1} = b(y_t - y_{t-i}) + \eta_t$$

$$\eta_t = \varepsilon_t - \varepsilon_{t-1}$$

Error term pass DW test, however b is smaller, which does not make much economic sense and worse, is unstable, depending on data.

Reasons: difference throw out all of the information in the data.

- 5.Addressing model misspecification

1)The ML philosophy offers an answer to model misspecifications: specify the right model,and then do ML

2)The GMM framework allows us to evaluate misspecified models.

3)The GMM framework allows us to flexibly incorporate statistical model misspecifications in the distribution theory.

- 6.Auxiliary model
- ML requires an auxiliary statistical model. For example, we need to assume the returns and factors are jointly i.i.d. normal.
- A convenient feature of GMM is that it does not require such as auxiliary statistical model.

1) Finite-Sample Distribution

Many users argue that they prefer regression tests and GRS statistic since it has a finite sample distribution theory, and distrust the finte-sample performance of the GMM asymptotic distribution theory.

The finite-sample distribution only holds if returns really are normal and i.i.d,and if the factor is perfectly measured.

Once you have picked the estimation method, finding its Finite-Sample Distributions, given an auxiliary statistical model,is simple. For example, Mante Carlo or bootstrap. It is therefore nonsense to pick an estimation method just due to the finite sample distribution.

2)Finite-sample quality of asymptotic distributions, and “nonparametric” estimates

Several investigations have found cases in which GMM asymptotic distribution theory is a poor approximation to a finite-sample distribution theory. This could be true:

(i) Non-parametric correction for autocorrelation and heteroskedasticity-Solution: use parametric rather than nonparametric.

(ii) The moments one uses is very inefficient which could happen if you put a lot of instrument-solution: Use efficient moments.

- The case for ML

1)In the classic setup,the efficiency gain of ML over GMM on the pricing errors is tiny.

2)but several studies show that ML has important efficiency advantages.

- The case for ML

$$\frac{dS_t}{S_t} = \mu dt + V_t dZ_{1t}$$

$$dV_t = \mu_v(V_t)dt + \sigma(V_t)dZ_{2t}$$

- However, this study presumes the model is true. Whether ML perform better under some other DGP is open to discussion.

- 7.Statistical philosophy

1) It seems that "it takes a model to beat a model."

The CAMP was taught and believed in and used for years despite formal statistical rejection.

It only fell when coherent views of the world were offered by multifactor models.

The multifactor models are also rejected!

2)Even when evaluating a specific model, most of the calculations come from examining specific alternatives rather than overall pricing error tests.

The original CAMP focus on the intercept of regression, and whether individual variance and other characteristic enter cross section regression.

3)Statistical testing is one of many questions we ask in evaluating theories ,and often not the most important one.

- Fama and French (1988), (1993) prove that long-horizon returns are predictable and suggest the use of three factor models. However, they are rejected by GRS test.

8.Summary

1) It is ok to do a first-stage or simple GMM estimate rather than an explicit ML estimate and test .

2)A simple first-stage GMM approach , focusing on economically interpretable moments, can be adequate efficient, robust to model misspecifications, and ultimately more persuasive.