



Tinjauan Terhadap Metode Ekonometrika Lanjutan

Telisa Aulia F.

Keywords: Econometric, Stationarity Test, Panel Data

ABSTRACT

Econometric models have been played an increasingly important role in empirical analysis in economics. This paper provides an overview on some advanced econometric methods that increasingly used in empirical studies.

A panel data combines features of both time series and cross section data. Because of increasing availability of panel data in economic sciences, panel data regression models are being increasingly used by researcher. Related to panel data model, there are some methods that will be discussed here such as fixed effect and random effect. A new approach to panel data that developed by Im, Shin, and Pesaran (2002) for testing unit root in heterogenous panel is included in this overview.

When we work with time series data, there are many problems that we must handle, most of them are unit root test, cointegration among non stationary variables, and autoregressive conditional heteroscedasticity. Provided these problems, author also review about ADF and Philips-Perron test. An approach to cointegration analysis developed by Pesaran (1999), ARCH and GARCH model are also interesting to be discussed here.

Bayesian econometric, that less known than classical econometric, is included in this overview. The genetic algorithm, a relatively new method in econometric, has been increasingly employed the behavior of economic agents in macroeconomic models. The genetic algorithm is based on the process of Darwin's Theory of Evolution. By starting with a set of potential solutions and changing them during several iterations, the Genetic Algorithm hopes to converge on the most 'fit' solutions.

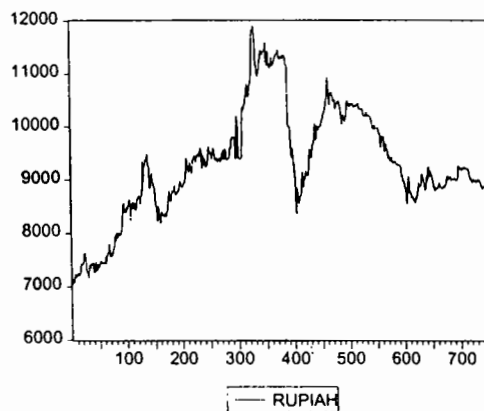
PENDAHULUAN

Ekonometrika adalah bagian yang tidak terpisahkan dari ilmu ekonomi. Perkembangan ilmu ekonomi didukung oleh perkembangan ekonometrika. Ekonometrika juga mengalami perkembangan dari waktu ke waktu. Semakin kompleksnya permasalahan dalam ilmu ekonomi menjadi tantangan bagi ekonometrika untuk terus berkembang. Perkembangan ini menarik penulis untuk melakukan tinjauan singkat atas metode-metode ekonometrika lanjutan (*advanced econometric*). Dengan tinjauan singkat ini diharapkan akan memberikan gambaran umum mengenai metode-metode ekonometrika tersebut. Dengan demikian diharapkan pembaca dapat tertarik untuk lebih lanjut mempelajari secara mendetail mengenai metode tersebut.

Dalam tinjauan ini, tidak semua metode ekonometrika akan dibahas. Ada beberapa metode yang akan dibahas dalam studi ini. Metode yang pertama adalah ARCH dan GARCH. Model ini banyak digunakan untuk *high frequency data* yang mengandung masalah korelasi serial dan heteroskedastisitas. Metode berikutnya adalah pengujian *unit root* untuk data *time series biasa* dan untuk data panel. Metode berikutnya yang tidak kalah menarik adalah *non linier least square*. Selama ini di tingkat ekonometrika dasar yang dipelajari adalah *linier least square*. Pada beberapa kondisi, metode *linier least square* tidak tepat digunakan sehingga kita memerlukan *non linier least square*. Metode kointegrasi terbaru yang dikembangkan oleh Pesaran (1999) yaitu *Autoregressive Distributed Lag (ARDL Model)* juga akan dibahas di sini. Tinjauan berikutnya mengenai pendekatan *fixed effect* dan *random effect* dalam regresi dengan data panel. Metode yang kurang populer seperti *Bayesian econometric* juga akan dibahas di sini. Perkembangan terbaru seperti *Genetic Algorithm* akan dibahas secara singkat dalam tinjauan ini.

MODEL ARCH DAN GARCH

Dalam model ekonometrik konvensional, *variance* dari *disturbance term* diasumsikan konstan. Namun, banyak data ekonomi yang berbentuk *time series* yang menunjukkan volatilitas yang besar pada periode tertentu dan diikuti oleh periode kestabilan. Contoh data-data yang menunjukkan perilaku demikian antara lain adalah nilai tukar dan indeks harga. Sebagai contoh adalah pergerakan nilai tukar rupiah terhadap dolar. Berikut ini grafik pergerakan nilai tukar rupiah terhadap dolar dengan data harian dari bulan Januari 2000 sampai dengan bulan Desember 2002. Observasi selama periode tersebut sekitar 750 observasi.



Dari grafik tersebut tampak bahwa terjadi volatilitas yang besar pada periode tertentu kemudian diikuti dengan periode kestabilan, dan seterusnya. Dalam kondisi ini asumsi bahwa varians konstan tidak tepat. Dengan model regresi :

$$y_t = x_t' \beta + \varepsilon_t \quad \dots\dots\dots (1)$$

Engle (1982) menunjukkan bahwa mungkin untuk memodelkan secara simultan *mean* dan *variance* dari suatu *series*. *Multiplicative conditionally heteroscedastic* model yang *proposed* oleh Engle adalah :

$$\varepsilon_t = vt\sqrt{ht} \quad \dots\dots\dots (2)$$

di mana $\sigma_v^2 = 1$

dan $ht = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2$

Conditional mean dan *variance* diberikan sebagai berikut :

$$E_{t-1} y_t = a_0 + a_1 y_{t-1} \quad \dots\dots\dots (3)$$

$$\begin{aligned} \text{var}(y_t | y_{t-1}, y_{t-2}, \dots) &= E_{t-1} (y_t - a_0 - a_1 y_{t-1})^2 \\ &= E_{t-1} (\varepsilon_t)^2 \\ &= \alpha_0 + \alpha_1 (\varepsilon_{t-1})^2 \quad \dots\dots\dots (4) \end{aligned}$$

Sedangkan solusi untuk y_t adalah :

$$y_t = a_0 / (1 - a_1) + \sum_{i=0}^{\infty} a_1^i \varepsilon_{t-i} \quad \dots\dots\dots (5)$$

Karena $E\varepsilon_t = 0$ untuk semua t , *unconditional expectation* dari persamaan di atas adalah :

$E y_t = a_0 / (1 - a_1)$. *Unconditional variance* untuk y_t adalah :

$$\text{Var}(y_t) = \sum a_1^{2i} \text{var}(\varepsilon_{t-i}) \quad \dots\dots\dots (6)$$

Point yang penting di sini adalah ARCH *error process* dapat digunakan untuk membuat model bagi periode volatilitas dalam kerangka *univariate*. Metode ini disebut dengan ARCH (*Autoregressive Conditional Heteroscedasticity*).

Bollerslev (1986) mengembangkan penemuan Engle dengan mengembangkan teknik yang memungkinkan *conditional variance* menjadi *ARMA process*. *Error process* sesuai dengan persamaan (2) yaitu : $\varepsilon_t = vt\sqrt{ht}$, di mana $\sigma_v^2 = 1$

dan $ht = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i h_{t-i}$

Karena vt adalah *white noise process* yang independen dari realisasi masa lalu dari ε_{t-i} , maka *conditional means* dan *unconditional means* dari ε_t akan sama dengan nol.

GARCH (p,q) mengijinkan baik untuk komponen *autoregressive* maupun *moving average* di dalam *heteroscedastic variance*. Keuntungan dari GARCH model adalah lebih mudah diestimasi untuk kasus ARCH model dengan ordo tinggi.

Karakteristik utama dari GARCH model adalah bahwa *conditional variance* dari *sequence* $[y_t]$ membentuk ARMA *process*.

Untuk mengestimasi GARCH dan ARCH model digunakan teknik *Maximum Likelihood Estimation*. Dengan teknik ini diharapkan akan didapatkan estimator yang secara asimtotik lebih efisien dibandingkan dengan estimator OLS.

METODE PENGUJIAN UNIT ROOT : ADF DAN PHILIPS PERRON TEST

Ketika membicarakan *series* yang stasioner dan tidak stasioner kita memerlukan tes yang menguji keberadaan *unit root* dalam rangka menghindari masalah *spurious regression*. Apabila suatu variabel mengandung *unit root*, maka regresi yang melibatkan variabel tersebut dapat mengimplikasikan hubungan ekonomi yang salah.

Ada beberapa cara untuk menguji keberadaan dari *unit root*. Salah satunya adalah dengan menggunakan *Dickey Fuller Test*. *Null hypothesis* dari *Dickey Fuller Test* adalah bahwa *series* mengandung *unit root*, sedangkan hipotesa alternatifnya adalah bahwa *series* bersifat stasioner. Bentuk paling sederhana dari *Dickey Fuller Test* adalah :

$$\Delta y_t = \mu + \gamma y_{t-1} + \varepsilon_t \quad \dots \dots \dots (7)$$

Hipotesa yang harus diuji adalah :

$$H_0 : \gamma = 0$$

$$H_0 : \gamma < 0$$

Apabila hipotesa nol diterima berarti *series* memang tidak stasioner, dengan kata lain mengandung *unit root*.

Augmented Dickey Fuller Test

Augmented Dickey Fuller Test merupakan pengembangan dari *Dickey Fuller Test*. Uji ADF menggunakan koreksi parameter untuk *higher order correlation* dengan mengasumsikan bahwa *series* y mengikuti proses AR (p) dan melakukan penyesuaian terhadap metodologi pengujian. Pendekatan ADF mengontrol untuk *higher order correlation* dengan menambah *lag* dari *difference term* untuk variabel dependen y di persamaan sebelah kanan dari regresi. Persamaan tersebut tampak dalam (8) :

$$\Delta y_t = \mu + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \delta_2 \Delta y_{t-2} + \dots + \delta_{p-1} \Delta y_{t-p-1} + \delta_p \Delta y_{t-p} + \varepsilon_t \dots (8)$$

Augmented specification kemudian digunakan untuk menguji :

$$H_0 : \gamma = 0$$

$$H_1 : \gamma < 0$$

Jika H_0 diterima, maka memang *series* tersebut mengandung *unit root*.

Philips Perron Unit Root Test

Philips and Perron (1988) menggunakan metode non parametrik dalam mengontrol *higher order serial correlation* dalam *series*. Regresi yang diuji oleh PP test adalah proses AR(1) tampak dalam persamaan (9) :

$$\Delta y_t = \alpha + \beta y_{t-1} + e_t \quad \dots \dots \dots (9)$$

Jika ADF tes mengoreksi untuk *higher order correlation* dengan menambahkan *lag* dari *differenced terms* di sisi kanan persamaan, PP test membuat koreksi terhadap t statistik dan koefisien γ dari regresi dengan AR(1) untuk memperhitungkan korelasi serial dalam c .

Koreksi bersifat non parametrik karena kita menggunakan estimasi dari spektrum e pada frekuensi nol dan *robust* terhadap heteroskedastisitas dan autokorelasi dari bentuk yang tidak diketahui. Dalam mengestimasi digunakan *Newey-West heteroscedasticity autocorrelation consistent estimate* :

$$\omega^2 = \gamma_0 + 2 \sum_{j=1}^q \left(1 - \frac{j}{q+1}\right) \gamma_j$$

di mana :

$$\gamma_j = 1/T \sum_{t=j+1}^T \hat{e}_t \hat{e}_{t-j}$$

Nilai PP t statistik dihitung sebagai berikut :

$$t_{pp} = \frac{\gamma_0^{1/2} t_b}{\omega} - \frac{\omega^2 - \gamma_0 T s_b}{2\omega \hat{\sigma}}$$

t_b adalah t statistik

s_b adalah *standard error* dari β

$\hat{\sigma}$ adalah *standard error* dari persamaan regresi yang diuji

Asymtootic distribution dari PP t *test statistics* sama dengan ADF *test statistics*. Nilai kritis dalam program yang digunakan sama dengan yang digunakan dalam ADF *test statistics* yaitu *MacKinnon critical values*. Dalam tes ini juga harus diputuskan apakah akan memasukkan *intercept* dan *trend* atau tidak, seperti halnya dalam uji ADF. Dalam PP *test* akan dipilih *truncation lag* (q) untuk menentukan *Newey West Correction*, untuk menentukan berapa jumlah periode yang dimasukkan untuk menghindari masalah *serial correlations*.

MENGUJI UNIT ROOT PADA DATA PANEL

Data panel merupakan data gabungan antara *cross section* dan *time series*. Seperti diketahui, data *time series* bisa mengandung *unit root*. Pengujian *unit root* untuk data *time series* biasa tidak bisa diterapkan pada data panel. Apabila data panel mengandung *unit root* maka ADF *test* dan Philips Perron *test* biasa yang disebutkan sebelumnya tidak bisa digunakan. Ada beberapa tes yang dikembangkan untuk menguji ada tidaknya *unit root* pada data panel, misalkan IPS *test* atau Levin-Lin *test*.

Motivasi utama dari panel data *unit root test* adalah untuk meningkatkan *power* dari *unit root test* dengan meningkatkan ukuran sampel. Namun alternatif dengan menggunakan deret data yang lebih panjang menyebabkan masalah karena munculnya perubahan struktural. Namun memang tidak terlalu jelas apakah masalah akibat *long time series data* lebih parah dari masalah heterogenitas karena menggunakan data panel dan struktural. *Unit root test* yang sering digunakan, yaitu uji ADF, tidak terlalu *powerful*, dan dengan menggunakan data panel kita akan mendapatkan tes yang lebih *powerful*.

Im Pesaran Shin test (IPS test) merupakan *unit root test* untuk *dynamic heterogenous panels* berdasarkan rata-rata dari individual *unit root statistics*. Secara khusus, tes ini mengusulkan *standardized t bar statistics* berdasarkan DF/ADF statistik yang dirata-ratakan di antara grup. Statistik ini dapat ditunjukkan konvergen dalam probabilita terhadap standard normal secara sekuensial seiring dengan T (*time series dimension*) menuju tak hingga, yang diikuti dengan N (*cross sectional dimension*) menuju tak hingga.

Konvergensi diagonal antara T dan N akan menuju tak hingga sedangkan N/T menuju k, di mana k adalah *finite non negative* konstan.

Dalam kasus khusus di mana *error* dalam *individual Dickey Fuller regression* adalah *serrially uncorelated*, versi t yang dimodifikasi yaitu *standardized t bar statistic*, yang dilambangkan dengan Z_{tbar} , akan terdistribusi dengan standar normal ketika N menuju tak hingga untuk suatu T yang tetap (*fixed*), selama $T > 5$ dalam kasus regresi DF dengan intersep dan $T > 6$ untuk kasus regresi DF dengan intersep dan *linier time trends*. Tes untuk seberapa *fixed* N dan T juga dikembangkan dengan menggunakan rata-rata dari DF *statistics*. Hasil Monte Carlo menunjukkan bahwa jika ordo yang cukup besar yang dipilih dalam ADF regression, maka *small sample performance* dari t bar *test* sangat memuaskan dan lebih baik daripada uji yang dilakukan oleh Levin Lin.

Panel based unit root telah dikembangkan oleh Quah (1992, 1994) dan Levin Lin (1993). Tes yang dikembangkan oleh Quah tidak mengakomodasi heterogenitas antar kelompok seperti efek spesifik individu dan pola yang berbeda dari *residual serial correlations*. Levin Lin *test* lebih umum untuk diaplikasikan, dan sudah mengakomodasi heterogenitas antar kelompok dan mensyaratkan N/T menuju nol ketika N dan T menuju tak hingga.

Dengan menggunakan kerangka *likelihood*, *IPS test* dihitung berdasarkan rata-rata dari DF/ADF *test* untuk setiap grup dalam panel, ini disebut sebagai t bar *test*. Seperti prosedur Levin Lin, prosedur ini mengakomodasi *residual serial correlations* dan efek spesifik individu dan *error variance* di antara kelompok. Tes untuk seberapa *fixed* N dan T (*exact test*) juga dikembangkan dengan menggunakan rata-rata dari DF *statistics*. Berdasarkan *stochastic simulations*, ditunjukkan bahwa Z_{tbar} menyediakan pendekatan yang bagus untuk *exact test* meskipun untuk N yang relatif kecil.

METODE NON LINIER LEAST SQUARE

Selama ini di tingkat ekonometrika dasar kita terfokus pada model linier. Untuk beberapa kasus aplikasi, spesifikasi non linier dalam parameter tidak dapat dihindari. Contohnya untuk kasus fungsi produksi CES (*Constant Elasticity of Substitution*). Fungsi produksi CES memiliki bentuk sebagai berikut :

$$Y = b1*(b2*L^b3 + (1-b2)*K^b3)^(b4/b3).....(10)$$

Fungsi produksi CES ini jelas berbentuk tidak linier dalam parameter. Misalkan Fungsi produksi CES secara umum memiliki bentuk :

$$Y=f(xt,\beta) + et.....(11)$$

Dengan *non linier least squares* berarti ingin mencari estimator b yang meminimumkan:

$$S(\beta)=[y-f(\beta)]'[y-f(\beta)]..... (12)$$

Sehingga *First Order Conditions* adalah :

$$\frac{\partial S}{\partial \beta} \Big|_b = -2Z(b)'[y - f[b]] = 0$$

$$\text{di mana } Z(\beta) = \frac{\partial f}{\partial \beta'} \Big|_{\beta} = \begin{bmatrix} \frac{\partial f_1}{\partial \beta_1} \Big|_{\beta, \dots} & \frac{\partial f_k}{\partial \beta_k} \Big|_{\beta} \\ \frac{\partial f_T}{\partial \beta_1} \Big|_{\beta, \dots} & \frac{\partial f_T}{\partial \beta_k} \Big|_{\beta} \end{bmatrix}$$

adalah matriks TxK yang terdiri dari turunan pertama dari f terhadap parameter.

Dalam kasus *non linier least square*, b secara umum adalah penaksir yang bias. Hanya dalam kondisi khusus, *asymptootic properties* dari b dapat diturunkan berdasarkan Amemiya (1983).

$$\beta_{n+1} = \left[z(\beta_n)' z(\beta_n) \right]^{-1} z(\beta_n)' [y - f(x, \beta_n) + z(\beta_n) \cdot \beta_n] \dots \dots \dots (13)$$

$$\beta_{n+1} = \beta_n \left[z(\beta_n)' z(\beta_n) \right]^{-1} z(\beta_n)' [y - f(x, \beta_n)] \dots \dots \dots (14)$$

Persamaan (13) merupakan *Gauss Newton Algorithm*. FOC untuk minimum mensyaratkan $z(\beta)' [y - f(x, \beta)] = 0$. Jika memenuhi FOC persamaan (13) menjadi :

$$\beta_{n+1} = \beta_n \dots \dots \dots (14)$$

Artinya pada kondisi tersebut konvergensi sudah tercapai. Masalah konvergensi terkait dengan algoritma yang dipakai. Untuk membedakan algoritma maka persamaan (15) yang merupakan bentuk lain persamaan (13) perlu diperhatikan :

$$\beta_{n+1} = \beta_n - t \cdot p(n) \cdot \gamma(n) \dots \dots \dots (15)$$

t menggambarkan panjangnya langkah

$\gamma(n)$ menggambarkan gradien

$t \cdot p(n) \cdot \gamma(n)$ menggambarkan *step size*

$p(n)$ menggambarkan algoritma. Dengan $p(n)$ yang berbeda-beda berarti algoritma yang digunakan juga berbeda-beda. Program berbeda-beda karena algoritma yang digunakan berbeda-beda.

Ada beberapa jenis algoritma antara lain:

- Gauss-Newton
- Marquardt-Levenberg
- Steepest Descent
- Newton-Raphson
- Quadratic Hill Climbing

MODEL AUTOREGRESSIVE DISTRIBUTED LAG

Efek dinamis dapat diinvestigasi dengan cara memasukkan nilai lag dari independen variabel di sisi kanan dari model regresi. Dalam model dinamis, dapat dilakukan dengan cara *distributed lag* atau *autoregressive model*. Kombinasi di antara kedua model ini

disebut dengan *Autoregressive Distributed Lag* (ARDL). Model ARDL telah digunakan secara luas dalam model-model yang menggunakan data *time series*.

Model ARDL digunakan untuk menganalisis hubungan jangka panjang ketika variabel penjelas tidak stasioner atau I(1). Estimator ARDL akan menghasilkan koefisien jangka panjang yang super konsisten. Salah satu keunggulan dari *ARDL approach* adalah menghasilkan estimasi yang konsisten dengan koefisien jangka panjang yang secara asimtotik normal walaupun regresor I(0) ataupun I(1). Dalam kasus di mana hubungan jangka panjang dari yang kita teliti bersifat *trend stationarity*, dengan ARDL dapat dilakukan *detrending* terhadap *series* dan memodelkan *detrended series* sebagai *distributed lag* yang stasioner. Estimasi dan inferensi yang berhubungan dengan *long-run properties* dari model dapat dijalankan dengan menggunakan teori *standard asymptotic normal*.

Literatur terkini yang membahas kointegrasi berhubungan dengan analisa hubungan jangka panjang antara variabel I(1), berdasarkan literatur tersebut *traditional ARDL approach* menjadi tidak lagi dapat diaplikasikan. Sebagai konsekuensinya, berkembang prosedur alternatif yang mengembangkan model untuk variabel I(1), antara lain Engle and Granger (1987), Philips dan Hansen (1990) dan Philips and Loretan(1991).

Keunggulan dari ARDL adalah estimator dari koefisien jangka panjang yang bersifat super konsisten dan inferensi yang valid terhadap koefisien jangka panjang dapat dibuat dengan menggunakan *standard normal asymptotic theory*. Berdasarkan *Monte Carlo experiment* dari Gerrard dan Godfrey (1998) mengindikasikan bahwa pendekatan ARDL lebih baik digunakan dalam mengestimasi koefisien jangka panjang dari hubungan kointegrasi.

Sekumpulan variabel yang termasuk dalam ARDL model dapat diikuti dengan deterministik regresor seperti *intercept term*, *time trend*, dan regresor dengan *fixed lag*. Dalam spesifikasi ARDL harus ditentukan *maximum lag ordo-nya*.

Berdasarkan Pesaran dan Shin (1999), model umum ARDL(p,q) diberikan oleh persamaan (16) sebagai berikut :

$$y_t = \alpha_0 + \alpha_1 t + \sum_{i=1}^p \phi_i y_{t-i} + \hat{a}' x_t + \sum_{i=0}^{q-1} \beta_i * \Delta x_{t-i} + u_t$$

$$\Delta x_t = P_1 \Delta x_{t-1} + P_2 \Delta x_{t-2} + \dots + P_s \Delta x_{t-s} + \varepsilon_t \dots \dots \dots (16)^1$$

Di mana x_t adalah k-dimensional, yang merupakan *forcing variable* I(1) yang tidak terkointegrasi di antara mereka sendiri. u_t dan ε_t adalah *serially uncorrelated disturbance* yang memiliki rata-rata nol dan *constant variance-covariance*, dan P_i adalah k X k koefisien matriks sedemikian rupa sehingga *vector autoregressive process* dalam Δx adalah stabil. Dalam kasus *bivariate cointegration*, kita menetapkan k=1. Dengan menetapkan k=1 kita menghindari masalah kointegrasi di antara "forcing variable" x_t ,

Dalam formulasi di atas $\alpha_0, \alpha_1, \beta_1^*, \dots, \beta_{q-1}^*, \phi = (\phi_1, \dots, \phi_p)$ adalah parameter jangka pendek yang penting dalam mengestimasi koefisien jangka panjang yang didefinisikan dengan rasio $\delta = \alpha_1 / \phi(1)$ dan $\theta = \beta / \phi(1)$, dimana $\phi(1) = 1 - \sum \phi(i)$. Pendekatan ARDL juga mengasumsikan adanya hubungan jangka panjang yang stabil di antara dua variabel X dan

¹ M. Hashem Pesaran and Yongcheol Shin, "An Autoregressive Distributed Lag Modelling Approach to Cointegration Analysis", dipresentasikan dalam *Simposium di Centennial of Ragnar Frisch, The Norwegian Academy of Science and Letters* 3-5 Maret 1995, chapter 11, Cambridge University Press, 1999, hlm. 1-2.

Y. Dalam kasus di mana u_t dan ϵ_t berkorelasi, spesifikasi ARDL di atas ditambah dengan lag perubahan dalam regresor dengan jumlah lag yang cukup. Derajat dari penambahan ini tergantung pada apakah $q > s+1$ atau tidak. *Augmented model* ini menjadi :

$$y_t = \alpha_0 + \alpha_1 t + \sum_{i=1}^p \phi_i y_{t-i} + \hat{a}'x_t + \sum_{i=0}^{m-1} \pi_i * \Delta x_{t-i} + \eta_t \dots\dots\dots(17)$$

di mana $m = \max(q, s+1)$, $\pi_i = \beta_i * \rho^i$, di mana ρ adalah 1×1 vektor yang berisi *contemporaneous correlation* antara u_t dan ϵ_t . Jadi, pendekatan ARDL mensyaratkan untuk memasukkan lag yang cukup dari *forcing variables* x_t dalam rangka mengendogenisasikan y_t . Dengan melakukan hal ini, masalah *endogenous regressor* dan *serial autocorrelation* dapat secara simultan dikoreksi.

Model ARDL yang lebih sederhana dapat dinyatakan sebagai :

$$\phi(L)y_t = \alpha_0 + \alpha_1 w_t + \beta'(L)x_t + u_t \dots\dots\dots(18)$$

di mana L adalah lag operator dan w_t adalah $s \times 1$ vektor dari variabel deterministik seperti intersep, tren, variabel *dummy*, dan variabel eksogen dengan *fixed lag*.

Dengan ARDL kita bisa mengestimasi *autoregressive distributed lag* (ARDL) model dengan ordo $(p, q_1, q_2, \dots, q_k)$, di mana p adalah ordo dari *distributed lag polynomial* dari variabel dependen atau variabel yang ada di sebelah kiri, dan q_1, q_2, \dots, q_k adalah ordo dari *distributed lag polynomials* dari regresor x . Model juga dapat terdiri dari deterministik atau regresor lain dengan *fixed lag distributions*. Pilihan ini menawarkan prosedur alternatif (seperti kriteria AIC dan SBC) untuk seleksi dari ordo lag p, q_1, q_2, \dots, q_k . Implikasi dari ARDL model juga adalah estimasi *Error Correction Model*. Seleksi model antara lain dengan AIC atau SBC. *Akaike Information Criterion* biasanya digunakan dalam seleksi model yang bersifat *non nested*. Makin kecil nilai AIC semakin baik. Untuk menentukan spesifikasi dari lag maka digunakan AIC, yaitu dengan memilih nilai AIC yang terendah. *Schwarz Bayesian Criterion* memiliki fungsi yang sama dengan AIC. Hanya saja SBC memberikan penalti yang lebih besar untuk tambahan koefisien.

Ordo dari *distributed lag function* terhadap y_t dan *forcing variable* x_t dipilih dengan menggunakan *Akaike Information Criterion* (AIC) atau *Schwartz Bayesian Criterion* (SBC). Dengan menetapkan *maximum orders* dari p dan q sama dengan 12 (data bulanan), kita membandingkan nilai maksimum dari fungsi *log likelihood* dari $(m+1)^{k+1} = (m+1)^2$ dari model ARDL yang berbeda. Kita memilih model akhir dengan menemukan nilai p dan q yang mengoptimumkan kriteria seleksi yang sudah disebutkan di atas. Ketika model sudah terpilih, digunakan OLS untuk mendapatkan parameter jangka pendek. Kemudian, kita mengestimasi koefisien jangka panjang dari hubungan kointegrasi $y_t = a + \delta_t + \theta x_t + v_t$ (dari persamaan 17) dengan :

$$\hat{a} = [\hat{a}_0 / (1 - \phi_1 - \dots - \phi_p)]$$

$$\hat{\delta} = [\hat{a}_1 / (1 - \phi_1 - \dots - \phi_p)]$$

$$\hat{\theta} = [(\beta_0 + \beta_1 + \dots + \beta_q) / (1 - \phi_1 - \dots - \phi_p)]$$

Koefisien jangka panjang menangkap efek dari *lag variable* terhadap hubungan kointegrasi. Pesaran dan Shin (1998) menunjukkan bahwa ARDL estimator dari koefisien jangka panjang adalah super konsisten dan inferensi *valid* dari estimator ini dapat

digunakan dengan menggunakan *standard normal asymptotic*. *Standard error* dari penduga untuk koefisien jangka panjang dapat diperoleh dengan *delta method*.

Jika memang terdapat hubungan kointegrasi, setelah mengestimasi koefisien jangka panjang, kemudian kita ingin mendapatkan *error correction representation* dari model ARDL (p,q) yang sudah dipilih. *Error correction model* adalah model ARDL yang sudah dipilih diperluas dengan *error correction term*. *Error correction term* ini dihitung berdasarkan ARDL, berdasarkan koefisien jangka panjang yang sudah diestimasi. Representasi *error correction* dari ARDL (p,q) yang telah dipilih tampak dalam persamaan (19) berikut :

$$\Delta y_t = \delta_0 + \delta_1 t + \sum_{i=1}^p d_i \Delta y_{t-i} + \sum_{i=1}^m g_i \Delta x_{t-i} + \gamma ect_{t-1} + \eta_t \dots \dots \dots (19)$$

Di mana $m = \max(q, s+1)$ dan ect_{t-1} adalah *lagged error correction term* yang diestimasi dengan menggunakan pendekatan ARDL. Koefisien dari ect_{t-1} diekspektasikan bernilai negatif dan signifikan secara statistik.

Berdasarkan Pesaran dan Shin, ketika ordo dari ARDL telah terbentuk, maka estimasi dan identifikasi dapat dilakukan dengan OLS. Karena adanya hubungan jangka panjang antar variabel tetap menjadi masalah yang penting maka Pesaran dan Shin menunjukkan bahwa *valid asymptotic inferences* dari parameter jangka pendek dan jangka panjang dapat dibuat untuk *least square estimate* dari model ARDL. Hal tersebut terjadi jika ordo dari model ARDL secara tepat ditambah untuk mengijinkan *contemporaneous correlations* antara komponen stokastik dari proses menghasilkan data yang termasuk dalam estimasi. Sehingga, estimasi dengan ARDL valid dapat digunakan meskipun variabel penjelas bersifat endogen dan karena keberadaan dari hubungan jangka panjang independen dari apakah variabel penjelas $I(1)$ atau $I(0)$. Jadi, metode ARDL memiliki keunggulan tidak membutuhkan identifikasi yang persis mengenai ordo integrasi dari data yang bersangkutan.

Langkah-langkah dalam estimasi dengan model ARDL adalah :

- Mengestimasi dan menganalisa model ARDL termasuk seleksi model dan melakukan *diagnostic test* untuk menguji ada atau tidaknya pelanggaran asumsi sebelum melangkah ke prosedur berikutnya
- Membuat *error correction model* berdasarkan model yang telah dipilih dan menguji ada / tidaknya hubungan kointegrasi jangka panjang dengan *Wald test*
- Menganalisa hasil output dari *error correction model* untuk mengetahui dinamika jangka pendek (*short run*)
- Menganalisa *long run coefficient* dari model ARDL

REGRESI DENGAN DATA PANEL : *FIXED EFFECT* DAN *RANDOM EFFECT*

Kelebihan data panel dari data *cross section* dan *time series* adalah: ²

1. Estimasi data panel dapat mempertimbangkan heterogenitas dengan memperkenalkan variabel-variabel individu spesifik.
2. Data panel dapat memberikan data yang lebih informatif, lebih bervariasi, kurang kolinearitas antar variabel, derajat bebas yang lebih besar, dan lebih efisien.
3. Data panel lebih sesuai untuk mempelajari dinamika perubahan (*dynamics of change*).
4. Data panel dapat secara lebih baik mendeteksi dan mengukur efek yang tidak dapat diamati dalam data *cross section* dan *time series*.

² Damodar Gujarati, *Basic Econometrics*, (New York: McGraw Hill, 2003).

5. Data panel dapat digunakan untuk mempelajari model-model perilaku yang kompleks.
6. Data panel dapat meminimalisir bias yang mungkin ditimbulkan oleh agregasi data individu.

Terdapat dua pendekatan dalam mengestimasi model regresi data panel, yaitu pendekatan *fixed effect* dan *random effect*.

Pendekatan Fixed Effect

Dalam *fixed effect*, bentuk umum regresi data panel:

$$Y_{it} = \beta_1 + \beta_2 X_{2it} + \beta_3 X_{3it} + \dots + \beta_n X_{nit} + u_{it} \dots\dots\dots(20)$$

dimana:

- i = banyaknya unit cross section ke i
- t = banyaknya periode observasi ke t
- n = banyaknya variabel

β_i disebut sebagai *individual effect* yang dianggap konstan sepanjang waktu *t* dan tertentu untuk unit *cross section* individu *i*.

Estimasi persamaan (20) tergantung pada asumsi mengenai intersep, koefisien kemiringan, dan galat yang dibuat. Ada beberapa kemungkinan yang dapat dibuat:

1. Asumsi bahwa koefisien *intercept* dan kemiringan konstan sepanjang waktu dan ruang dan galat menangkap selisih sepanjang waktu dan antar individu.
2. Koefisien kemiringan konstan tetapi *intercept*-nya berubah antar individu.
3. Koefisien kemiringan konstan tetapi *intercept*-nya berubah antar individu dan waktu.
4. Semua koefisien (baik *intercept* maupun kemiringan) berubah antar individu.
5. Koefisien *intercept* dan kemiringan berubah antar individu dan waktu.

Masing-masing kasus mempunyai kompleksitasnya sendiri dalam mengestimasi model regresi data panel, apalagi bila semakin banyak regressor yang digunakan dalam model karena semakin besar kemungkinan terjadinya kolinearitas antar regresor.

Pendekatan Random Effect

Dalam pendekatan *random effect*, bentuk umum regresi data panel:

$$Y_{it} = \beta_{1i} + \beta_2 X_{2it} + \beta_3 X_{3it} + \dots + \beta_n X_{nit} + u_{it} \dots\dots\dots(21)$$

dimana β_{1i} tidak lagi dianggap *fixed* tetapi dianggap sebagai variabel random dengan nilai rata-rata β_1 . Dengan demikian nilai intersep untuk masing-masing individu *i* dapat dinyatakan sebagai:

$$\beta_{1i} = \beta_1 + \epsilon_i \dots\dots\dots(22)$$

dimana ϵ_i adalah galat random dengan nilai rerata nol dan varians σ_c^2 .

Substitusi persamaan (12) ke (11):

$$Y_{it} = \beta_1 + \beta_2 X_{2it} + \beta_3 X_{3it} + \dots + \beta_n X_{nit} + w_{it} \dots\dots\dots(23)$$

dimana:

1. $w_{it} = \varepsilon_{it} + u_{it}$. Galat w_{it} terdiri atas dua komponen, yaitu ε_{it} yang merupakan komponen galat dari individu spesifik atau *cross section*, dan u_{it} yang merupakan komponen galat gabungan dari *time series* dan *cross section*.
2. $E(w_{it}) = 0$
3. $var(w_{it}) = \sigma_\varepsilon^2 + \sigma_u^2$
4. Dari butir (3) di atas, galat w_{it} adalah homoskedastis. Namun untuk $t \neq s$, w_{it} dan w_{is} berkorelasi—yaitu galat untuk unit *cross section* pada dua titik waktu yang berbeda berkorelasi. Sehingga koefisien korelasinya dapat dinyatakan sebagai:

$$corr(w_{it}, w_{is}) = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_u^2}. \text{ Ada dua hal mengenai koefisien korelasi.}$$

Pertama, untuk unit *cross section* tertentu, nilai korelasi antara galat pada dua waktu yang berbeda adalah tetap. Kedua, struktur korelasi tetap sama untuk semua unit *cross section*—artinya identik untuk semua individu.

Fixed Effect atau Random Effect

Pertanyaan berikutnya adalah bagaimana memilih pendekatan—*fixed* atau *random effects*— yang akan digunakan dalam mengestimasi regresi data panel.

1. Jika banyaknya data *time series* besar dan banyaknya unit *cross section* sedikit, maka tidak terdapat perbedaan nilai dalam mengestimasi parameter dengan pendekatan *fixed* maupun *random*. Keputusannya dengan demikian lebih pada aspek kemudahan dan kenyamanan peneliti. *Fixed effect* lebih diutamakan.
2. Jika banyaknya unit *cross section* besar dan banyaknya data *time series* juga besar, penggunaan kedua pendekatan tersebut akan membawa hasil yang berbeda. Inferensi statistik bersifat kondisional berdasarkan unit *cross section* dalam sampel yang diamati. Jika diyakini bahwa unit *cross section* yang diambil dari sampel yang lebih besar tidak bersifat *random*, maka pendekatan *fixed effect* lebih tepat. Sebaliknya jika unit *cross section* diambil secara *random*, maka pendekatan yang lebih tepat adalah *random effect*.
3. Jika komponen galat ε_{it} dan satu atau salah satu dari regressor berkorelasi, maka estimasi dengan *fixed effect* tidak bias.
4. Jika banyaknya unit *cross section* besar dan banyaknya data *time series* sedikit, dan apabila asumsi *random effect* terpenuhi, maka estimasi dengan menggunakan pendekatan *random effect* menjadi lebih efisien daripada pendekatan *fixed effect*.

Kondisi yang tepat untuk *random effect* adalah jika kita yakin bahwa unit *cross section* yang kita ambil sebagai sampel diambil dari suatu populasi yang besar. Sedangkan jika kita yakin bahwa hampir semua atau semua individu sudah terambil maka kita menggunakan *fixed effects*, karena pengambilan sampel seperti itu tidak bisa disebut *random*. Namun kadang-kadang pedoman ini tidak bisa selalu digunakan.

Untuk itu perlu dilakukan *Hausman test*. Uji ini dilakukan untuk menguji ortogonalitas dari *random effect* dan regressor. Hipotesa nol dari tes ini adalah bahwa antara estimasi *fixed effect* dan *random effect* tidak akan berbeda secara statistik. Untuk menguji hipotesa ini digunakan *Wald statistics*.

$$W = \chi^2[K] = [b - \beta^*]' \Sigma^{-1} [b - \beta^*] \dots\dots\dots(24)$$

Di mana:

- Σ = Var[b] – Var [β^*]
- b = estimasi dari *fixed effect*
- β^* = adalah estimasi dari *random effect*
- var(β^*)= ini sudah mengeluarkan *constant term*

W secara asimtotik terdistribusi sebagai *Chi Square* dengan derajat kebebasan sebesar K. Hipotesa nol adalah bahwa *individual effect* tidak berkorelasi dengan variabel independen. Jika hipotesa nol ditolak berarti *individual effect* berkorelasi dengan variabel independen. Jika hipotesis nol ditolak, maka kesimpulan yang dapat diambil adalah estimasi dengan *random effect* tidak sesuai sehingga akan lebih baik jika menggunakan *fixed effect*—dimana inferensi statistika akan bersifat kondisional berdasarkan ε_i dalam sampel.

BAYESIAN ECONOMETRICS

Salah satu karakteristik penting dari analisa Bayesian adalah ketidakpastian mengenai nilai dari parameter yang tidak diketahui dapat diekspresikan dalam suatu distribusi probabilitas. Seperti apapun nilai parameter maka parameter tersebut dapat memiliki suatu *probability density function*. Hal ini berbeda dengan kerangka klasik. Dalam kerangka klasik, parameter bersifat *fixed* dan *probability distribution* tidak dapat di-*assigned* terhadap parameter, atau secara sederhana memiliki *trivial distribution* di mana probabilita sama dengan satu pada *true parameter value* dan nol untuk parameter lainnya.

Bayesian subjective probability distribution terhadap suatu parameter merangkum pengetahuan individu mengenai parameter. Pengetahuan ini dapat ada sebelum mengobservasi informasi tentang sampel, dalam kasus ini distribusi disebut sebagai *prior distribution*. Jadi dalam *prior distribution* terdapat *judgement* peneliti mengenai bentuk distribusi dari parameter. Pengetahuan juga bisa diperoleh dari *prior* dan *sample information* yang disebut sebagai *posterior distribution*. *Sample information* adalah informasi yang diperoleh dari data sampel. Prosedur yang mengkombinasikan *prior distribution* dengan *sample information* untuk membentuk *posterior distribution* disebut Bayes' theorem.

Dengan *Bayesian regression* kita tidak terjebak untuk menemukan apakah suatu parameter itu positif atau negatif. Dalam *classical regression* kita biasanya memaksakan diri untuk mendapatkan apakah suatu parameter itu positif atau negatif disesuaikan dengan teori. Dalam *Bayesian regression* estimasi titik menjadi tidak penting karena kita akan mendapatkan keseluruhan bentuk distribusi dari parameter yang kita cari. Inti dari *Bayesian econometrics* bukanlah mengestimasi titik, tetapi mengetahui bentuk distribusinya secara keseluruhan sehingga kita tidak akan terjebak pada hasil estimasi titik yang bisa *misleading* dan bahkan kadang tidak sesuai dengan teori.

Prior distribution dapat berbentuk *informative prior* dan dapat berbentuk *uninformative prior*. Ada tiga alasan utama mengapa kita menggunakan *noninformative prior* (*uninformative prior*):

- Ada kondisi di mana kita hanya memiliki sedikit informasi tentang parameter sehingga kita membutuhkan *prior distribution* yang akan menggambarkan kekurangan ini

- Meskipun kita memiliki *prior information*, ada suatu kebiasaan di mana akan lebih objektif jika menyajikan *posterior distribution* yang hanya menggambarkan *sample information* dibandingkan dengan adanya bias dari pandangan kita sendiri
- Seringkali sangat sulit untuk memformulasikan distribusi *informative prior* yang tepat, sehingga untuk keluar dari masalah ini digunakan *noninformative prior*

Bentuk dari *uninformative prior* dari β dan σ adalah :

$$g(\beta, \sigma) \propto \sigma^{-1} \dots\dots\dots(25)$$

Dengan *uninformative prior* tersebut, maka *posterior density* menjadi :

$$g(\beta, \sigma | y) \propto g(\beta, \sigma) \ell(\beta, \sigma | y) \dots\dots\dots (26)$$

$$g(\beta, \sigma | y) \propto \sigma^{-(T+1)} \exp \{-1/2\sigma^2 [(T-K)\delta^2 + (\beta-b)'X'X(\beta-b)]\} \dots\dots\dots (27)$$

Joint density function ini merangkum semua informasi tentang *post sample information* tentang β dan σ .

Untuk mengkarakterisasi $g(\beta, \sigma | y, x)$ diperlukan *Markov Chain Monte Carlo*. MCMC menjadi mempermudah proses ini. Untuk menjalankan MCMC kita perlu mempunyai $g(\beta | \sigma)$ dan $g(\sigma | \beta)$. Untuk mendapatkan $g(\sigma | \beta)$ memanfaatkan bahwa :

$$\frac{y'My}{\sigma^2} \approx \chi^2_T$$

$g(\sigma | \beta) \sim \chi^2_T$ diciptakan secara acak

Metode *Bayesian econometrics* ini memang kurang populer dibandingkan dengan *classical econometrics* yang biasa kita pelajari.³ Hal ini disebabkan karena metode ini lebih sulit secara teknis matematis. Untuk aplikasinya dalam *software* siap pakai – pun masih jarang dan lebih sulit untuk dipahami karena kita bekerja dengan bentuk distribusi yang memang sulit untuk dipelajari bagi orang umum. Karena hal tersebut maka tidak heran jika *Bayesian Econometrics* jarang dipelajari secara lebih mendalam.

PENDEKATAN *GENETIC ALGORITHM*

Genetic Algorithm didasarkan pada proses teori evolusi dari Darwin. Dimulai dengan sekumpulan solusi yang mungkin dan mengubahnya melalui beberapa iterasi maka diharapkan akan konvergen pada solusi yang paling fit. Proses dimulai dengan sekumpulan solusi yang mungkin atau kromosom (dalam bentuk *bit strings*) yang secara random dipilih atau dihasilkan. Sekumpulan kromosom ini membentuk suatu *population*. Kromosom yang terlibat selama proses iterasi disebut sebagai *generation*. Generasi yang baru dihasilkan melalui teknik *crossover* dan *mutation*. Kromosom kemudian dievaluasi menggunakan kriteria *fitness* dan yang terbaik akan disimpan sedangkan yang lain akan dibuang. Proses ini berulang sampai satu kromosom memiliki *fitness* yang terbaik dan diambil sebagai solusi terbaik dari problem atau *objective* kita.

³ Penulis mengucapkan terimakasih kepada Dr. Muhammad Syamsudin, Dosen Ekonometrika Program Pasca Sarjana Ilmu Ekonomi, yang telah memperkenalkan metode *Bayesian Econometric*.

Ada beberapa keuntungan dari *Genetic Algorithm* :

- Bekerja dengan baik untuk *global optimization*, khususnya ketika *objective function* bersifat diskontinu atau memiliki beberapa *local minima*.
- Dapat digunakan baik untuk masalah optimisasi tanpa kendala maupun dengan kendala
- Dapat diaplikasikan pada *non linier programming* dan *stochastic programming* Namun di sisi lain memang ada kelemahan dari *Genetic Algorithm* antara lain:
- Karena tidak menggunakan informasi tambahan seperti gradien maka *Genetic Algorithm* memiliki tingkat pencapaian konvergensi yang lambat untuk *well-behaved objective function*
- Kalau sudah mendekati *the best solution* maka tingkat pencapaian konvergensi cukup lambat

Langkah-langkah dalam estimasi dengan *Genetic Algorithm* adalah :

1. Menentukan masalah optimisasi
Masalah optimisasi tergantung pada kebutuhan kita. Contoh masalah optimisasi adalah memaksimalkan *likelihood function*.
2. Representasi
Dalam representasi, kita melakukan encoding terhadap *decision variable* menjadi *binary strings*.
3. *Initial Population*
Initial population dihasilkan secara random.
4. *Evaluation*
Langkah berikutnya setelah menghasilkan generasi adalah menghitung *fitness value* dari setiap anggota dalam populasi.
5. Membentuk populasi baru
Setelah melakukan evaluasi maka kita menciptakan populasi baru dari generasi terkini.
 - Reproduksi, di mana dua kromosom yang memiliki *fitness* terbaik akan dipertahankan dan diidutkan dalam produksi di generasi mendatang
 - *Crossover*, di mana akan diputuskan kromosom mana yang akan dipilih untuk *crossover*. Seleksi dilakukan dengan *cumulative probability*.
 - *Mutation*, dilaksanakan setelah *crossover*. *Mutation* mengubah satu atau lebih gen dengan probabilitas sama dengan *mutation rate*.
6. Hasil
Metode *Genetic Algorithm* dalam ekonomi dapat digunakan untuk kasus volatilitas dalam nilai tukar. Estimasi volatilitas nilai tukar dengan metode ARCH dan GARCH dapat menggunakan kerangka *Genetic Algorithm*. Dengan kerangka *Genetic Algorithm* kita juga dapat mencari tingkat ekuilibrium dari nilai tukar dengan mencari populasi nilai tukar yang memiliki nilai *fitness* yang paling tinggi. Ekuilibrium nilai tukar yang terpilih dihasilkan dari proses reproduksi, seleksi, dan *crossover* dengan mempertahankan individu yang terbaik di setiap generasi.

DAFTAR PUSTAKA

- Arifovic, Jasmina dan Gencay, Ramazan, 1998, "Statistical Properties of Genetic Learning in a Model of Exchange Rate", *Department of Economics*, University of Windsor.
- Enders, Walter, 1995, " *Applied Econometric Time Series*". John Wiley and Sons, Inc.
- Greene, William H, " *Econometric Analysis*. New Jersey : Prentice Hall 4th ed.
- Gujarati, Damodar, 2003, " *Basic Econometrics*. 4th " edition McGraw Hill.
- Im, So Kyung, Pesaran, M. Hashem, dan Shin, Yongcheol, 2002, "Testing for Unit Roots in Heterogenous Panels" *DAE Working Paper No. 9526*, University of Cambridge.
- Judge, et al, 1985," *The Theory and Practice of Econometrics* ", 2nd edition, John Willey and Sons.
- _____, 1988," *Introduction to The Theory and Practice of Econometrics* ". 2nd edition, John Willey and Sons.
- Kanas, Angelos dan Kouretas, Georgios P , 2000, "A Cointegration Approach to Lead-Lag Effect Among Size Sorted Equity Portfolios" , paper presented at seminar in Athens University Economics and Business.
- Maddala, G.S dan Kim, In Moo, 2000, " *Unit Roots, Cointegration, and Structural Change*", Cambridge University Press, Cambridge.
- Manual E-Views 3.1**
- Pesaran, M. Hashem dan Shin, Yongcheol, 1995, " An Autoregressive Distributed Lag Modelling Approach to Cointegration Analysis" dipresentasikan dalam *Symposium di Centennial of Ragnar Frisch, The Norwegian Academy of Science and Letters*. 3-5 Maret, chapter 11, Cambridge University Press,
- Pesaran, M. Hashem, Shin, Yongcheol dan Smith, Richard J, 1996, "Bounds Testing Approach to the Analysis of Long Run Relationship" dalam *DAE Working Paper, Vol 9622*, Department of Applied Economics, University of Cambridge.
- Proposal Inflasi Regional**, Februari 2003, Laboratorium Ilmu Ekonomi dan Studi Pembangunan.
- Pyndyck, Robert S, 1991," *Econometric Models and Economic Forecast* ",. 3rd ed. Mc. Graw Hill, Inc., New York.
- Zheng, Yingsong, 1999, "Genetic Algorithm Applications" ,Assignment 2 for Dr. Z Dong.