

## Regresi Data Panel (2) "Tahap Analisis"

**Author :** Muhammad Iqbal

**Date :** 20 January 2015

### Tahapan Analisis Regresi Data Panel

Berikut ini adalah tahapan analisis regresi data panel:

#### (1) Estimasi Model Regresi Data Panel

Model persamaan data panel yang merupakan gabungan dari data *cross section* dan data *time series* adalah sebagai berikut:

$$Y_{it} = \alpha + \beta_1 X_{1it} + \beta_2 X_{2it} + \dots + \beta_n X_{nit} + e_{it}$$

dimana:

$Y_{it}$  = variabel terikat (*dependent*)

$X_{it}$  = variabel bebas (*independent*)

$i$  = entitas ke- $i$

$t$  = periode ke- $t$

Persamaan di atas merupakan model regresi linier berganda dari beberapa variabel bebas dan satu variabel terikat. Estimasi model regresi linier berganda bertujuan untuk memprediksi parameter model regresi yaitu nilai konstanta ( $\alpha$ ) dan koefisien regresi ( $\beta_j$ ). Konstanta biasa disebut dengan intersep dan koefisien regresi biasa disebut dengan slope. Regresi data panel memiliki tujuan yang sama dengan regresi linier berganda, yaitu memprediksi nilai intersep dan slope. Penggunaan data panel dalam regresi akan menghasilkan intersep dan slope yang berbeda pada setiap entitas/ perusahaan dan setiap periode waktu. Model regresi data panel yang akan diestimasi membutuhkan asumsi terhadap intersep, slope dan variabel gangguannya. Menurut Widarjono (2007) ada beberapa kemungkinan yang akan muncul atas adanya asumsi terhadap intersep, slope dan variabel gangguannya.

- 1) Diasumsikan intersep dan slope adalah tetap sepanjang periode waktu dan seluruh entitas/perusahaan. Perbedaan intersep dan slope dijelaskan oleh variabel gangguan (residual).
- 2) Diasumsikan slope adalah tetap tetapi intersep berbeda antar entitas/perusahaan.
- 3) Diasumsikan slope tetap tetapi intersep berbeda baik antar waktu maupun antar individu.

- 4) Diasumsikan intersep dan slope berbeda antar individu.
- 5) Diasumsikan intersep dan slope berbeda antar waktu dan antar individu.

Dari berbagai kemungkinan yang disebutkan di atas muncullah berbagai kemungkinan model/teknik yang dapat dilakukan oleh regresi data panel. Dalam banyak literatur hanya asumsi pertama sampai ketiga saja yang sering menjadi acuan dalam pembentukan model regresi data panel.

Menurut Widarjono (2007, 251), untuk mengestimasi parameter model dengan data panel, terdapat tiga teknik (model) yang sering ditawarkan, yaitu:

### 1. Model *Common Effect*

Teknik ini merupakan teknik yang paling sederhana untuk mengestimasi parameter model data panel, yaitu dengan mengkombinasikan data *cross section* dan *time series* sebagai satu kesatuan tanpa melihat adanya perbedaan waktu dan entitas (individu). Dimana pendekatan yang sering dipakai adalah metode *Ordinary Least Square* (OLS). Model *Common Effect* mengabaikan adanya perbedaan dimensi individu maupun waktu atau dengan kata lain perilaku data antar individu sama dalam berbagai kurun waktu.

### 2. Model Efek Tetap (*Fixed Effect*)

Pendekatan model *Fixed Effect* mengasumsikan bahwa intersep dari setiap individu adalah berbeda sedangkan slope antar individu adalah tetap (sama). Teknik ini menggunakan variabel dummy untuk menangkap adanya perbedaan intersep antar individu.

### 3. Model Efek Random (*Random Effect*)

Pendekatan yang dipakai dalam *Random Effect* mengasumsikan setiap perusahaan mempunyai perbedaan intersep, yang mana intersep tersebut adalah variabel random atau stokastik. Model ini sangat berguna jika individu (entitas) yang diambil sebagai sampel adalah dipilih secara random dan merupakan wakil populasi. Teknik ini juga memperhitungkan bahwa error mungkin berkorelasi sepanjang *cross section* dan *time series*.

## **(2) Pemilihan Model (Teknik Estimasi) Regresi Data Panel**

Pada dasarnya ketiga teknik (model) estimasi data panel dapat dipilih sesuai dengan keadaan penelitian, dilihat dari jumlah individu bank dan variabel penelitiannya. Namun demikian, ada beberapa cara yang dapat digunakan untuk menentukan teknik mana yang paling tepat dalam mengestimasi parameter data panel. Menurut Widarjono (2007: 258), ada tiga uji untuk memilih teknik estimasi data panel. Pertama, uji statistik F digunakan untuk memilih antara metode *Common Effect* atau metode *Fixed Effect*. Kedua, uji Hausman yang digunakan untuk memilih antara metode *Fixed Effect* atau metode *Random Effect*. Ketiga, uji *Lagrange Multiplier* (LM)

digunakan untuk memilih antara metode *Common Effect* atau metode *Random Effect*.

Menurut, Nachrowi (2006, 318), pemilihan metode *Fixed Effect* atau metode *Random Effect* dapat dilakukan dengan pertimbangan tujuan analisis, atau ada pula kemungkinan data yang digunakan sebagai dasar pembuatan model, hanya dapat diolah oleh salah satu metode saja akibat berbagai persoalan teknis matematis yang melandasi perhitungan. Dalam *software Eviews*, metode *Random Effect* hanya dapat digunakan dalam kondisi jumlah individu bank lebih besar dibanding jumlah koefisien termasuk intersep. Selain itu, menurut beberapa ahli Ekonometri dikatakan bahwa, jika data panel yang dimiliki mempunyai jumlah waktu ( $t$ ) lebih besar dibandingkan jumlah individu ( $i$ ), maka disarankan menggunakan metode *Fixed Effect*. Sedangkan jika data panel yang dimiliki mempunyai jumlah waktu ( $t$ ) lebih kecil dibandingkan jumlah individu ( $i$ ), maka disarankan menggunakan metode *Random Effect*.

### a) Uji Statistik F (Uji Chow)

Untuk mengetahui model mana yang lebih baik dalam pengujian data panel, bisa dilakukan dengan penambahan variabel dummy sehingga dapat diketahui bahwa intersepanya berbeda dapat diuji dengan uji Statistik F. Uji ini digunakan untuk mengetahui apakah teknik regresi data panel dengan metode *Fixed Effect* lebih baik dari regresi model data panel tanpa variabel dummy atau metode *Common Effect*.

Hipotesis nul pada uji ini adalah bahwa intersep sama, atau dengan kata lain model yang tepat untuk regresi data panel adalah *Common Effect*, dan hipotesis alternatifnya adalah intersep tidak sama atau model yang tepat untuk regresi data panel adalah *Fixed Effect*.

Nilai Statistik F hitung akan mengikuti distribusi statistik F dengan derajat kebebasan (*deggre of freedom*) sebanyak  $m$  untuk numerator dan sebanyak  $n - k$  untuk denominator.  $m$  merupakan jumlah restriksi atau pembatasan di dalam model tanpa variabel dummy. Jumlah restriksi adalah jumlah individu dikurang satu.  $n$  merupakan jumlah observasi dan  $k$  merupakan jumlah parameter dalam model *Fixed Effect*. Jumlah observasi ( $n$ ) adalah jumlah individu dikali dengan jumlah periode, sedangkan jumlah parameter dalam model *Fixed Effect* ( $k$ ) adalah jumlah variabel ditambah jumlah individu. Apabila nilai F hitung lebih besar dari F kritis maka hipotesis nul ditolak yang artinya model yang tepat untuk regresi data panel adalah model *Fixed Effect*. Dan sebaliknya, apabila nilai F hitung lebih kecil dari F kritis maka hipotesis nul diterima yang artinya model yang tepat untuk regresi data panel adalah model *Common Effect*.

### b) Uji Hausman

Hausman telah mengembangkan suatu uji untuk memilih apakah metode *Fixed Effect* dan metode *Random Effect* lebih baik dari metode *Common Effect*. Uji Hausman ini didasarkan pada ide bahwa *Least Squares Dummy Variables* (LSDV) dalam metode metode *Fixed Effect* dan *Generalized Least Squares* (GLS) dalam metode *Random Effect* adalah efisien sedangkan *Ordinary Least Squares* (OLS) dalam metode *Common Effect* tidak efisien. Dilain pihak, alternatifnya adalah metode OLS efisien dan GLS tidak efisien. Karena itu, uji hipotesis nulnya adalah hasil estimasi keduanya tidak berbeda sehingga uji Hausman bisa dilakukan berdasarkan perbedaan estimasi tersebut.

Statistik uji Hausman mengikuti distribusi statistik *Chi-Squares* dengan derajat kebebasan (*df*) sebesar jumlah variabel bebas. Hipotesis nulnya adalah bahwa model yang tepat untuk regresi data panel adalah model *Random Effect* dan hipotesis alternatifnya adalah model yang tepat untuk regresi data panel adalah model *Fixed Effect*. Apabila nilai statistik Hausman lebih besar dari nilai kritis *Chi-Squares* maka hipotesis nul ditolak yang artinya model yang tepat untuk regresi data panel adalah model *Fixed Effect*. Dan sebaliknya, apabila nilai statistik Hausman lebih kecil dari nilai kritis *Chi-Squares* maka hipotesis nul diterima yang artinya model yang tepat untuk regresi data panel adalah model *Random Effect*.

### c) Uji *Lagrange Multiplier*

Menurut Widarjono (2007: 260), untuk mengetahui apakah model *Random Effect* lebih baik dari model *Common Effect* digunakan *Lagrange Multiplier* (LM). Uji Signifikansi *Random Effect* ini dikembangkan oleh Breusch-Pagan. Pengujian didasarkan pada nilai residual dari metode *Common Effect*.

Uji LM ini didasarkan pada distribusi *Chi-Squares* dengan derajat kebebasan (*df*) sebesar jumlah variabel independen. Hipotesis nulnya adalah bahwa model yang tepat untuk regresi data panel adalah *Common Effect*, dan hipotesis alternatifnya adalah model yang tepat untuk regresi data panel adalah *Random Effect*. Apabila nilai LM hitung lebih besar dari nilai kritis *Chi-Squares* maka hipotesis nul ditolak yang artinya model yang tepat untuk regresi data panel adalah model *Random Effect*. Dan sebaliknya, apabila nilai LM hitung lebih kecil dari nilai kritis *Chi-Squares* maka hipotesis nul diterima yang artinya model yang tepat untuk regresi data panel adalah model *Common Effect*.

### (3) Pengujian Asumsi Klasik (Multikolinieritas dan Heteroskedastisitas)

Regresi data panel memberikan alternatif model, *Common Effect*, *Fixed Effect* dan *Random Effect*. Model *Common Effect* dan *Fixed Effect* menggunakan pendekatan *Ordinary Least Squared* (OLS) dalam teknik estimasinya, sedangkan *Random Effect* menggunakan *Generalized Least Squares* (GLS) sebagai teknik estimasinya. Uji asumsi klasik yang digunakan dalam regresi linier dengan pendekatan *Ordinary Least Squared* (OLS) meliputi uji Linieritas, Autokorelasi, Heteroskedastisitas, Multikolinieritas dan Normalitas. Walaupun demikian, tidak semua uji asumsi klasik harus dilakukan pada setiap model regresi linier dengan pendekatan OLS.

Uji linieritas hampir tidak dilakukan pada setiap model regresi linier. Karena sudah diasumsikan bahwa model bersifat linier. Kalaupun harus dilakukan semata-mata untuk melihat sejauh mana tingkat linieritasnya.

Autokorelasi hanya terjadi pada data *time series*. Pengujian autokorelasi pada data yang tidak bersifat *time series* (*cross section* atau panel) akan sia-sia semata atau tidaklah berarti.

Multikolinieritas perlu dilakukan pada saat regresi linier menggunakan lebih dari satu variabel

bebas. Jika variabel bebas hanya satu, maka tidak mungkin terjadi multikolinieritas.

Heteroskedastisitas biasanya terjadi pada data *cross section*, dimana data panel lebih dekat ke ciri data *cross section* dibandingkan *time series*.

Uji normalitas pada dasarnya tidak merupakan syarat BLUE (*Best Linier Unbias Estimator*) dan beberapa pendapat tidak mengharuskan syarat ini sebagai sesuatu yang wajib dipenuhi.

Dari penjelasan di atas dapat ditarik kesimpulan bahwa pada regresi data panel, tidak semua uji asumsi klasik yang ada pada metode OLS dipakai, hanya multikolinieritas dan heteroskedastisitas saja yang diperlukan.

### **Uji Multikolinieritas**

Regresi data panel tidak sama dengan model regresi linier, oleh karena itu pada model data panel perlu memenuhi syarat terbebas dari pelanggaran asumsi-asumsi dasar (asumsi klasik). Meskipun demikian, adanya korelasi yang kuat antara variabel bebas dalam pembentukan sebuah model (persamaan) sangatlah tidak dianjurkan terjadi, karena hal itu akan berdampak kepada keakuratan pendugaan parameter, dalam hal ini koefisien regresi, dalam memperkirakan nilai yang sebenarnya. Korelasi yang kuat antara variabel bebas dinamakan *multikolinieritas*.

Menurut Chatterjee dan Price dalam Nachrowi (2002), adanya korelasi antara variabel-variabel bebas menjadikan interpretasi koefisien-koefisien regresi mejadi tidak benar lagi. Meskipun demikian, bukan berarti korelasi yang terjadi antara variabel-variabel bebas tidak diperbolehkan, hanya kolinieritas yang sempurna (*perfect collinierity*) saja yang tidak diperbolehkan, yaitu terjadinya korelasi linier antara sesama variabel bebasnya. Sedangkan untuk sifat kolinier yang hampir sempurna (hubungannya tidak bersifat linier atau korelasi mendekati nol) masih diperbolehkan atau tidak termasuk dalam pelanggaran asumsi.

Ada beberapa cara untuk mengidentifikasi adanya *multikolinieritas*, dan cara yang paling mudah adalah dengan mencari nilai koefisien korelasi antar variabel bebas. Koefisien korelasi antara dua variabel yang bersifat kuantitatif dapat menggunakan *coefficient correlation pearson*, dengan rumus sebagai berikut:

Dimana  $X_i$  dan  $Y_i$  adalah variabel bebas yang akan dicari nilai koefisien korelasinya dan  $n$  adalah jumlah data dari kedua variabel bebas tersebut. Nilai mutlak dari koefisien korelasi besarnya dari nol sampai satu. Semakin mendekati satu, maka dapat dikatakan semakin kuat hubungan antara kedua variabel tersebut dan artinya semakin besar kemungkinan terjadinya *multikolinieritas*.

### **Uji Heteroskedastisitas**

Regresi data panel tidak sama dengan model regresi linier, oleh karena itu pada model data panel perlu memenuhi syarat BLUE (*Best Linear Unbiased Estimator*) atau terbebas dari pelanggaran asumsi-asumsi dasar (asumsi klasik). Jika dilihat dari ketiga pendekatan yang

dipakai, maka hanya uji heteroskedastisitas saja yang relevan dipakai pada model data panel.

Uji heteroskedastisitas digunakan untuk melihat apakah residual dari model yang terbentuk memiliki varians yang konstan atau tidak. Suatu model yang baik adalah model yang memiliki varians dari setiap gangguan atau residualnya konstan. Heteroskedastisitas adalah keadaan dimana asumsi tersebut tidak tercapai, dengan kata lain dimana adalah ekspektasi dari eror dan adalah varians dari eror yang berbeda tiap periode waktu.

Dampak adanya heteroskedastisitas adalah tidak efisiennya proses estimasi, sementara hasil estimasinya tetap konsisten dan tidak bias. Eksistensi dari masalah heteroskedastisitas akan menyebabkan hasil Uji-*t* dan Uji-F menjadi tidak berguna (*miss leading*).

Ada beberapa metode yang dapat digunakan untuk mendeteksi heteroskedastisitas, tetapi dalam penelitian ini hanya akan dilakukan dengan menggunakan *White Heteroskedasticity Test* pada *consistent standard error & covariance*. Hasil yang diperlukan dari hasil uji ini adalah nilai F dan *Obs\*R-squared*, dengan hipotesis sebagai berikut:

$H_0$  : *Homoskedasticity*

$H_1$  : *Heteroskedasticity*

Kemudian kita bandingkan antara nilai *Obs\*R-squares* dengan nilai tabel dengan tingkat kepercayaan tertentu dan derajat kebebasan yang sesuai dengan jumlah variabel bebas. Jika nilai Uji Heteroskedastisitas tabel maka  $H_0$  diterima, dengan kata lain tidak ada masalah heteroskedastisitas.

#### **(4) Uji Kelayakan (*Goodness of Fit*) Model Regresi Data Panel**

##### **Uji Hipotesis**

Menurut Nachrowi (2006), uji hipotesis berguna untuk menguji signifikansi koefisien regresi yang didapat. Artinya, koefisien regresi yang didapat secara statistik tidak sama dengan nol, karena jika sama dengan nol maka dapat dikatakan bahwa tidak cukup bukti untuk menyatakan variabel bebas mempunyai pengaruh terhadap variabel terikatnya. Untuk kepentingan tersebut, maka semua koefisien regresi harus diuji. Ada dua jenis uji hipotesis terhadap koefisien regresi yang dapat dilakukan, yaitu:

##### 1. Uji-F

Uji-F diperuntukkan guna melakukan uji hipotesis koefisien (*slope*) regresi secara bersamaan, dengan kata lain digunakan untuk memastikan bahwa model yang dipilih layak atau tidak untuk menginterpretasikan pengaruh variabel bebas terhadap variabel terikat.

##### 2. Uji-*t*

Jika Uji-F dipergunakan untuk menguji koefisien regresi secara bersamaan, maka Uji-*t* digunakan untuk menguji koefisien regresi secara individu. Pengujian dilakukan terhadap koefisien regresi populasi, apakah sama dengan nol, yang berarti variabel bebas tidak mempunyai pengaruh signifikan terhadap variabel terikat, atau tidak sama dengan nol, yang berarti variabel bebas mempunyai pengaruh signifikan terhadap variabel terikat.

### **Koefisien Determinasi**

Koefisien Determinasi (*Goodness of Fit*) dinotasikan dengan *R-squares* yang merupakan suatu ukuran yang penting dalam regresi, karena dapat menginformasikan baik atau tidaknya model regresi yang terestimasi. Nilai Koefisien Determinasi mencerminkan seberapa besar variasi dari variabel terikat dapat diterangkan oleh variabel bebasnya. Bila nilai Koefisien Determinasi sama dengan 0, artinya variasi dari variabel terikat tidak dapat diterangkan oleh variabel-variabel bebasnya sama sekali. Sementara bila nilai Koefisien Determinasi sama dengan 1, artinya variasi variabel terikat secara keseluruhan dapat diterangkan oleh variabel-variabel bebasnya. Dengan demikian baik atau buruknya suatu persamaan regresi ditentukan oleh *R-squares*-nya yang mempunyai nilai antara nol dan satu.

#### Referensi:

Baltagi, Bagi (2005). *Econometric Analysis of Panel Data*, Third Edition. John Wiley & Sons.

Nachrowi, D.N. dan H. Usman (2002). *Penggunaan Teknik Ekonometrika*. Jakarta: PT Raja Grafindo Persada.

Widarjono, Agus (2007). *Ekonometrika: Teori dan Aplikasi Untuk Ekonomi dan Bisnis*, edisi kedua. Yogyakarta: Ekonisia FE Universitas Islam Indonesia.