

## **BAB III**

### **METODOLOGI PENELITIAN**

#### **A. Tujuan penelitian**

Berdasarkan masalah-masalah yang telah peneliti rumuskan, maka tujuan penelitian ini adalah untuk mendapatkan pengetahuan yang tepat dan dapat dipercaya tentang Pengaruh PDRB dan Investasi terhadap Pendapatan Asli Daerah Kabupaten/Kota di Jawa Barat.

#### **B. Objek Penelitian**

Penelitian ini dilaksanakan dengan mengambil data PDRB, Investasi, dan penerimaan Pendapatan Asli Daerah Kabupaten/Kota di Jawa Barat. Data tersebut diperoleh dari *website* World Bank, Badan Pusat Statistik (BPS) Jawa Barat, dan Dinas Pendapatan Daerah (Dispenda) Jawa Barat.

Penelitian ini dilaksanakan pada bulan Juli 2012. Waktu ini dipilih karena merupakan waktu yang paling luang untuk melakukan penelitian. Pada bulan ini sudah tidak ada lagi jadwal perkuliahan di dalam kelas sehingga peneliti dapat lebih fokus pada saat penelitian.

#### **C. Metode Penelitian**

Metode penelitian yang digunakan dalam penelitian ini adalah berbentuk time series dari tahun 2001 sampai tahun 2010 dan cross section, yaitu 11 Kabupaten dan 4 Kota, dengan menghilangkan 6 kabupaten yaitu

kabupaten Ciamis, Kabupaten Garut, kabupaten Indramayu, kabupaten Kuningan, kabupaten tasikmalaya, kabupaten Bandung Barat dan 5 kota yaitu kota Sukabumi, kota Tasikmalaya, kota Cimahi, kota Banjar, dan kota Bogor di Jawa Barat atau dikenal dengan analisis panel. Alasan menghilangkan 11 kabupaten dan kota tersebut dikarenakan tidak tersedianya data investasi daerah tersebut. Sehingga jumlah data secara keseluruhan dengan menggunakan cross section dan time series dalam bentuk panel menjadi 150 data.

Sedangkan pendekatan yang digunakan adalah pendekatan korelasional yang bertujuan untuk mengetahui seberapa besar pengaruh antara variabel-variabel yang diteliti yaitu Pendapatan Asli Daerah sebagai variabel terikat, Produk Domestik Regional Bruto sebagai variabel bebas pertama dan Investasi sebagai variabel bebas kedua.

#### **D. Jenis dan Sumber Data**

Data yang digunakan dalam penelitian ini adalah data sekunder yang berbentuk time series dari tahun 2001 sampai tahun 2010. Data sekunder adalah jenis data yang diperoleh dan digali melalui hasil pengolahan pihak kedua dari hasil penelitian lapangannya, baik berupa data kualitatif maupun kuantitatif.<sup>37</sup> Data yang digunakan meliputi data PDRB, Investasi, dan penerimaan Pendapatan Asli Daerah Kabupaten/kota Jawa Barat yang diperoleh dari *website* World Bank, Badan Pusat Statistik (BPS) dan Dinas Pendapatan Daerah (Dispenda) Jawa Barat.

---

<sup>37</sup> Muhammad Teguh, *Metodologi Penelitian Ekonomi* (Jakarta: PT Raja Grafindo Persada, 2005), p. 121

## **E. Teknik Operasional Variabel**

### **1. Pendapatan Asli Daerah (PAD)**

#### a. Definisi Konseptual

Pendapatan Asli Daerah (PAD) adalah penerimaan daerah yang diperoleh dari dalam daerahnya sendiri, yang dipungut berdasarkan ketentuan yang berlaku. PAD terdiri dari hasil pajak daerah, hasil retribusi daerah, hasil perusahaan milik daerah dan hasil pengelolaan kekayaan daerah lainnya yang dipisahkan, dan lain-lain Pendapatan Asli Daerah yang sah.

#### b. Definisi Operasional

Pendapatan Asli Daerah adalah perolehan daerah yang dipungut berdasarkan peraturan daerah sesuai dengan perundang-undangan untuk mengumpulkan dana guna membiayai kegiatan dan keperluan daerah. Data Pendapatan Asli Daerah yang dipergunakan dalam penelitian ini adalah data tahun 2001-2010 yang diambil dari BPS.

### **2. Produk Domestik Regional Bruto (PDRB)**

#### a. Definisi Konseptual

Produk Domestik Regional Bruto (PDRB) adalah nilai produk barang dan jasa akhir yang dihasilkan oleh berbagai unit produksi di dalam suatu daerah atau wilayah dalam jangka waktu tertentu (satu tahun), yang diproduksi bukan hanya oleh perusahaan yang ada di wilayah tersebut, tetapi juga warga asing yang tinggal di wilayah tersebut.

b. Definisi Operasional

PDRB adalah jumlah nilai tambah yang dihasilkan oleh seluruh unit usaha dalam suatu daerah tertentu, atau merupakan jumlah nilai barang dan jasa akhir yang dihasilkan oleh seluruh unit ekonomi dalam suatu waktu tertentu. Data PDRB yang dipergunakan dalam penelitian ini adalah data tahun 2001-2010 yang diambil dari BPS.

**3. Investasi**

a. Definisi Konseptual

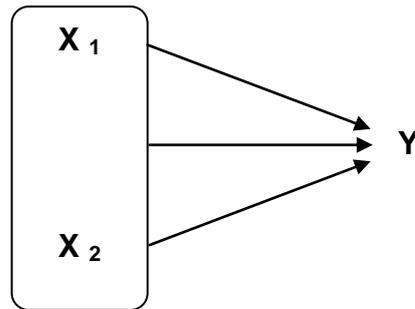
Berinvestasi berarti menempatkan sejumlah dana atau barang dalam jangka waktu tertentu, untuk memperoleh manfaat ekonomi, sosial, dan atau manfaat lainnya di masa yang akan datang. Investasi dapat dilakukan oleh pihak pemerintah maupun pihak swasta (kelompok atau perorangan).

b. Definisi Operasional

Investasi adalah menanamkan atau menempatkan asset baik berupa harta maupun dana, pada sesuatu dengan harapan akan memberikan hasil atau akan meningkatkan nilainya di masa akan datang. Data yang digunakan dalam penelitian ini adalah data tahun 2001 – 2010 yang diambil dari BPS.

**F. Konstelasi Pengaruh Antar Variabel**

Konstelasi pengaruh antar variabel dalam penelitian ini bertujuan untuk memberikan arahan atau gambaran dari penelitian ini, yang dapat digambarkan sebagai berikut:



Keterangan:

X1 = Produk Domestik Regional Bruto (PDRB)

X2 = Investasi

Y = Pendapatan Asli Daerah

→ = Arah Pengaruh

## G. Teknik Analisis Data

### 1. Panel Data

Data panel adalah jenis data yang merupakan gabungan antara data runtun waktu (*time series*) dan lintas sektor (*cross section*). Oleh karenanya, data panel memiliki gabungan karakteristik kedua jenis data tadi, yaitu terdiri atas beberapa objek dan meliputi beberapa periode.<sup>38</sup> Estimasi panel data akan meningkatkan derajat kebebasan, mengurangi kolinearitas antara variabel penjelas dan memperbaiki efisiensi estimasi. Verbeek (2000:310) mengemukakan bahwa keuntungan regresi dengan data panel adalah kemampuan regresi data panel dalam mengidentifikasi parameter-parameter regresi secara pasti tanpa asumsi restriksi atau kendala.

---

<sup>38</sup> Wing Wahyu Winarno, *Analisis Ekonometrika dan Statistika dengan Eviews Edisi Ke-3* (Yogyakarta; STIM YKPN, 2011), p.1.2

Menurut Baltagi (2001), keunggulan penggunaan data panel dibanding data runtun waktu dan data lintas sektor adalah :

- 1) Estimasi data panel dapat menunjukkan adanya heterogenitas dalam tiap unit.
- 2) Dengan data panel, data lebih informatif, mengurangi kolinieritas antara variabel, meningkatkan derajat kebebasan dan lebih efisien.
- 3) Data panel cocok digunakan untuk menggambarkan adanya dinamika perubahan.
- 4) Data panel dapat lebih mampu mendeteksi dan mengukur dampak.
- 5) Data panel bisa digunakan untuk studi dengan model yang lebih lengkap.
- 6) Data panel dapat meminimumkan bias yang mungkin dihasilkan dalam regresi.

Data panel merupakan gabungan dari data *cross section* dan data *time series*, dalam hal modelnya dapat dituliskan sebagai berikut:

$$Y_{it} = \alpha + \beta X_{it} + \varepsilon_{it} \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T \quad (3.1)$$

Keterangan :

N = banyaknya observasi

T = banyaknya waktu

N x T = banyaknya data panel

Untuk mengestimasi parameter model dengan data panel, terdapat beberapa teknik yang ditawarkan, yaitu:

**a. Model *Common Effect***

Model *common effects* atau *pooled regression* merupakan model regresi data panel yang paling sederhana. Model ini pada dasarnya mengabaikan struktur panel dari data, sehingga diasumsikan bahwa perilaku antar individu sama dalam berbagai kurun waktu atau dengan kata lain pengaruh spesifik dari masing-masing individu diabaikan atau dianggap tidak ada. Dengan demikian, akan dihasilkan sebuah persamaan regresi yang sama untuk setiap unit *cross-section*. Sesuatu yang secara realistis tentunya kurang dapat diterima. Karena itu, model ini sangat

jarang digunakan dalam analisis data panel. Persamaan regresi untuk model *common effect* dapat dituliskan sebagai berikut:

$$Y_{it} = \alpha + \beta X_{it} + \varepsilon_{it} \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T \quad (3.2)$$

Keterangan :

Y = variabel dependen

A = koefisien regresi

X = variabel independen

B = estimasi parameter

Eit= *aerror term*

N = jumlah (individu)

T = jumlah periode waktu.

Ekananda (2005) menyatakan bahwa berdasarkan asumsi struktur matriks varians-covarians residual, maka pada model *common effects*, terdapat 4 metode estimasi yang dapat digunakan, yaitu:

- (1) *Ordinary Least Square (OLS)*, jika struktur matriks varians-kovarians residualnya diasumsikan bersifat homoskedastik dan tidak ada *cross sectional correlation*,
- (2) *Generalized Least Square (GLS) / Weighted Least Square (WLS): Cross Sectional Weight*, jika struktur matriks varians-kovarians residualnya diasumsikan bersifat heteroskedastik dan tidak ada *cross sectional correlation*,
- (3) *Feasible Generalized Least Square (FGLS)/Seemingly Uncorrelated Regression (SUR)* atau *Maximum Likelihood Estimator (MLE)*, jika struktur matriks varians-kovarians residualnya diasumsikan bersifat heteroskedastik dan ada *cross sectional correlation*,
- (4) *Feasible Generalized Least Square (FGLS)* dengan proses *autoregressive (AR)* pada *error term*-nya, jika struktur matriks varians-kovarians residualnya diasumsikan bersifat heteroskedastik dan ada korelasi antar waktu pada residualnya.

### b. Model *Fixed Effect*

Jika model *common effect* cenderung mengabaikan struktur panel dari data dan pengaruh spesifik masing-masing individu, maka model *fixed effect* adalah sebaliknya. Pada model ini, terdapat efek spesifik individu  $\alpha_i$  dan diasumsikan berkorelasi dengan variabel penjelas yang teramati  $X_{it}$ .

Ekananda (2005) menyatakan bahwa berdasarkan asumsi struktur matriks varians-kovarians residual, maka pada model *fixed effect*, terdapat 3 metode estimasi yang dapat digunakan, yaitu:

- (1) *Ordinary Least Square (OLS/LSDV)*, jika struktur matriks varians-kovarians residualnya diasumsikan bersifat homoskedastik dan tidak ada *cross sectional correlation*,
- (2) *Weighted Least Square (WLS)*, jika struktur matriks varians-kovarians residualnya diasumsikan bersifat heteroskedastik dan tidak ada *cross sectional correlation*,
- (3) *Seemingly Uncorrelated Regression (SUR)*, jika struktur matriks varians-kovarians residualnya diasumsikan bersifat heteroskedastik dan ada *cross sectional correlation*.

Model data panel dengan OLS, ada asumsi yang menyatakan bahwa dalam persamaan 3.2,  $\alpha$  dan  $\beta$  konstan untuk setiap individu ( $i$ ) dan waktu ( $t$ ) kurang realistis.

$$Y_{it} = \alpha + \beta X_{it} + \epsilon_{it}; \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T$$

Maka dari itu, dalam model efek tetap hal tersebut diatasi yang mana model ini memungkinkan adanya perubahan  $\alpha$  pada setiap  $i$  dan  $t$ . Secara matematis, model efek tetap dapat dituliskan sebagai berikut:

$$Y_{it} = \alpha + \beta X_{it} + \gamma_2 W_{2i} + \gamma_3 W_{3i} + \dots + \gamma_N W_{Ni} + \sigma_2 Z_{i2} + \sigma_3 Z_{i3} + \dots + \sigma_i Z_{iT} + \epsilon_{it} \quad (3.3)$$

Keterangan :

$Y_{it}$  = variabel terikat untuk kota ke- $i$  dan tahun ke- $t$

$X_{it}$  = variabel bebas untuk kota ke- $i$  dan tahun ke- $t$

$W_{it}$  yang didefinisikan sebagai berikut:

$W_{it} = 1$ ; untuk negara  $i$ ;  $i = 1, 2, \dots, N = 0$  ; lainnya

Berdasarkan model di atas, terlihat bahwa sesungguhnya model efek tetap adalah sama dengan regresi yang menggunakan *dummy variable* sebagai variabel bebas, sehingga dapat diestimasi dengan OLS. Dengan diestimasi model tersebut dengan OLS, maka akan diperoleh estimator yang tidak bias dan konsisten.

### c. Model *Random Effect*

Keputusan untuk memasukkan peubah *dummy* dalam model *fixed effects* akan menimbulkan konsekuensi tersendiri yaitu dapat mengurangi banyaknya derajat kebebasan yang pada akhirnya akan mengurangi efisiensi dari parameter yang diestimasi. Untuk mengatasi masalah tersebut maka dapat digunakan model *random effects*.

Dalam model ini, parameter yang berbeda antar individu maupun antar waktu dimasukkan ke dalam *error*, karena hal inilah model ini sering juga disebut sebagai *error component model*. Bentuk model *random effects* dapat dijelaskan dengan persamaan berikut:

$$Y_{it} = \alpha + \beta X_{it} + \epsilon_{it} \quad \epsilon_{it} = u_i + v_t + w_{it} \quad (3.4)$$

Keterangan :

$u_i \sim N(0, \delta U^2) = \text{error component cross section}$

$v_t \sim N(0, \delta V^2) = \text{error component time series}$

$w_{it} \sim N(0, \delta W^2) = \text{error component combinations}$

Melihat persamaan di atas, maka dapat dinyatakan bahwa Model Efek Random menganggap efek rata-rata dari data *cross-section* dan *time-series* direpresentasikan dalam *intercept*. Sedangkan deviasi efek secara random untuk data *time-series* direpresentasikan dalam  $v_t$  dan deviasi untuk data *cross-section* dinyatakan dalam  $u_i$ .

Telah diketahui bahwa:  $\epsilon_{it} = u_i + v_t + w_{it}$ . Dengan demikian varians dari *error* tersebut dapat dituliskan dengan:

$$\text{Var}(\epsilon_{it}) = \delta U^2 + \delta V^2 + \delta W^2 \quad (3.5)$$

Hal ini tentunya berbeda dengan model OLS yang diterapkan pada panel data, sebagaimana telah dijelaskan di atas, yang mempunyai varian *error* sebesar:

$$\text{Var}(\epsilon_{it}) = \delta W^2 \quad (3.6)$$

Dengan demikian, Model Efek Random bisa diestimasi dengan OLS bila  $\delta U^2 = \delta V^2 = 0$ . Kalau tidak demikian, Model Efek Random diestimasi dengan metode *Generalized Least Square* (GLS). Asumsi yang digunakan dalam Model Efek Random ini adalah *error* secara individual tidak saling berkorelasi, begitu pula dengan *error* kombinasinya. Penggunaan pendekatan *random effects* dapat menghemat derajat kebebasan dan tidak mengurangi jumlahnya seperti pada pendekatan *fixed effects*. Hal ini berimplikasi pada parameter hasil estimasi akan menjadi efisien. Semakin efisien maka model akan semakin baik. Terkait dengan

beberapa pilihan teknik untuk permodelan panel data, sebelum model diestimasi dengan model yang tepat, terlebih dahulu dilakukan uji spesifikasi apakah *Common Effect*, *Fixed Effect* dan atau *Random Effect* memberikan hasil yang sama. Penyeleksian model estimasi data panel antara lain:

**(1) Pengujian Signifikansi *Common Effects* atau *Fixed Effects***

Signifikansi model *fixed effect* dapat dilakukan dengan uji statistik F. Uji F digunakan untuk mengetahui apakah teknik regresi data panel dengan *fixed effect* lebih baik dari model regresi data panel tanpa variabel *dummy* (*common effect*) dengan melihat *residual sum squares* (*RSS*).

$$F = \frac{(RSS_1 - RSS_2) / (n-1)}{RSS_2 / (nT - n - k)} \quad (3.7)$$

Dimana n adalah jumlah individu, T adalah periode waktu, k adalah parameter dalam model *fixed effect*, RSS1 dan RSS2 masing-masing merupakan *residual sum of squares* teknik tanpa variabel *dummy* dan teknik *fixed effect* dengan variabel *dummy*. Nilai statistik F hitung akan mengikuti distribusi statistik F dengan derajat bebas sebanyak (n-1) untuk numerator dan (nT-n-k) untuk *denominator*. Jika nilai statistik F hitung lebih besar dari F tabel, maka hipotesis nul akan ditolak, yang berarti koefisien intersep dan slope adalah sama tidak berlaku, sehingga teknik regresi data panel dengan *fixed effect* lebih baik dari *common effect*.

## (2) Pengujian Signifikansi Common Effects atau Random Effects

Breush dan Pagan (1980) telah mengembangkan pengujian *Lagrange Multiplier* untuk mengetahui signifikansi dari *random effects* berdasarkan residual dari OLS (*common effects*). Secara matematis, statistik uji untuk *LM test* (*Lagrange Multiplier*) dapat dituliskan sebagai berikut:

$$LM = \frac{nT}{2(T-1)} \left[ \frac{\sum_{i=1}^n (\sum_{t=1}^T e_{it})^2}{\sum_{i=1}^n \sum_{t=1}^T e_{it}^2} - 1 \right]^2 = \frac{nT}{2(T-1)} \left[ \frac{\sum_{i=1}^n [T\bar{e}_i]^2}{\sum_{i=1}^n \sum_{t=1}^T e_{it}^2} - 1 \right]^2 \quad (3.8)$$

Di bawah hipotesis nul, LM mengikuti sebaran *chi-square* dengan derajat bebas satu. Jika hasil LM statistik lebih besar dari nilai kritis statistik *chi-square*, maka hipotesis nul akan ditolak, yang berarti estimasi yang tepat untuk regresi data panel adalah metode *random effect* dibandingkan metode *common effects*.

## (3) Pengujian Signifikansi Fixed Effects atau Random Effects

Setelah menguji signifikansi antara *common effects* atau *fixed effects* serta *common effects* atau *random effects*, maka selanjutnya jika terbukti *fixed effects* dan *random effects* sama-sama lebih baik dari *common effects* adalah melakukan pengujian signifikansi *fixed effects* atau *random effects*.

Uji ini dilakukan dengan membandingkan untuk *subset* dari koefisien variabel-variabel yang bervariasi antar unit waktu (*time-varying variables*). Secara matematis dengan menggunakan notasi matriks, statistik uji Hausman (*H*) dapat dituliskan sebagai berikut:

$$H = \left( \hat{\beta}_{FE} - \hat{\beta}_{RE} \right)' \left[ \text{var} \left( \hat{\beta}_{FE} \right) - \text{var} \left( \hat{\beta}_{RE} \right) \right]^{-1} \left( \hat{\beta}_{FE} - \hat{\beta}_{RE} \right) \quad (3.9)$$

Di bawah hipotesis nul, statistik uji ini mengikuti sebaran *chi-square* dengan derajat bebas  $M$ , di mana  $M$  adalah jumlah variabel penjelas yang nilainya bervariasi antar unit waktu di dalam model.

Hipotesis nul pada uji Hausman adalah efek spesifik individu tidak berkorelasi dengan peregresi atau dengan kata lain model *random effect* lebih baik bila dibandingkan dengan model *fixed effect*. Di bawah hipotesis nul, pendugaan parameter dengan menggunakan *random effect* adalah konsisten dan efisien, sedangkan pendugaan dengan *fixed effect* meskipun tetap konsisten, tetapi tidak lagi efisien. Di bawah hipotesis alternatif, estimasi dengan *random effect* menjadi tidak konsisten, sebaliknya estimasi dengan *fixed effect* tetap konsisten. Jika nilai statistik Hausman lebih besar daripada nilai kritis statistik *chi-square*, maka hipotesis nul akan ditolak, yang berarti estimasi yang tepat untuk regresi data panel adalah metode *fixed effect* dari pada metode *random effect*.

Judge *et al.* dalam Gujarati (2003) memberikan sejumlah pertimbangan terkait pilihan apakah menggunakan model *fixed effect* (FE) ataukah model *random effect* (RE). Pertimbangan-pertimbangan itu adalah sebagai berikut:

- Jika jumlah data *time series* (T) besar dan jumlah data *cross-section* (N) kecil, ada kemungkinan perbedaan nilai parameter yang diestimasi dengan FE dan RE cukup kecil. Karena itu, pilihan ditentukan berdasarkan kemudahan perhitungan. Dalam hal ini, adalah model FE.
- Ketika N besar dan T kecil estimasi kedua metode dapat berbeda secara signifikan. Pada kondisi seperti ini, pilihan ditentukan berdasarkan keyakinan apakah individu yang diobservasi merupakan sampel acak yang diambil dari populasi tertentu atau tidak. Jika observasi bukan merupakan sampel acak, maka digunakan model FE. Jika sebaliknya, maka digunakan model RE.
- Jika efek individu tidak teramati  $\alpha_i$  berkorelasi dengan satu atau lebih variabel bebas, maka estimasi dengan RE bias, sedangkan estimasi dengan FE tidak bias.
- Jika N besar dan T kecil, serta semua asumsi yang disyaratkan oleh model RE terpenuhi, maka estimasi dengan menggunakan RE lebih efisien dibanding estimasi dengan FE.

Dalam penelitian ini, penentuan apakah model FE atau RE yang akan digunakan selain didasarkan pada sejumlah pertimbangan yang telah disebutkan, juga akan didasarkan pada kriteria ekonomi (*make sense* secara ekonomi). Dalam hal ini, adalah kesesuaian tanda hasil estimasi koefisien regresi setiap variabel di dalam model dengan teori dan kewajaran besaran nilai koefisien hasil estimasi tersebut.

Setelah menentukan spesifikasi model yang akan digunakan, tahapan selanjutnya adalah memilih metode estimasi (estimator) yang tepat sesuai dengan struktur varian kovarian residual. Konsekuensi yang muncul ketika membangun model regresi dengan data panel adalah bertambahnya komponen residual, karena adanya dimensi *cross-section* dan *time-series* pada data. Kondisi ini menyebabkan matriks varian kovarian residual menjadi sedikit lebih kompleks bila dibandingkan dengan model regresi klasik yang hanya menggunakan data *cross-section* atau data *time-series*.

Pada model regresi klasik, pelanggaran terhadap asumsi klasik terkait residual, seperti heterokedastisitas dan autokorelasi merupakan masalah serius yang mengakibatkan penduga parameter regresi yang diestimasi dengan OLS tidak lagi bersifat BLUE (*best linear unbiased estimator*). Tindakan yang biasa dilakukan untuk mengatasi masalah ini adalah dengan melakukan penghitungan *robust standard error*. Dalam pemodelan regresi dengan data panel, terjadinya pelanggaran asumsi regresi linier klasik pada residual adalah hal yang sangat sulit dihindari, dan tidak seperti pada regresi klasik, pelanggaran dapat diakomodasi

untuk menentukan metode estimasi terbaik bagi spesifikasi model yang digunakan.

Terdapat beberapa kemungkinan struktur varian kovarian residual yang mungkin terjadi pada model regresi data panel. Berbagai kemungkinan yang dibahas pada bagian ini adalah yang biasa dijumpai pada estimasi model dengan *common effects* dan *fixed effect*. Karena itu, metode-metode estimasi yang dapat digunakan terkait struktur varian kovarian residual yang dipaparkan pada bagian ini hanya akan diterapkan pada model yang diestimasi dengan *common effects* atau *fixed effect*.

- **Pemilihan Estimator Struktur Homoskedastik atau Heteroskedastik dengan Uji Lagrange Multiplier (LM)**

Pada pengujian ini, hipotesis nul ( $H_0$ ) yang digunakan adalah bahwa struktur varians-covarians residual bersifat homoskedastik. Sementara hipotesis alternatif ( $H_1$ ) adalah bahwa struktur varians-covarians residual bersifat heteroskedastik.

Secara matematis, statistik uji yang digunakan dapat dirumuskan sebagai berikut :

$$LM = \frac{T}{2} \sum_{i=1}^n \left[ \frac{\hat{\sigma}_i^2}{\hat{\sigma}^2} - 1 \right]^2 \quad (3.10)$$

di mana  $T$  adalah jumlah periode waktu,  $n$  adalah jumlah individu, adalah varians residual persamaan ke- $i$  pada kondisi homoskedastik, dan adalah *Sum Square Residual (SSR)* persamaan *system* pada kondisi homoskedastik.

Statistik uji LM ini mengikuti distribusi statistik *chi-square* dengan derajat bebas sebanyak  $n-1$ . Jika nilai statistik LM lebih besar dari nilai kritis statistik *chi-*

*square*, maka hipotesis nul ditolak, yang berarti struktur varians-covarians residual bersifat heteroskedastik.

- **Pemilihan Estimator Struktur Heteroskedastik dan Ada Cross – sectional Correlation**

Pengujian ini dilakukan apabila hasil pengujian LM menunjukkan bahwa struktur varians-covarians residual bersifat heteroskedastik. Pada pengujian ini, hipotesis nul (H0) yang digunakan adalah bahwa struktur varians-covarians residual bersifat heteroskedastik dan tidak ada *cross sectional correlation*. Sementara hipotesis alternatifnya (H1) adalah bahwa struktur varians-covarians residual bersifat heteroskedastik dan ada *cross sectional correlation* (*Seemingly Uncorrelated Regression/SUR*). Secara matematis, statistik uji yang digunakan dapat dirumuskan sebagai berikut:

$$\lambda LM = T \sum_{i=2}^n \sum_{j=1}^{i-1} r_{ij}^2 \quad (3.11)$$

di mana  $T$  adalah jumlah periode waktu,  $n$  adalah jumlah individu, dan  $r_{ij}$  adalah *residual correlation coefficient* antara persamaan ke- $i$  dan ke- $j$ .

Statistik uji LM ini mengikuti distribusi statistik *chi-square* dengan derajat bebas sebanyak  $n(n-1)/2$ . Jika nilai statistik LM lebih besar dari nilai kritis statistik *chi-square*, maka hipotesis nul akan ditolak, yang berarti struktur varians-covarians residual bersifat heteroskedastik dan ada *cross sectional correlation* (*Seemingly Uncorrelated Regression/SUR*).

Pengujian statistik juga dilakukan untuk mengetahui apakah model regresi non linier merupakan model yang tepat untuk menggambarkan hubungan antar variabel dan apakah ada hubungan yang signifikan diantara variabel-variabel

dependen dengan variabel-variabel penjelas (seperti : uji statistik t dan uji statistik F) selain itu kita bisa melihat nilai hasil estimasi untuk  $R^2$  (koefisien determinasi).

## 2. Uji Hipotesis

### a. Uji Statistik F

Uji statistik F merupakan pengujian koefisien regresi secara keseluruhan. Pengujian ini menunjukkan apakah semua variabel bebas yang dimasukkan kedalam model mempunyai pengaruh secara bersama-sama terhadap variabel terikat.

Hipotesis pengujian yang digunakan adalah:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$$

$H_1$  : paling sedikit salah satu nilai  $\beta_i \neq 0$ , dengan  $i = 1, 2, \dots, k$ .

Statistik uji  $F$  dihitung dengan formula sebagai berikut:

$$F_{(n-k-1, nT-n-k)} = \frac{R^2 / (n + k - 1)}{(1 - R^2) / (nT - n - k)} \quad (3.12)$$

Hipotesis nol ditolak jika yang berarti bahwa minimal ada satu variabel bebas yang signifikan berpengaruh terhadap variabel tidak bebasnya. Keputusan ini dapat juga didasarkan pada perbandingan nilai  $p$ -value dengan tingkat signifikansinya. Hipotesis nol ditolak jika nilai  $p$ -value lebih kecil dari tingkat signifikansi ( $\alpha$ ).

### b. Koefisien determinasi ( $R^2$ )

$R^2$  digunakan untuk mengukur kebaikan atau kesesuaian suatu model persamaan regresi. Besaran  $R^2$  dihitung dengan rumus:

$$R^2 = \frac{\sum (\hat{Y}_i - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2} = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS} \quad (3.13)$$

Keterangan :

$ESS$ : Jumlah kuadrat yang dijelaskan

$RSS$ : Jumlah kuadrat residual.

$TSS$ : Jumlah kuadrat total.

$n$ : Jumlah observasi (negara)

$T$ : Jumlah periode waktu.

$k$ : Banyaknya variabel bebas tanpa intersep.

*Adjusted  $R^2$*  digunakan karena sudah menghilangkan pengaruh penambahan variabel bebas dalam model, karena nilai  $R^2$  akan terus naik seiring dengan penambahan variabel bebas. Karena itu kita harus berhati-hati dalam menggunakan nilai  $R^2$  ketika menilai kebaikan dan kesesuaian suatu model persamaan regresi. Penggunaan *adjusted  $R^2$*  sudah memperhitungkan jumlah derajat bebas.

### c. Uji Statistik $t$

Uji  $t$  digunakan untuk mengetahui apakah variabel-variabel bebas secara parsial berpengaruh signifikan terhadap variabel tak bebasnya.

Hipotesis pengujian:

$$H_0 : \beta_i = 0$$

$$H_1 : \beta_i \neq 0$$

Statistik uji yang digunakan adalah statistik uji *t-student*. Adapun formulanya adalah sebagai berikut:

$$t_{hitung} = \frac{\hat{\beta}_i}{se(\hat{\beta}_i)} \quad (3.14)$$

adalah nilai penduga parameter ke-*i*, adalah simpangan baku dari nilai penduga parameter ke-*i*.

Hipotesis nol ditolak jika . Keputusan ini dapat juga didasarkan pada perbandingan nilai *p-value* dengan tingkat signifikansinya ( $\alpha$ ). Hipotesis nol ditolak jika nilai *p-value* lebih kecil dari ( $\alpha$ ) . Hal ini berarti secara parsial variabel bebas ke-*i* signifikan memengaruhi variabel tidak bebasnya dengan tingkat kepercayaan sebesar  $(1-\alpha) \times 100$  persen.

### 3. Uji Asumsi klasik

Untuk membangun persamaan regresi panel yang terbaik dari kriteria ekonometrika, perlu dilakukan penyelidikan dan penanganan adanya masalah-masalah yang berkaitan dengan pelanggaran asumsi dasar. Berikut ini adalah asumsi-asumsi yang diperlukan dalam analisis regresi:

#### a. Normalitas

Pengujian normalitas dilakukan untuk mengetahui apakah residual berdistribusi normal atau tidak. Berdasarkan dalil limit pusat (*central limit theorem*), ada kecenderungan *residual* yang terjadi sebenarnya menyebar secara

normal. Jika residual  $\varepsilon$ , merupakan jumlah *residual* dari beberapa sumber maka apapun sebaran peluang masing-masing *residual* itu, akan mendekati sebaran normal bila komponen *residual*-nya semakin banyak. Pemeriksaan kenormalan terhadap residual dapat dilakukan menggunakan plot persentil-persentil (P-P Plot). Jika plot mengikuti garis lurus, maka residual mengikuti sebaran normal (Draper dalam Firmansyah, 2009).

Pengujian normalitas dilakukan untuk mengetahui apakah residual berdistribusi normal atau tidak. Berdasarkan dalil limit pusat (*central limit theorem*), ada kecenderungan residual yang terjadi sebenarnya menyebar secara normal. Jika residual merupakan jumlah residual dari beberapa sumber, maka apapun sebaran peluang masing-masing residual itu, akan mendekati sebaran normal bila komponen residual semakin banyak. Salah satu metode yang dapat digunakan untuk menguji Normalitas adalah *Jarque-Bera test*. Uji statistik ini dapat dihitung dengan rumus berikut:

$$JB = n \left[ \frac{\mu_3^2}{6\mu_2^3} + \frac{(\frac{\mu_4}{\mu_2} - 3)^2}{24} \right] \quad (3.15)$$

Keterangan :

$n$  = jumlah sampel

$\mu_2$  = varians

$\mu_3$  = skewness

$\mu_4$  = kurtosis

*Jarque-Bera test* mempunyai distribusi *chi square* dengan derajat bebas dua. Jika hasil *Jarque-Bera test* lebih besar dari nilai *chi square* pada  $\alpha=5$  persen,

maka tolak hipotesis nul yang berarti tidak berdistribusi normal. Jika hasil *Jarque-Bera test* lebih kecil dari nilai *chi square* pada  $\alpha=5$  persen, maka terima hipotesis nul yang berarti *error term* berdistribusi normal.

#### **b. Heteroskedastisitas**

Asumsi dalam model regresi adalah nilai residual memiliki nilai rata-rata nol, residual memiliki varian yang konstan serta residual suatu observasi tidak saling berhubungan dengan residual lainnya sehingga menghasilkan estimator yang BLUE.

Apabila asumsi tidak terpenuhi, yang terpengaruh hanyalah slope estimator dan ini tidak membawa konsekuensi serius dalam analisis ekonometris. Sedangkan apabila asumsi residual memiliki varian yang konstan serta residual suatu observasi tidak saling berhubungan dengan residual lainnya sehingga menghasilkan estimator yang BLUE ini dilanggar, maka akan membawa dampak serius bagi prediksi dengan model yang dibangun.

Kenyataannya, nilai residual sulit memiliki varian yang konstan. Hal ini sering terjadi pada data yang bersifat *cross section* dibanding data *time series*. Untuk mengidentifikasi ada tidaknya masalah *heterokedastisitas*, ada beberapa metode yang dapat digunakan dalam program *Eviews 6.0* ini. Dalam hal ini peneliti menggunakan Uji *Lagrange Multiplier* (LM). Untuk menghilangkan masalah ini peneliti menggunakan penimbang *white cross-section* pada *fixed effect* model.