

## **BAB III**

### **METODOLOGI PENELITIAN**

#### **A. Tujuan Penelitian**

Berdasarkan masalah-masalah yang telah peneliti rumuskan, maka tujuan penelitian ini adalah untuk mendapatkan pengetahuan yang tepat (sahih, benar dan valid) dan dapat dipercaya (dapat diandalkan, reliable) tentang:

1. Pengaruh pertumbuhan sektor pertanian terhadap penyerapan tenaga kerja di Indonesia.
2. Pengaruh pertumbuhan sektor pengangkutan dan komunikasi terhadap penyerapan tenaga kerja di Indonesia.
3. Pengaruh pertumbuhan sektor pertanian dan sektor pengangkutan- komunikasi terhadap penyerapan tenaga kerja di Indonesia.

#### **B. Tempat dan Waktu Penelitian**

Tempat penelitian dipilih di Indonesia. Alasannya karena Indonesia merupakan negara yang mempunyai jumlah penduduk yang tinggi dengan jumlah pengangguran yang mencapai jutaan orang. Tetapi tidak diimbangi dengan pertumbuhan sektor yang ada dalam menyerap tenaga kerja.

Penelitian ini dilakukan menggunakan periode waktu tahun 2005 sampai dengan 2011. Permulaan pada tahun 2005 karena pada tahun ini perekonomian telah dalam kondisi stabil dari masa krisis yang menyebabkan resesi pada tahun

1998, sedangkan batas akhir tahun 2011 karena keterbatasan data yang telah dipublikasikan oleh BPS berada pada tahun tersebut.

### C. Metode Penelitian

Metode penelitian yang digunakan dalam penelitian ini adalah metode *expost facto*. Menurut Kerlinger, penelitian ekspos fakto merupakan pencarian empirik yang sistematis di mana peneliti tidak dapat mengendalikan variabel bebasnya karena peristiwa itu telah terjadi atau sifatnya tidak dapat dimanipulasi. Cara menerapkan metode penelitian ini yaitu dengan menganalisis peristiwa-peristiwa yang terjadi dari tahun-tahun sebelumnya untuk mengetahui faktor-faktor yang dapat menimbulkan kejadian tersebut<sup>42</sup>.

Metode ini dipilih sesuai dengan tujuan penelitian yaitu untuk mengetahui seberapa besar pengaruh antara variabel-variabel yang diteliti yaitu penyerapan tenaga kerja sebagai variabel terikat, serta pertumbuhan sektor pertanian sebagai variabel bebas pertama, dan pertumbuhan sektor pengangkutan dan komunikasi sebagai variabel bebas kedua. Juga dengan analisis regresi data panel dapat melihat fluktuasi suatu tempat pada periode tertentu dan perbedaan beberapa tempat pada satu waktu<sup>43</sup>.

---

<sup>42</sup> Husein Umar. *Metode Penelitian untuk Skripsi dan Tesis Bisnis Edisi 2*. (Jakarta: PT Raja Grafindo Persada, 2009), p. 28.

<sup>43</sup> Nachrowi D Nachrowi dan Hardius Usman, *Pendekatan Populer dan Praktis Ekonometrika untuk Analisis Ekonomi dan Keuangan*. (Jakarta: LPFEUI, 2006), p. 310.

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

Jenis data yang digunakan dalam penelitian ini adalah data sekunder berupa data panel yang merupakan gabungan antara data deret berkala (*time series*) dari tahun 2005-2011 dan *cross section* 30 provinsi di Indonesia. Pada penelitian ini data sekunder diperoleh dari Badan Pusat Statistik Indonesia (BPS), yang meliputi data pertumbuhan PDRB sektor pertanian berdasarkan harga konstan tahun 2000, pertumbuhan PDRB sektor pengangkutan dan komunikasi berdasarkan harga konstan tahun 2000 dan data penduduk yang di daerah maupun perkotaan pada 30 provinsi di Indonesia.

#### **E. Instrumen Penelitian**

##### **1. Penyerapan Tenaga Kerja**

###### **a. Definisi Konseptual**

Penyerapan tenaga kerja adalah banyaknya lapangan kerja yang sudah terisi yang tercermin dari banyaknya jumlah penduduk bekerja. Penduduk yang bekerja terserap dan tersebar di berbagai sektor perekonomian.

###### **b. Definisi Operasional**

Penyerapan tenaga kerja dilihat dari jumlah penduduk baik laki-laki maupun perempuan yang berusia diatas 15 tahun dan bekerja berdasarkan kegiatan utama di daerah dan perkotaan yang terdapat pada 30 provinsi di Indonesia tahun 2005-2011.

## **2. Pertumbuhan Sektor Pertanian**

### **a. Definisi Konseptual**

Sebagai suatu ukuran kuantitatif yang menggambarkan peningkatan output yang dihasilkan dalam suatu tahun tertentu apabila dibandingkan dengan tahun sebelumnya, yang diukur berdasarkan laju dari Produk Domestik Bruto (PDB) semua kegiatan sektor pertanian yang meliputi penyediaan komoditi tanaman bahan makanan, perkebunan, peternakan, kehutanan, dan perikanan.

### **b. Definisi Operasional**

Sektor pertanian diukur melalui data pertumbuhan Produk Domestik Regional Bruto (PDRB) sektor tersebut berdasarkan harga konstan tahun 2000 yang dikeluarkan oleh Badan Pusat Statistik pada 30 provinsi di Indonesia tahun 2005-2011.

## **3. Pertumbuhan Sektor Pengangkutan dan Komunikasi**

### **a. Definisi Konseptual**

Sebagai suatu ukuran kuantitatif yang menggambarkan peningkatan output yang dihasilkan dalam suatu tahun tertentu apabila dibandingkan dengan tahun sebelumnya, yang diukur berdasarkan laju dari Produk Domestik Bruto (PDB) semua kegiatan sektor pengangkutan dan komunikasi yang meliputi semua kegiatan angkutan, jasa, penunjang angkutan, dan komunikasi.

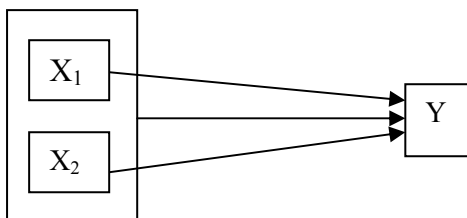
### **b. Definisi Operasional**

Sektor pengangkutan dan komunikasi diukur melalui data pertumbuhan Produk Domestik Regional Bruto (PDRB) sektor tersebut berdasarkan harga

konstan tahun 2000 yang dikeluarkan oleh Badan Pusat Statistik pada 30 provinsi di Indonesia tahun 2005-2011.

## F. Konstelasi Pengaruh Antar Variabel

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



Keterangan:

$X_1$  = Pertumbuhan Sektor Pertanian

$X_2$  = Pertumbuhan Sektor Pengangkutan dan Komunikasi

$Y$  = Penyerapan Tenaga Kerja

→ = Arah Pengaruh

## G. Teknik Analisis Data

### 1. Regresi Data Panel

Model ini menggabungkan observasi lintas sektor dan runtun waktu sehingga jumlah observasi meningkat. Estimasi panel data akan meningkatkan derajat kebebasan, mengurangi kolinearitas antara variabel penjelas dan memperbaiki efisiensi estimasi.

Keunggulan penggunaan data panel dibanding data runtun waktu dan data lintas sektor adalah.<sup>44</sup>

- a. Teknik estimasi data panel dapat mengatasi heterogenitas tersebut secara eksplisit dengan memberikan variabel spesifik-subjek. Subjek disini merupakan istilah sederhana yang mencakup unit-unit mikro seperti individu, perusahaan, negara bagian, dan negara.
- b. Adanya penggabungan observasi *time series* dan *cross section*, data panel memberi lebih banyak informasi, lebih banyak variasi, sedikit kolinearitas antarvariabel, lebih banyak *degree of freedom*, dan lebih efisien.
- c. Data panel cocok untuk mempelajari dinamika perubahan, karena dapat mempelajari observasi *cross section* yang berulang-ulang,
- d. Data panel paling baik untuk mendeteksi dan mengukur dampak yang secara sederhana tidak bias dilihat pada data *cross section* murni atau *time series* murni.
- e. Data panel memudahkan untuk mempelajari model perilaku yang rumit. Contohnya fenomena keekonomian berskala (*economies of scale*) dan perubahan teknologi lebih tepat dipelajari menggunakan data panel daripada data *cross section* murni atau *time series* murni.
- f. Data panel dapat meminimumkan bias yang bisa terjadi jika kita mengagregasi individu-individu atau perusahaan-perusahaan ke dalam agregasi besar.

---

<sup>44</sup>Damodar N Gujarati, 2012. *Dasar-Dasar Ekonometrika*. Buku 2 edisi 5. (Jakarta: Salemba Empat),p.237

Terdapat beberapa teknik untuk mengestimasi parameter model dengan data panel:

**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:

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

$Y$  adalah variabel dependen,  $\alpha$  adalah koefisien regresi,  $X$  adalah variabel independen,  $\beta$  adalah estimasi parameter,  $\epsilon_{it}$  adalah error term,  $N$  adalah jumlah (individu) dan  $T$  adalah jumlah periode waktu.

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 Efek Tetap (*Fixed Effect*)**

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  $X_{it}$ .

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.

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

Maka 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:

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

Keterangan:

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

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

$W_{it}$  dan  $Z_{it}$  variabel dummy yang didefinisikan sebagai berikut:

$$\begin{aligned} W_{it} &= 1; \text{ untuk negara } i; & i &= 1, 2, \dots, N=0; \text{ lainnya} \\ Z_{it} &= 1; \text{ untuk tahun } t; & t &= 1, 2, \dots, T=0; \text{ lainnya} \end{aligned}$$

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 Efek Random (*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*.

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:

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

Keterangan:

$$\begin{aligned} u_i \sim N(0, \delta_u^2) &= \text{error component cross section} \\ v_t \sim N(0, \delta_v^2) &= \text{error component time series} \\ w_{it} \sim N(0, \delta_w^2) &= \text{error component combinations} \end{aligned}$$

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  $u_i$ .

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

$$\text{Var} (\varepsilon_{it}) = \delta_U^2 + \delta_V^2 + \delta_W^2 \quad (3.5)$$

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

$$\text{Var} (\varepsilon_{it}) = \delta_W^2 \quad (3.6)$$

Jadi 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* atau *Random Effect* memberikan hasil yang sama.

## 2. Pemilihan Estimator dengan Struktur Varian Kovarian Residual

### a. Pengujian Signifikansi *Common Effect* atau *Fixed effect*

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).

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

Di mana,  $n$  adalah jumlah individu,  $T$  adalah periode waktu,  $k$  adalah parameter dalam model *fixed effect*,  $RSS_1$  dan  $RSS_2$  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*.

#### **b. Pengujian Signifikansi *Common Effects* atau *Random Effect***

Pengujian *Lagrange Multiplier* dilakukan untuk mengetahui signifikansi dari *random effect* berdasarkan residual dari OLS (*common effect*). Secara matematis, statistik uji untuk LM test (*Lagrange Multiplier*) dapat dituliskan sebagai berikut:

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

Dibawah 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*.

### c. Pengujian Signifikansi *Fixed Effect* atau *Random effect*

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  $\hat{\beta}_{FE}$  dan  $\hat{\beta}_{RE}$  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:

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

Terletak dibawah 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*.

### 3. Uji Hipotesis

#### a. Uji t

Uji t yaitu suatu uji untuk mengetahui signifikansi dari pengaruh variabel independen terhadap variabel dependen secara individual dan menganggap dependen yang lain konstan. Signifikansi pengaruh tersebut dapat diestimasi dengan membandingkan antara nilai  $t_{\text{tabel}}$  dengan nilai  $t_{\text{hitung}}$ . Apabila nilai  $t_{\text{hitung}}$  lebih besar daripada  $t_{\text{tabel}}$  maka variabel independen secara individual mempengaruhi variabel dependen, sebaliknya jika nilai  $t_{\text{hitung}}$  lebih kecil daripada  $t_{\text{tabel}}$  maka variabel independen secara individual tidak mempengaruhi variabel dependen. Tahap-tahap yang digunakan:<sup>45</sup>

#### 1) Merumuskan Hipotesis

- a)  $H_0 : b = 0$ , yaitu tidak ada pengaruh yang signifikan pada variabel independen terhadap variabel dependen.
- b)  $H_a : b \neq 0$ , yaitu terdapat pengaruh yang signifikan pada variabel independen terhadap variabel dependen.

#### 2) Mencari t hitung

$$t_h = \frac{\text{koefisien } \beta}{\text{standar error}}$$

---

<sup>45</sup> Imam Ghozali. *Ekonometrika*. Semarang : Badan Penelitian Universitas Diponegoro, 2009.,p.59

### 3) Kriteria Pengujian

$t_{hitung} > t_{kritis}$  berarti  $H_0$  ditolak dan menerima  $H_a$

$t_{hitung} \leq t_{kritis}$  berarti  $H_0$  diterima dan menolak  $H_a$

Uji t juga bisa dilihat pada tingkat signifikansinya :

- a) Jika tingkat signifikansi  $< 0.05$  maka  $H_0$  ditolak dan  $H_a$  diterima.
- b) Jika tingkat signifikansi  $> 0.05$  maka  $H_0$  diterima dan  $H_a$  ditolak.

### b. Uji F

Uji F atau uji koefisien regresi secara simultan, Uji F ini dilakukan dengan menggunakan analisis varian. Analisis varian dalam regresi berganda pada hakikatnya diperlukan untuk menunjukkan sumber-sumber variasi yang menjadi komponen dari variabel total model regresi. Dengan analisis varian ini akan dapat diperoleh pengertian tentang bagaimana pengaruh sekelompok variabel bebas secara bersama-sama terhadap variabel tidak bebas<sup>46</sup>.

Hipotesis penelitiannya :

- 1)  $H_0$  ;  $b_1 = b_2 = 0$ , Artinya variabel  $x_1$  dan  $x_2$  secara serentak tidak berpengaruh terhadap Y
- 2)  $H_a$  :  $b_1 \neq b_2 \neq 0$ , Artinya variabel  $X_1$  dan  $X_2$  secara serentak berpengaruh terhadap Y.

Kriteria pengambilan keputusan, yaitu:

---

<sup>46</sup> Muhammad Firdaus, *Ekonometrika Suatu Pendekatan Aplikatif*. (Jakarta: PT Bumi Aksara. 2008), p. 88.

F hitung  $\leq$  F tabel, jadi  $H_0$  diterima

F hitung  $\geq$  F tabel, jadi  $H_0$  ditolak

### c. Uji Koefisien Determinasi ( $R^2$ )

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

$$R^2 = \frac{\sum (\hat{Y}_i - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2} = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS} \quad (3.10)$$

*Adjusted R<sup>2</sup>* dihitung dengan rumus:

$$\overline{R^2} = 1 - (1 - R^2) \frac{nT - 1}{nT - n - k} \quad (3.11)$$

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<sup>2</sup>* 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<sup>2</sup>* sudah memperhitungkan jumlah derajat bebas.



#### 4. Uji Asumsi Klasik

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

##### a. Normalitas

Pengujian normalitas dilakukan untuk melihat apakah suatu data berdistribusi normal atau tidak. Normalitas dapat dilihat dari gambar histogram, namun sering kali polanya tidak mengikuti bentuk kurva normal sehingga sulit disimpulkan. Untuk memudahkan dalam mengetahui normalitas model dilihat dari koefisien *Jarque-Bera test* dan probabilitasnya. Kedua angka ini bersifat saling mendukung yakni:<sup>47</sup>

- Bila nilai J-B tidak signifikan (lebih kecil dari 2,0) maka data berdistribusi normal.
- Bila probabilitasnya lebih besar dari 5% (0,05) maka data berdistribusi normal (hipotesis nol dianggap berdistribusi normal).

##### b. Multikolinieritas

Multikolinieritas merupakan suatu keadaan dimana satu atau lebih variabel bebas terdapat korelasi dengan variabel bebas lainnya atau suatu variabel bebas merupakan fungsi linear dari variabel bebas lainnya. Adanya multikolinieritas

---

<sup>47</sup> Nachrowi Djalal, dkk. *Penggunaan Teknik Ekonometri Edisi Revisi*. (Jakarta: PT Raja Grafindo persada, 2002), p.22.

menurut Heir et.al. dilihat dari *tolerance value* atau *variance inflation factor (FIV)*. *Tolerance Value* adalah suatu jumlah yang menunjukkan bahwa variabel bebas tidak dapat dijelaskan oleh variabel bebas lainnya dalam suatu nilai yang menunjukkan tidak adanya multikolinieritas dalam persamaan regresi. Batas dari *tolerance value* adalah 0,1 maka terjadi multikolinieritas<sup>48</sup>.

*Variance Inflation Factor (VIF)* merupakan kebalikan dari *tolerance value*, karena  $VIF = 1/\textit{tolerance value}$ , VIF merupakan suatu jumlah yang menunjukkan bahwa suatu variabel bebas dapat dijelaskan oleh variabel bebas lainnya dalam persamaan regresi atau dapat dikatakan VIF menunjukkan adanya multikolinieritas dalam persamaan regresi. Batas VIF adalah 10 jika nilai VIF diatas 10 maka terjadi multikolinieritas. Menurut Gujarati ada atau tidaknya multikolinieritas dapat dihitung dengan *Variance Inflation Factor (VIF)* dengan rumus:<sup>49</sup>

$$VIF = \frac{1}{1 - R^2_i}$$

Keterangan :

VIF : *Variance Inflation Factor*

$R^2_i$  : Koefisien determinasi yang diperoleh dengan meregresikan satu variabel bebas dengan variabel bebas lainnya.

---

<sup>48</sup> Anonim. *Ekonometrika Dasar, Modul Ajar Ekonometrika*. p.59 (digilib.uns.ac.id) diakses pada 28 Juni 2012 pukul 21.05

<sup>49</sup> *Ibid.*,p.60

### c. Heteroskedastisitas

pengujian ada atau tidaknya masalah heteroskedastisitas dalam penelitian menggunakan uji White. Uji White dilakukan dengan mengendalikan residual kuadrat sebagai variabel dependen ditambah dengan kuadrat variabel independen, kemudian ditambahkan lagi dengan perkalian dua variabel independen. Prosedur pengujian dilakukan dengan hipotesis sebagai berikut:<sup>50</sup>

- $H_0$  : Tidak ada heteroskedastisitas
- $H_1$  : Ada heteroskedastisitas

Jika  $\alpha = 5\%$ , maka tolak  $H_0$  jika  $obs * R\text{-square} > X^2$  atau  $P\text{-value} < \alpha$ . Sebaliknya model terkena masalah heteroskedastisitas jika  $obs * R\text{-square} < X^2$  atau  $P\text{-value} > \alpha$ .

### d. Autokorelasi

Rumus yang digunakan adalah rumus Uji Durbin Watson (Uji DW) yaitu:

$$DW = \frac{\sum_{t=1}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2}$$

Keterangan:

DW : Nilai Durbin Watson

$e_t$  : Nilai residual periode t

$e_{t-1}$  : Nilai residual periode t-1

- Menentukan nilai  $d_U$  dan  $d_L$  dari tabel d berdasarkan observasi (n) dan jumlah variabel. Menentukan kaidah keputusan dengan uji 2 arah, di mana:

$H_0$  :  $\rho \leq 0$  maka tidak ada autokorelasi

---

<sup>50</sup> Wing Wahyu Winarno. *Analisis Ekonometrika dan Statistika Dengan Eviews (Edisi 2)*. (Yogyakarta: UPP STIM YKPN, 2001), P.54

$H_1 : \rho > 0$  maka ada gejala autokorelasi

- Membandingkan  $d$  hitung dengan  $d$  tabel
  - $d < d_L$ , tolak  $H_0$  : ada autokorelasi
  - $d > d_U$ , gagal tolak  $H_0$  : tidak ada autokorelasi
  - $d_L \leq d \leq d_U$  : tidak dapat disimpulkan
- Dapat juga dengan melihat nilai-nilai sebagai berikut:
  1. Kurang dari 1,10 = Ada autokorelasi
  2. 1,10 s/d 1,54 = Tanpa kesimpulan
  3. 1,55 s/d 2,46 = Tidak ada autokorelasi
  4. 2,46 s/d 2,90 = Tanpa kesimpulan
  5. Lebih dari 2,91 = Ada autokorelasi