

BAB III

METODOLOGI PENELITIAN

A. Tujuan Penelitian

Berdasarkan masalah – masalah yang telah peneliti rumuskan, maka tujuan penelitian ini adalah mutlak :

1. Mengetahui besarnya pengaruh antara pertumbuhan ekonomi terhadap Indeks Pembangunan Manusia di Indonesia
2. Mengetahui besarnya pengaruh antara pengeluaran pemerintah sektor pendidikan terhadap Indeks Pembangunan Manusia di Indonesia
3. Mengetahui besarnya pengaruh antara pertumbuhan ekonomi dan pengeluaran pemerintah sektor pendidikan terhadap Indeks Pembangunan Manusia di Indonesia

B. Objek Penelitian

Penelitian ini dilaksanakan dengan mengambil data Pertumbuhan Ekonomi, dan Indeks Pembangunan Manusia (IPM) diambil dari Badan Pusat Statistik, serta data pengeluaran pemerintah pada sektor pendidikan diambil dari Kementerian Pendidikan dan Kebudayaan (KEMENDIKBUD).

Data yang digunakan adalah data panel yaitu data PDRB atas dasar harga berlaku, Realisasi APBD Pendidikan dan Indeks Pembangunan Manusia (IPM) pada tahun 2007-2010 di 33 provinsi di Indonesia, total data adalah 132.

C. Metode Penelitian

Metode yang digunakan dalam penelitian ini adalah metode *expost facto*, yang merupakan suatu penelitian yang dilakukan untuk meneliti peristiwa yang telah terjadi dan kemudian meruntut kebelakang untuk mengetahui faktor yang dapat menimbulkan kejadian tersebut. Metode ini digunakan untuk memperoleh data sekunder.⁴⁸

D. Jenis dan Sumber Data

Jenis data yang digunakan dalam penelitian ini adalah data sekunder yang bersifat kuantitatif yaitu data yang telah tersedia dalam bentuk angka. Sedangkan data yang digunakan dalam penelitian ini adalah data panel yang merupakan gabungan antara data *time series* (runtun waktu) dan data *cross section* (silang).⁴⁹ Data *time series* sebanyak empat tahun dari tahun 2007-2010 dan data *cross section* sebanyak 33 provinsi di Indonesia. Data tersebut diperoleh dari studi kepustakaan yang bersumber dari publikasi yang diterbitkan oleh Badan Pusat Statistik (BPS) untuk memperoleh data pertumbuhan ekonomi, dan indeks pembangunan manusia, serta Pengeluaran Pemerintah sektor Pendidikan yang dipublikasikan oleh Kementerian Pendidikan dan Kebudayaan (Kemendikbud) dengan jumlah data sebanyak 132.

⁴⁸Sugiyono. *Metode Penelitian Bisnis*. (Jakarta: Alfabeta, 2004), p. 7

⁴⁹Nachrowi, *Pendekatan Populer dan Praktis Ekonometrika untuk Analisis Ekonomi dan Keuangan* (Jakarta: LPFE UI, 2006), p.309

E. Operasionalisasi Variabel Penelitian

1. IPM

a. Definisi Konseptual

Indeks Pembangunan Manusia (IPM) / Human Development Index (HDI) adalah pengukuran perbandingan dari harapan hidup, melek huruf, pendidikan dan standar hidup untuk semua negara seluruh dunia.

b. Definisi Operasional

Indeks Pembangunan Manusia (IPM) didapat dari Rata-rata indikator IPM. Data IPM dilihat dari data Indeks Pembangunan Manusia yang diterbitkan oleh Badan Pusat Statistik (BPS). Data IPM yang dipergunakan dalam penelitian ini adalah data IPM Per provinsi di Indonesia pada tahun 2007 – 2010.

2. Pertumbuhan Ekonomi

a. Definisi Konseptual

Pertumbuhan ekonomi adalah proses kegiatan perekonomian yang menyebabkan perubahan kapasitas produksi dalam jangka panjang yang ditunjukkan dengan adanya perubahan Gross Domestic Product (GDP).

b. Definisi Operasional

Data pertumbuhan ekonomi untuk penelitian ini menggunakan data Laju Produk Domestik Bruto (PDRB). Data pertumbuhan

ekonomi ini diperoleh dari data laju pertumbuhan Produk Domestik Regional Bruto (PDRB) atas dasar harga konstan 2000 yang dikeluarkan oleh Badan Pusat Statistik (BPS). Data yang dipergunakan dalam penelitian ini adalah data tahun 2007 – 2010 di 33 provinsi di Indonesia.

3. Pengeluaran Pemerintah Sektor Pendidikan

a. Definisi Konseptual

Pengeluaran pemerintah daerah sektor pendidikan adalah pengeluaran keuangan pemerintah yang mencerminkan pilihan kebijaksanaan untuk suatu periode pengeluaran yang dikategorikan sebagai investasi pembangunan pada sektor pendidikan.

b. Definisi Operasional

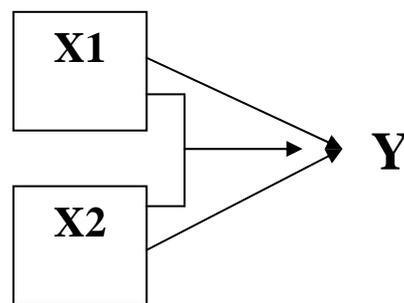
Data Pengeluaran pemerintah sector pendidikan dapat dilihat dari data Realisasi Anggaran Pemerintah Sektor Pendidikan. Data realisasi anggaran pemerintah sektor pendidikan ini diperoleh dari laporan Kementerian Pendidikan dan Kebudayaan (KEMENDIKBUD). Data yang dilihat yaitu Realisasi Anggaran sektor pendidikan Perprovinsi di Indonesia pada tahun 2007-2010.

F. Konstelasi Pengaruh Antar Variabel

Variabel penelitian ini terdiri dari tiga variabel yaitu variabel bebas (pertumbuhan ekonomi digambarkan dengan simbol X_1 dan pengeluaran

pemerintah sektor pendidikan digambarkan dengan simbol X2) dan variabel terikat tingkat IPM yang digambarkan dengan simbol Y.

Sesuai dengan hipotesis yang diajukan bahwa terdapat pengaruh variabel X1 dan X2 terhadap variabel Y, maka konstelasi pengaruh variabel X1 dan X2 terhadap variabel Y adalah :



Gambar III.2

Keterangan :

Variabel Bebas (X1) : pertumbuhan ekonomi

Variabel Bebas (X2) : pengeluaran pemerintah sektor pendidikan

Variabel Terikat (Y) : IPM (Indeks Pembangunan Manusia)

—————> : Menunjukkan 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.⁵⁰ 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 :

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:

$$IPM = \alpha + \beta PE + \beta Peng + \varepsilon_{it} \quad i=1, 2..N; t=1, 2.. T \quad (3.1)$$

dimana :

- IPM = Indeks Pembangunan Manusia
- N = banyaknya observasi (provinsi)
- T = banyaknya waktu
- N x T = banyaknya data panel
- PE = Pertumbuhan Ekonomi
- Peng = Pengeluaran Pemerintah Sektor Pendidikan

⁵⁰ Wing Wahyu Winarno, *Analisis Ekonometrika dan Statistika dengan Eviews Edisi Ke-3* (Yogyakarta; STIM YKPN, 2011). p.1.2

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

a. Model *Common Effects*

Model *common effects* 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 effects* dapat dituliskan sebagai berikut:

$$IPM = \alpha + \beta PE + \beta Peng + \varepsilon_{it} \quad i=1, 2..N; t=1, 2.. T \quad (3.2)$$

dimana:

IPM	= variabel dependen (Indeks Pembangunan Manusia)
α	= koefisien regresi
PE	= variabel independen (Pertumbuhan Ekonomi)
Peng	= variabel independen (Pengeluaran Pemerintah sektor Pendidikan)
β	= estimasi parameter
ε_{it}	= <i>error term</i>
N	= jumlah (provinsi)
T	= 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 *Fixed Effects*

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

Berdasarkan asumsi struktur matriks varians-kovarians residual, maka pada model *fixed effects*, 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*.

c. Model *Random Effects*

Keputusan untuk memasukkan variabel *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:

$$\text{IPM} = \alpha + \beta \text{PE} + \beta \text{Peng} + \varepsilon_{it} \quad \varepsilon_{it} = u_i + v_t + w_{it} \quad (3.4)$$

Di mana:

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

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

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

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

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

$$F = \frac{(RSS_1 - RSS_2) / (n-1)}{RSS_2 / (nT - n - k)} \quad (3.5)$$

Di mana, n adalah jumlah individu, T adalah periode waktu, k adalah parameter dalam model *fixed effects*, RSS_1 dan RSS_2 masing-masing merupakan *residual sum of squares*. 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 effects* lebih baik dari *common effects*.

2) Pengujian Signifikansi *Common Effects* atau *Random Effects*

Secara matematis, statistik uji untuk *LM test* (*Lagrange Multiplier*) dapat dituliskan sebagai berikut:

$$LM = \frac{nT}{2(T-1)} \left[\frac{\sum_{i=1}^n (\sum_{t=1}^T e_{it})^2}{\sum_{i=1}^n \sum_{t=1}^T e_{it}^2} - 1 \right]^2 \quad (3.6)$$

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 effects* 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:

$$H = (\hat{\beta}_{FE} - \hat{\beta}_{RE}) [\text{var}(\hat{\beta}_{FE}) - \text{var}(\hat{\beta}_{RE})]^{-1} (\hat{\beta}_{FE} - \hat{\beta}_{RE}) \quad (3.7)$$

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 effects* lebih baik bila dibandingkan dengan model *fixed effects*. Di

bawah hipotesis nul, pendugaan parameter dengan menggunakan *random effects* adalah konsisten dan efisien, sedangkan pendugaan dengan *fixed effects* meskipun tetap konsisten, tetapi tidak lagi efisien. Di bawah hipotesis alternatif, estimasi dengan *random effects* menjadi tidak konsisten, sebaliknya estimasi dengan *fixed effects* 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 effects* dari pada metode *random effects*.

Sementara itu, dalam memberikan sejumlah pertimbangan terkait pilihan apakah menggunakan model *fixed effects* ataukah model *random effects*. Pertimbangan-pertimbangan itu adalah sebagai berikut:

- a) Jika jumlah data *time series* (T) besar dan jumlah data *cross-section* (N) kecil, ada kemungkinan perbedaan nilai parameter yang diestimasi dengan *Fixed Effects* dan *Random Effects* cukup kecil. Karena itu, pilhan ditentukan berdasarkan kemudahan perhitungan. Dalam hal ini, adalah model FE.
- b) 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 *Fixed Effects*. Jika sebaliknya, maka digunakan model *Random Effects*.

- c) Jika efek individu tidak teramati α_i berkorelasi dengan satu atau lebih variabel bebas, maka estimasi dengan *Random Effects* bias, sedangkan estimasi dengan *Fixed Effects* tidak bias.
- d) Jika N besar dan T kecil, serta semua asumsi yang disyaratkan oleh model *Random Effects* terpenuhi, maka estimasi dengan menggunakan *Random Effects* lebih efisien dibanding estimasi dengan *Fixed Effects*.

Dalam penelitian ini, penentuan apakah model *Fixed Effects* atau *Random Effects* 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 Linier 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 effects*. 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 effects*.

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 :

$$LM = \frac{T}{2} \sum_{i=1}^n \left[\frac{\hat{\sigma}_i^2}{\hat{\sigma}^2} - 1 \right]^2 \quad (3.8)$$

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.

Uji Hipotesis

a. Uji 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:

$$F_{(n-k-1, nT-n-k)} = \frac{R^2 / (n+k-1)}{(1-R^2) / (nT-n-k)} \quad (3.10)$$

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 (α).

b. Koefisien Determinasi (R^2)

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

$$R^2 = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2} = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS} \quad (3.11)$$

Dan R^2 *adjusted* dihitung dengan rumus:

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

Dimana:

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 R²* sudah memperhitungkan jumlah derajat bebas.

c. Uji 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:

$$t_{\text{hitung}} = \frac{\hat{\beta}_i}{\text{se}(\hat{\beta}_i)} \quad (3.12)$$

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

Hipotesis nol ditolak jika $t_{\text{hitung}} > t_{\frac{\alpha}{2}}(nT - n - k)$. Keputusan ini dapat juga didasarkan pada perbandingan nilai *p-value* dengan tingkat signifikansinya (α). Hipotesis nol ditolak jika nilai *p-value* lebih kecil dari (α). Hal ini berarti secara parsial variabel bebas ke- i signifikan mempengaruhi variabel tidak bebasnya dengan tingkat kepercayaan sebesar $(1-\alpha) \times 100$ persen.

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 uji asumsi yang digunakan dalam penelitian ini:

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 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:

$$JB = \frac{n}{6} \left[S^2 + \frac{(K-3)^2}{4} \right] \quad (3.13)$$

Di mana:

n = jumlah sampel

S = kemencengan

K = peruncingan

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. Multikolinieritas

Multikolinieritas berarti adanya kolerasi antar variabel bebas, yang terjadi karena variabel – variabel bebas tersebut memiliki hubungan pada populasi atau hanya pada sampel. Cara mendeteksi adanya kolinieritas yaitu:

1. Dengan memeriksa *simple pairwise* (Pearson) correlation antar variabel independen. Batas nilai yang disarankan sebagai indikasi kolinieritas serius berbeda – beda. Suatu variabel independen berkorelasi tinggi dengan variabel independen lainnya jika r lebih dari 0,85.
2. Uji formal untuk mendeteksi keberadaan dari multikolinieritas yaitu dengan melihat nilai *Variance Inflation Factor* (VIF). Nilai VIF ini menunjukkan bagaimana varians dari sebuah estimator akan meningkat akibat adanya multikolinieritas. Nilai VIF diperoleh dengan formula berikut:

$$VIF_k = \frac{1}{1-R_k^2} \quad (3.14)$$

Di mana $k=1,2,\dots, p - 1$ dan R_k^2 merupakan koefisien determinasi dari regresi berganda ketika X_k diregresikan dengan $p - 2$ variabel lainnya dalam model. Jika nilai VIF lebih dari 10, maka hal tersebut dapat berindikasi bahwa multikolinieritas bersifat serius dan akan mempengaruhi estimasi yang menggunakan OLS karena meskipun estimator tetap bersifat *unbiased* namun sudah tidak lagi memiliki varians yang minimum. Selain itu, keberadaan multikolinieritas juga akan membuat estimator bersifat sensitif untuk perubahan yang kecil pada data, sehingga akan mengakibatkan kesalahan (*misleading*) dalam menginterpretasikan suatu model regresi. Cara mengatasi adanya multikolinieritas antara lain melepas satu atau lebih variabel yang memiliki korelasi yang tinggi, mentransformasi model, atau memperbesar jumlah sampel.

c. Heteroskedastisitas

Asumsi dalam model regresi adalah nilai residual memiliki nilai rata-rata nol, residual memiliki varians 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.

Dalam 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, secara sistematis, statistik uji yang digunakan dapat dirumuskan sebagai berikut:

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

Hasil uji LM dibandingkan dengan nilai *Chi-square* tabel dengan derajat (n-1) dengan ketentuan H0 ditolak jika nilai LM lebih besar dari *Chi-square* yang berarti model yang terbentuk mengandung masalah *heteroskedastisitas*.