

**PEMODELAN REGRESI DATA PANEL *RANDOM EFFECT* DENGAN
METODE *GENERALIZED LEAST SQUARE* (GLS)**

(Skripsi)

Oleh

**ROSITA SARI
NPM 1717031010**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2021**

ABSTRACT

RANDOM EFFECT PANEL DATA REGRESSION MODELING WITH GENERALIZED LEAST SQUARE (GLS) METHOD

By

ROSITA SARI

Panel data are combination of cross section data and time series data. One of panel data regression analysis models is the random effect model. Random effect model has the assumption that error are different for each individual, but the intercepts is constant. The random effect estimation is done using generalized least square method and applied the models to the human development index data in Lampung in year of 2015-2018. Based on the result, the estimated random effect model of the panel data regression analysis is,

$$\hat{Y}_{it} = 7,215355 + 0,512942X_{1it} + 1,573944X_{2it} + 0,001357X_{3it}$$

with,

\hat{Y}_{it} = human development index of district i time t

X_{1it} = life expectancy of district i time t

X_{2it} = average length of school of district i time t

X_{3it} = adjusted per capita income of district i time t

Keywords: panel data regression analysis, random effect models, GLS

ABSTRAK

PEMODELAN REGRESI DATA PANEL *RANDOM EFFECT* DENGAN METODE *GENERALIZED LEAST SQUARE* (GLS)

Oleh

ROSITA SARI

Data panel merupakan gabungan antara data *cross section* dan data *time series*. Salah satu model analisis regresi data panel adalah model *random effect*. Model *random effect* mengasumsikan bahwa *error* berbeda untuk setiap individu, tetapi intersep konstan. Estimasi model *random effect* dilakukan dengan menggunakan metode *generalized least square* dan menerapkannya pada data indeks pembangunan manusia di provinsi Lampung tahun 2015-2018. Berdasarkan hasil penelitian diperoleh model regresi data panel *random effect* yaitu,

$$\hat{Y}_{it} = 7,215355 + 0,512942X_{1it} + 1,573944X_{2it} + 0,001357X_{3it}$$

dengan,

\hat{Y}_{it} = Indeks Pembangunan Manusia (IPM) di kabupaten ke-i dan tahun ke-t

X_{1it} = Angka Harapan Hidup (AHH) di kabupaten ke-i dan tahun ke-t

X_{2it} = Rata-rata Lama Sekolah (RLS) di kabupaten ke-i dan tahun ke-t

X_{3it} = Pendapatan Perkapita Disesuaikan di kabupaten ke-i dan tahun ke-t

Kata kunci: analisis regresi data panel, model *random effect*, GLS

ESTIMASI PARAMETER MODEL REGRESI DATA PANEL *RANDOM EFFECT* DENGAN METODE *GENERALIZED LEAST SQUARE (GLS)*

Oleh

ROSITA SARI

Skripsi

Sebagai Salah Satu Syarat Untuk Mencapai Gelar
SARJANA MATEMATIKA

Pada

Jurusan Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2021**

Judul Skripsi : **PEMODELAN REGRESI DATA PANEL RANDOM
EFFECT DENGAN METODE GENERALIZED
LEAST SQUARE (GLS)**

Nama Mahasiswa : **Rosita Sari**

Nomor Pokok Mahasiswa : **1717031010**

Jurusan : **Matematika**

Fakultas : **Matematika dan Ilmu Pengetahuan Alam**



1. Komisi Pembimbing

Dr. Khoirin Nisa, S.Si., M.Si.
NIP 19740726 200003 2 001

Dr. Aang Nuryaman, S.Si., M.Si.
NIP 19740316 200501 1 001

2. Ketua Jurusan Matematika

Dr. Aang Nuryaman, S.Si., M.Si.
NIP 19740316 200501 1 001

MENGESAHKAN

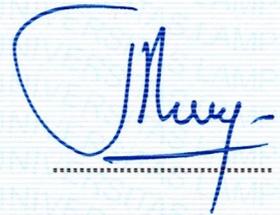
1. Tim Penguji

Ketua Penguji : **Dr. Khoirin Nisa, S.Si., M.Si.**



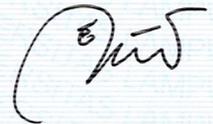
.....

Sekretaris : **Dr. Aang Nuryaman, S.Si., M.Si.**



.....

Penguji
Bukan Pembimbing : **Drs. Eri Setiawan, M.Si.**



.....



2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam



Dr. Eng. Suropto Dwi Yuwono, M.T.
NIP 19740705 200003 1 001

Tanggal Lulus Ujian Skripsi : **19 Oktober 2021**

PERNYATAAN SKRIPSI MAHASISWA

Saya yang bertanda tangan di bawah ini:

Nama : **Rosita Sari**

Nomor Pokok Mahasiswa : **1717031010**

Jurusan : **Matematika**

Judul Skripsi : **Pemodelan Regresi Data Panel *Random Effect* dengan Metode *Generalized Least Square (GLS)***

Dengan ini menyatakan bahwa skripsi ini merupakan hasil pekerjaan saya sendiri dan semua tulisan yang tertuang dalam skripsi ini telah mengikuti kaidah penulisan karya ilmiah Universitas Lampung. Apabila kemudian hari terbukti bahwa skripsi ini merupakan hasil salinan atau dibuat oleh orang lain, maka saya bersedia menerima sanksi sesuai dengan ketentuan akademik yang berlaku.

Bandar Lampung, Oktober 2021

Penulis,



Rosita Sari
NPM. 17170031010

RIWAYAT HIDUP

Penulis dilahirkan di Waykamal pada tanggal 30 November 1998 sebagai anak pertama dari pasangan Bapak Rohimuddin dan Ibu Marhaiyah.

Penulis menempuh pendidikan di Sekolah Dasar Negeri (SDN) 1 Negeri Ratu pada tahun 2005-2011, Sekolah Menengah Pertama Negeri (SMPN) 1 Kota Agung pada tahun 2011-2014 dan Sekolah Menengah Atas Negeri (SMAN) 1 Kota Agung pada tahun 2014-2017.

Pada tahun 2017 penulis terdaftar sebagai mahasiswa Program Studi S1 Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam (FMIPA) Universitas Lampung melalui jalur masuk SNMPTN. Penulis juga merupakan mahasiswa penerima beasiswa bidikmisi. Selama menjadi mahasiswa penulis memiliki pengalaman organisasi diantaranya yaitu menjadi Anggota Biro Kesekretariatan Himpunan Mahasiswa Jurusan Matematika (HIMATIKA) tahun 2018, Anggota Biro Kemuslimahan Rohani Islam (ROIS) FMIPA Universitas Lampung tahun 2018, Sekretaris Biro Kesekretariatan HIMATIKA tahun 2019, Sekretaris Dinas Advokasi dan Kesejahteraan Mahasiswa Badan Eksekutif Mahasiswa (BEM) FMIPA Universitas Lampung tahun 2020.

Pada bulan Januari sampai Februari tahun 2020 penulis melaksanakan Kerja Praktek di Badan Pusat Statistik (BPS) Kabupaten Tanggamus sebagai bentuk penerapan ilmu yang telah diperoleh selama kuliah. Pada bulan Juli sampai Agustus 2020 penulis melaksanakan Kuliah Kerja Nyata (KKN) di Desa Babakan, Kecamatan Pugung, Kabupaten Tanggamus sebagai bentuk pengabdian mahasiswa dan menjalankan Tri Dharma Perguruan Tinggi.

KATA INSPIRASI

Untuk jiwa pemilik semangat yang tak pernah padam.

Untuk jiwa pemilik mimpi yang tak pernah habis.

Untuk jiwa pemilik doa yang tak pernah putus.

Hidup seperti membaca buku, jika tidak pernah membuka halaman maka tidak akan pernah tau apa yg akan terjadi berikutnya.

Menyerah atau bertahan, sama-sama berat. Tapi pada akhirnya memilih bertahan akan lebih baik, karena barangkali ada penyesalan, lebih baik menyesal telah bertahan dari pada menyesal tidak bertahan dan tak pernah melakukan apapun.

Kesedihan yang menumpuk tidak akan pernah lepas dari akhir yang bahagia.

Tak apa untuk tidak baik-baik saja. Butuh waktu, butuh sabar, butuh keyakinan untuk sembuh.

Tak apa untuk rehat sejenak, lalu lekas mengusahakan. Karena tidak ada ujian yang mudah untuk hidup yang lebih indah.

Terima kasih sudah sangat hebat dalam menghadapi dunia dan kerumitannya.

Mimpi setinggi dan sebanyak apapun. Tetap berusaha hingga nanti doa dan Tuhan yang akan menentukan hasilnya.

Untukmu, titipan terindah dari Tuhan untuk ayah dan Ibu

(Rosita Sari)

PERSEMBAHAN

Puji Syukur kepada Allah SWT yang telah memberikan kemudahan dalam menyelesaikan skripsi ini. Dengan segala kerendahan hati, ku persembahkan karya syarat perjuangan kelulusan ku ini kepada,

Ayahanda Rohimuddin dan Ibu Marhaiyah

Terima kasih telah memberikan kasih sayang yang tulus, tetes keringat pengorbanan, semangat dan motivasi yang tiada henti, doa yang tak pernah terputus, dan sabar yang tak pernah habis untuk menanti keberhasilanku. Atas doa dan ridho kalian, Allah beri kemudahan dalam menjalankan kehidupan ini.

Kakak, adik dan keluarga

Yang telah memberi semangat, bantuan serta doa yang tulus untuk selalu berusaha dan berikhtiar kepada Allah SWT.

Dosen Pembimbing dan Pembahas

Yang senantiasa membimbing, mengarahkan dan memberi motivasi sejak awal hingga terselesaikannya skripsi ini.

Almamater tercinta, Universitas Lampung

SANWACANA

Penulis mengucapkan puji syukur kehadirat Allah SWT, karena berkat ridho dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul “Pemodelan Regresi Data Panel *Random Effect* dengan Metode *Generalized Least Square (GLS)*”. Selesaiannya penulisan skripsi ini adalah berkat bimbingan serta motivasi dari berbagai pihak. Dengan segala kerendahan dan ketulusan hati penulis ingin menyampaikan ucapan terima kasih kepada :

1. Dr. Khoirin Nisa, M.Si., selaku pembimbing pertama yang telah memberikan arahan, bimbingan, saran serta motivasi kepada penulis dalam menyelesaikan skripsi ini.
2. Dr. Aang Nuryaman, S.Si., M.Si., selaku pembimbing kedua sekaligus Ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam (FMIPA) Universitas Lampung yang telah memberikan arahan, bimbingan dan saran kepada penulis dalam menyelesaikan skripsi ini.
3. Drs. Eri Setiawan, M.Si., selaku pembahas yang telah memberikan kritik dan saran hingga terselesaikannya skripsi ini.
4. Dr. Fitriani, M. Sc., selaku pembimbing akademik yang telah memberikan bimbingan dan motivasi kepada penulis selama menuntut ilmu di Universitas Lampung.
5. Dr. Eng. Suripto Dwi Yuwono, M.T., selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
6. Para Dosen dan Staf Jurusan Matematika FMIPA Universitas Lampung.
7. Emak, Ebak, Kakak, Adik dan Keluarga tersayang yang selalu memberi semangat, motivasi dan doa yang tak terputus kepada penulis.

8. Srikandi Ku (Dwi, Simun dan Fifi) yang telah memberikan kebahagiaan dan keceriaan selama menjalani perkuliahan, yang telah membantu, memberi semangat kepada penulis.
9. Teman-teman jurusan Matematika angkatan 2017 serta keluarga HIMATIKA.
10. Semua pihak yang terlibat dalam penyelesaian skripsi ini yang tidak dapat penulis sebutkan satu per satu.

Penulis menyadari bahwa masih ada kekurangan dalam skripsi ini, akan tetapi besar harapan penulis bahwa skripsi ini dapat bermanfaat bagi penulis maupun pembaca.

Bandar Lampung, Oktober 2021
Penulis

Rosita Sari

DAFTAR ISI

	Halaman
DAFTAR TABEL	iv
DAFTAR GAMBAR	v
I. PENDAHULUAN	
1.1 Latar Belakang	1
1.2 Tujuan Penelitian	2
1.3 Manfaat Penelitian	3
II. TINJAUAN PUSTAKA	
2.1 Estimasi Parameter.....	4
2.2 Analisis Regresi	4
2.3 Analisis Regresi Data Panel.....	6
2.4 Estimasi Model Regresi Data Panel	8
2.4.1 <i>Common Effect Model</i> (CEM)	8
2.4.2 <i>Ordinary Least Square</i> (OLS)	9
2.4.3 <i>Fixed Effect Model</i> (FEM)	10
2.4.4 <i>Least Square Dummy Variable</i> (LSDV)	11
2.4.5 <i>Random Effect Model</i> (REM)	13
2.4.6 <i>Generalized Least Square</i> (GLS)	14
2.5 Pemilihan Model Analisis Regresi Data Panel	15
2.5.1 Uji Chow	15
2.5.2 Uji Hausman	16
2.5.3 Uji Lagrange <i>Multiplier</i>	17
2.6 Uji Asumsi Klasik.....	18
2.6.1 Uji Normalitas	18
2.6.2 Uji Heteroskedastisitas	19
2.6.3 Uji Autokorelasi	20
2.7 Uji Statistik Model.....	21
2.7.1 Uji F Statistik	21
2.7.2 Uji T Statistik	22
2.8 Uji Koefisien Determinasi	23
III. METODOLOGI PENELITIAN	
3.1 Waktu dan Tempat.....	25
3.2 Data Penelitian	25
3.3 Metode Penelitian	25

IV. HASIL DAN PEMBAHASAN

4.1	Estimasi Model Regresi Data Panel.....	27
4.1.1	<i>Common Effect Model</i> (CEM)	27
4.1.2	<i>Fixed Effect Model</i> (FEM)	28
4.1.3	<i>Random Effect Model</i> (REM)	29
4.2	Pemilihan Model Estimasi Regresi Data Panel	29
4.2.1	Uji Chow	30
4.2.2	Uji Hausman.....	30
4.3	Uji Asumsi Model Regresi Data Panel	31
4.3.1	Uji Normalitas	32
4.3.2	Uji Heteroskedastisitas	33
4.3.3	Uji Autokorelasi	33
4.4	Pengujian Parameter Model Regresi	34
4.4.1	Uji F	35
4.4.2	Uji T	35
4.5	Uji Koefisien Determinasi	36
4.6	Model Akhir Regresi Data Panel	37

V. KESIMPULAN

DAFTAR PUSTAKA

LAMPIRAN

DAFTAR TABEL

Tabel	Halaman
1. Kriteria Pengujian Autokorelasi	21
2. Estimasi Model <i>Common Effect</i>	27
3. Estimasi Model <i>Fixed Effect</i>	28
4. Estimasi Model <i>Random Effect</i>	29
5. Hasil Uji Chow	30
6. Hasil Uji Hausman	31
7. Hasil Uji Heteroskedastisitas	33
8. Hasil Uji F.....	35
9. Hasil Uji T.....	36
10. Data Penelitian	43

DAFTAR GAMBAR

Gambar	Halaman
1. Hasil Uji Normalitas	32
2. Hasil Uji Autokorelasi	34

I. PENDAHULUAN

1.1 Latar Belakang

Analisis regresi adalah salah satu metode analisis yang digunakan untuk menentukan hubungan sebab akibat antara satu variabel dengan variabel lainnya. Istilah regresi pertama kali diperkenalkan pada tahun 1886 oleh Francis Gultom (Gujarati & Porter, 2009). Secara umum, analisis regresi digunakan untuk melihat ketergantungan antara satu variabel tak bebas pada satu atau lebih variabel bebas dengan maksud untuk menduga atau memprediksi nilai variabel tak bebas berdasarkan nilai variabel bebas yang diketahui.

Data panel merupakan data gabungan antara data runtun waktu (*time series*) dengan data silang (*cross section*). Data *time series* merupakan data yang terdiri dari satu objek namun meliputi beberapa periode waktu. Data ini memiliki nilai yang berubah-ubah seiring berjalannya waktu. Data *cross section* merupakan data yang terdiri dari beberapa objek namun hanya pada satu waktu/periode. Dengan kata lain, data panel merupakan data dari beberapa objek yang diamati dalam beberapa periode waktu.

Data panel memiliki kelebihan daripada hanya data *cross section* saja atau data *time series* saja. Data panel dapat memberikan peneliti dalam jumlah pengamatan yang besar, memiliki variabilitas yang besar dan mengurangi kolinearitas antara variabel penjelas, dimana dapat menghasilkan penduga yang efisien. Data panel dapat memberikan informasi lebih banyak yang tidak dapat diberikan oleh data

cross section atau *time series* saja, dan data panel dapat memberikan penyelesaian yang lebih baik dalam inferensi perubahan dinamis dibandingkan data *cross section*.

Menurut Jaya & Sunengsih (2009), analisis regresi data panel merupakan teknik regresi dengan menggunakan struktur data panel yaitu menggabungkan data *time series* dengan *cross section*. Dalam menentukan model yang sesuai dalam regresi data panel digunakan metode pendekatan. Salah satu metode pendekatan dalam menduga model regresi data panel yaitu model *random effect*. Estimasi model *random effect* menggunakan metode *Generalized Least Square* (GLS). Metode GLS dapat menghasilkan sifat-sifat penaksir yang baik diantaranya adalah tak bias, konsisten dan varians minimum. Metode ini juga memiliki kemampuan untuk menetralsir akibat pelanggaran asumsi homoskedastisitas (heteroskedastisitas) yang terjadi pada model dengan menggunakan metode *Ordinary Least Square* (OLS) sehingga akan didapatkan model penduga yang baik untuk berhadapan dengan gejala heteroskedastisitas.

Beberapa penelitian yang telah dilakukan menggunakan data panel antara lain penelitian yang dilakukan oleh Andini (2017) tentang Faktor-faktor yang Mempengaruhi Pertumbuhan Ekonomi di Provinsi Jawa Timur dengan Regresi Data Panel, dan Junia Rahma Nur Imani (2020) tentang Estimasi Model *Fixed Effect* pada Analisis Regresi Data Panel dengan Metode *Least Square Dummy Variabel* (LSDV). Oleh karena itu, dalam penelitian ini akan membahas pemodelan *random effect* menggunakan metode *Generalized Least Square* (GLS).

1.2 Tujuan Penelitian

Tujuan yang ingin dicapai dari penulisan penelitian ini adalah untuk menerapkan analisis regresi data panel model *random effect* dengan metode *Generalized Least Square* pada data riil.

1.3 Manfaat Penelitian

Manfaat yang diperoleh dari penelitian ini adalah:

1. Menambah pengetahuan dan wawasan bagi penulis dan pembaca mengenai analisis regresi data panel.
2. Salah satu bahan referensi dalam hal pengolahan data dengan menggunakan data panel.
3. Dapat mengetahui bentuk pemodelan model regresi data panel *random effect* dengan metode *Generalized Least Square*.

II. TINJAUAN PUSTAKA

2.1 Estimasi Parameter

Menurut Turmudi & Sri Hartini (2008), Parameter adalah hasil pengukuran yang menggambarkan karakteristik dari populasi. Sedangkan menurut Hasan (2002), estimasi (pendugaan) adalah proses yang menggunakan sampel statistik untuk menduga atau menaksir hubungan parameter populasi yang tidak diketahui. Pendugaan merupakan suatu pernyataan mengenai parameter populasi yang diketahui berdasarkan populasi sampel, dalam hal ini sampel random, yang diambil dari populasi yang bersangkutan. Sehingga dengan pendugaan, keadaan parameter populasi dapat diketahui. Secara umum parameter regresi diberi lambang β dan penduga diberi lambang $\hat{\beta}$.

2.2 Analisis Regresi

Analisis regresi adalah teknik analisis yang mencoba menjelaskan bentuk hubungan antara peubah-peubah yang mendukung sebab akibat. Prosedur analisisnya didasarkan atas distribusi probabilitas bersama peubah-peubahnya. Secara umum, dapat dikatakan, bahwa analisis regresi berkenaan dengan studi ketergantungan suatu variabel, yaitu variabel tak bebas (*dependent variable*), pada

satu atau lebih variabel bebas (*independent variable*), dengan maksud menduga dan atau meramalkan nilai rata-rata hitung (*mean*) dari variabel tak bebas (Firdaus, 2004).

Analisis regresi terdiri dari dua macam bentuk fungsi regresi yaitu regresi linear dan regresi nonlinear. Analisis regresi linear terdiri dari analisis regresi linear sederhana dan analisis regresi linear berganda. Analisis regresi linear sederhana didasarkan pada hubungan antara satu variabel tak bebas dan satu variabel bebas. Persamaan analisis regresi linear sederhana adalah sebagai berikut :

$$Y = \beta_0 + \beta_1 X_1 + \varepsilon \quad (2.1)$$

Y merupakan variabel tak bebas, X merupakan variabel bebas, dan ε merupakan variabel gangguan (*error*). Sedangkan β_0 dan β_1 adalah intersep dan koefisien regresi. (Gujarati & Porter, 2009).

Analisis regresi linear berganda adalah suatu analisis dalam statistika yang digunakan untuk mengetahui hubungan secara linear antara dua atau lebih variabel bebas (X_1, X_2, \dots, X_k) dengan variabel tak bebas (Y). Analisis ini bertujuan untuk mengetahui arah hubungan antara variabel bebas dan variabel tak bebas apakah masing-masing variabel bebas berhubungan positif atau negatif dan untuk memprediksi nilai dari variabel tak bebas apabila nilai variabel mengalami kenaikan atau penurunan (Draper & Smith, 1992). Persamaan regresi linear berganda adalah sebagai berikut:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + \varepsilon_i \quad (2.2)$$

dengan,

Y_i = nilai variabel tak bebas dalam observasi ke- i

β_0 = intersep

$\beta_1, \beta_2, \dots, \beta_k$ = koefisien kemiringan (*slope*)

X_{ik} = nilai variabel bebas

ε_i = variabel gangguan (*error*)

i = 1,2,3,...,N

2.3 Analisis Regresi Data Panel

Data panel merupakan data yang berasal dari penggabungan antara data *cross section* dan data *time series*. Pada data panel, unit *cross section* yang sama diukur selama beberapa periode waktu. Jadi, dapat disimpulkan bahwa data panel memiliki dimensi ruang dan waktu. Data panel adalah data yang merupakan hasil pengamatan pada beberapa individu atau unit yang masing-masing diamati dalam beberapa periode waktu yang berurutan (Baltagi, 2005).

Data panel dibedakan menjadi dua, *balance panel* dan *unbalanced panel*. *Balance panel* terjadi jika panjangnya waktu untuk setiap unit *cross section* sama. Sedangkan *unbalanced panel* terjadi jika panjangnya waktu tidak sama untuk setiap unit *cross section* (Gujarati & Porter, 2009).

Penggunaan data panel dalam penelitian memberikan beberapa keuntungan. Pertama, karena data panel adalah data gabungan antar data *cross section* dan data *time series* memungkinkan jumlah data meningkat dan memberikan sejumlah besar titik data yang dapat meningkatkan derajat bebas sehingga mengurangi terjadinya multikolinearitas antar variabel-variabel bebas yang terdapat dalam penelitian. Kedua, data lebih bervariasi dan dapat mengurangi masalah yang muncul apabila ada variabel yang dihilangkan. Ketiga, data panel juga dapat mengontrol heterogenitas individu (Hsiao, 2003). Hasil estimasi yang lebih baik akan diperoleh didalam analisis menggunakan data panel, yang mana hal ini seiring dengan peningkatan jumlah observasi.

Menurut Jaya & Sunengsih (2009) analisis regresi data panel adalah regresi yang didasarkan pada data panel untuk mengamati hubungan antara satu variabel tak bebas dengan satu atau lebih variabel bebas. Beberapa alternatif model yang dapat diselesaikan dengan data panel yaitu:

Model 1: semua koefisien baik *intercept* maupun *slope* koefisien konstan.

$$Y_{it} = \beta_0 + \sum_{K=1}^K \beta_k X_{kit} + \varepsilon_{it} \quad (2.3)$$

Model 2: *slope* koefisien konstan, tetapi *intercept* berbeda akibat perbedaan unit.

$$Y_{it} = \beta_{0i} + \sum_{K=1}^K \beta_k X_{kit} + \varepsilon_{it} \quad (2.4)$$

Model 3: *slope* koefisien konstan, tetapi *intercept* berbeda akibat perbedaan unit *cross section* dan berubahnya waktu.

$$Y_{it} = \beta_{0it} + \sum_{K=1}^K \beta_k X_{kit} + \varepsilon_{it} \quad (2.5)$$

Model 4: *intercept* dan *slope* koefisien berbeda akibat perbedaan unit *cross section*.

$$Y_{it} = \beta_{0i} + \sum_{K=1}^K \beta_{ki} X_{kit} + \varepsilon_{it} \quad (2.6)$$

Model 5: *intercept* dan *slope* koefisien berbeda akibat perbedaan unit *cross section* dan berubahnya waktu.

$$Y_{it} = \beta_{0it} + \sum_{K=1}^K \beta_{kit} X_{kit} + \varepsilon_{it} \quad (2.7)$$

dengan,

$$i = 1, 2, \dots, N$$

$$t = 1, 2, \dots, T$$

$$N = \text{banyak unit } \textit{cross section}$$

$$T = \text{banyak unit } \textit{time series}$$

$$Y_{it} = \text{Nilai variabel terikat } \textit{cross section} \text{ ke-}i \text{ dan } \textit{time series} \text{ ke-}t$$

$$X_{kit} = \text{nilai variabel bebas ke-}k \text{ untuk unit } \textit{cross section} \text{ ke-}i \text{ dan } \textit{time series} \text{ ke-}t$$

$$\beta_{kit} = \text{parameter yang ditaksir atau diduga}$$

$$\varepsilon_{it} = \text{nilai variabel gangguan untuk unit } \textit{cross section} \text{ ke-}i \text{ dan } \textit{time series} \text{ ke-}t$$

$$K = \text{banyaknya parameter parameter dalam regresi yang akan diduga}$$

Menurut Baltagi (2005), secara umum persamaan model regresi data panel dapat dituliskan sebagai berikut:

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \varepsilon_{it} \quad (2.8)$$

dengan,

$$Y_{it} = \text{variabel tak bebas untuk unit individu ke-}i \text{ dan waktu ke-}t$$

- X_{it} = variabel bebas untuk unit indivisu ke- i dan waktu ke- t
 β_0 = *intersep*
 β_1 = koefisien kemiringan (*slope*) untuk semua unit
 ε_{it} = *error* untuk individu ke- i dan waktu ke- j

2.4 Estimasi Model Regresi Data Panel

Berdasarkan asumsi pengaruh yang digunakan dalam analisis regresi data panel, model analisis regresi data panel dibagi menjadi tiga yaitu *common effect model*, *fixed effect model* dan *random effect model*.

2.4.1 Common Effect Model (CEM)

Common effect model merupakan pendekatan yang paling sederhana dengan mengabaikan dimensi *cross section* dan *time series*. Model ini mengasumsikan bahwa *intersep* dan koefisien *slope* konstan sepanjang waktu dan individu, dan *error term* menjelaskan perbedaan *intersept* dan koefisien *slope* sepanjang waktu dan individu tersebut. Regresi dilakukan dengan mengkombinasikan data *time series* dan *cross section*.

Model tanpa pengaruh individu atau CEM adalah pendugaan yang menggabungkan seluruh data *cross section* dan *time series*. Estimasi yang dilakukan yaitu dengan metode OLS. Metode OLS merupakan salah satu metode yang sering digunakan dalam teknik analisis regresi dengan meminimumkan kuadrat kesalahan atau *error* sehingga nilai regresinya akan mendekati nilai sesungguhnya (Andini, 2017). Dengan demikian, dalam model ini tidak ada efek individu.

Secara umum model dalam bentuk sistem persamaan adalah:

$$Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \dots + \beta_k X_{kit} + \varepsilon_{it} \quad (2.9)$$

dengan,

Y_{it} = variabel tak bebas untuk unit individu ke- i dan waktu ke- t

X_{it} = variabel bebas ke- k untuk unit indivisu ke- i dan waktu ke- t

β_0 = *intercept* model regresi

β_2, \dots, β_k = koefisien *slope*

ε_{it} = *error* untuk individu ke- i dan waktu ke- j .

Model *common effect* adalah model yang mengasumsikan bahwa β_0 dan β_1 akan sama untuk setiap data *time series* dan *cross section* mengakibatkan model ini memiliki kelemahan. Kelemahan tersebut yaitu ketidaksesuaian model dengan keadaan yang sesungguhnya yang mana kondisi tiap objek saling berbeda, bahkan kondisi objek pada suatu waktu dapat saja sangat berbeda dengan kondisi objek tersebut pada waktu yang lain.

2.4.2. Ordinary Least Square (OLS)

Menurut Nachrowi & Usman (2006), data panel akan memberikan observasi yang lebih banyak dibanding data *cross section* atau *time series* saja. Akibatnya, ketika data digabungkan menjadi *pooled data*, guna membuat regresi maka hasilnya cenderung akan lebih baik dibanding regresi yang hanya menggunakan data *cross section* atau *time series* saja. Model regresi data panel dengan OLS secara umum sebagai berikut:

$$Y_{it} = \alpha + \beta X_{it} + \varepsilon_{it} \quad (2.10)$$

Bila $cov(\varepsilon_{it}, \varepsilon_{jt}) = 0$; $cov(\varepsilon_{it}, \varepsilon_{it-1}) = 0$; $E(\varepsilon_{it}) = 0$; dan $var(\varepsilon_{it}) = \sigma^2$, kita dapat menduga model tersebut dengan memisahkan waktunya sehingga ada T regresi dengan N pengamatan. Atau dapat dituliskan dengan:

$$Y_{i1} = \alpha + \beta X_{i1} + \varepsilon_{i1};$$

$$Y_{i2} = \alpha + \beta X_{i2} + \varepsilon_{i2}$$

$$\begin{aligned} & \vdots \\ Y_{iT} &= \alpha + \beta X_{iT} + \varepsilon_{iT} \end{aligned} \quad (2.11)$$

Model juga dapat diduga dengan memisahkan *cross section*-nya sehingga didapat N regresi dengan masing-masing T pengamatan. Atau dapat ditulis dengan:

$$\begin{aligned} i &= 1; Y_{1t} = \alpha + \beta X_{1t} + \varepsilon_{1t}; \\ i &= 2; Y_{2t} = \alpha + \beta X_{2t} + \varepsilon_{2t}; \\ & \vdots \\ i &= N; Y_{Nt} = \alpha + \beta X_{Nt} + \varepsilon_{Nt}; \end{aligned} \quad (2.12)$$

Apabila asumsi bahwa α dan β akan sama (konstan) untuk setiap data *time series* dan *cross section*, maka α dan β dapat diduga dengan model berikut dengan menggunakan $N \times T$ pengamatan.

$$Y_{it} = \alpha + \beta X_{it} + \varepsilon_{it}; \quad (2.13)$$

2.4.3 Fixed Effect Model (FEM)

Kelemahan asumsi pada metode dalam model *common effect* yang tidak sesuai dengan keadaan yang sesungguhnya memerlukan suatu metode lain yang lebih sesuai, yang mana metode tersebut dapat menunjukkan adanya perbedaan dalam model. Salah satu cara paling sederhana untuk mengetahui adanya perbedaan adalah dengan mengasumsikan bahwa *intersep* berbeda antar unit *cross section*, sedangkan kemiringan (*slope*) nya tetap sama antar unit *cross section* (Gujarati & Porter, 2009). Model ini dikenal dengan model *fixed effect*, suatu model regresi yang bertujuan untuk menduga data panel. Persamaan untuk model *fixed effect* ini adalah (Widarjono, 2013) :

$$Y_{it} = \beta_{0i} + \beta_1 X_{1it} + \beta_2 X_{2it} + \dots + \beta_k X_{kit} + \varepsilon_{it} \quad (2.14).$$

Subskrip i pada *intersep* dalam persamaan diatas menjelaskan bahwa *intersept* untuk setiap unit *cross section* dalam analisis memiliki perbedaan. Metode yang digunakan dalam model *fixed effect* adalah suatu metode yang dipakai guna mengestimasi parameter regresi linear dengan menggunakan OLS pada model

yang melibatkan variabel *dummy* sebagai salah satu variabel prediktornya. Metode pendekatan ini disebut dengan metode *Least Square Dummy Variabel* (LSDV). Variabel *dummy* adalah sebuah variabel yang hanya memiliki dua kemungkinan (Andini, 2017). Analisis pada variabel *dummy* ini yaitu dengan memberi kode 1 pada salah satu kategori dan lainnya diberi kode 0. Penyertaan variabel *dummy* dalam analisis regresi data panel untuk model *fixed effect* diharapkan dapat mewakili menangkap adanya perbedaan dalam *intersept* dan mewakili ketidaklengkapan informasi dalam pembuatan model (Jaya & Sunengsih, 2009). Model *fixed effect* dengan variabel *dummy* dapat dituliskan sebagai berikut:

$$Y_{it} = \beta_0 + \sum_{i=2}^N \beta_{0i} D_{ki} + \sum_{k=1}^K \beta_k X_{kit} + \varepsilon_{it} \quad (2.15)$$

dengan,

Y_{it} = variabel tak bebas untuk unit *cross section* ke- i dan *time series* ke- t

X_{kit} = variabel bebas ke- k untuk unit *cross section* ke- i dan *time series* ke- t

D_{ki} = variabel *dummy* ke- k untuk unit *cross section* ke- i

ε_{it} = variabel gangguan (*error*) untuk unit *cross section* ke- i dan *time series* ke- t

β_k = parameter untuk variabel ke- k .

Subskrip $0i$ pada konstanta β_{0i} menunjukkan bahwa i adalah objeknya. Dengan demikian dapat diartikan bahwa objek memiliki konstanta yang berbeda. Variabel *dummy* D_{1i} berarti bahwa 1 untuk objek pertama dan 0 untuk objek lainnya.

Variabel *dummy* yang dibentuk berjumlah $N-1$, dengan β_0 sebagai *intersept* untuk unit *cross section* yang pertama (Singh & Sachdeva, 2020).

2.4.4 Least Square Dummy Variabel (LSDV)

Menurut Greene (2007), pendugaan parameter model efek tetap dilakukan dengan LSDV, dimana LSDV merupakan suatu metode pendugaan parameter linear dengan menggunakan Metode OLS pada model yang melibatkan variabel boneka.

OLS merupakan teknik pengepasan garis lurus terbaik untuk menghubungkan variabel tak bebas (Y) dengan variabel bebas (X). Berikut adalah prinsip dasar OLS:

$$u = Y - X\beta \quad (2.16)$$

Pada pemodelan efek tetap grup, variabel boneka yang dibentuk adalah sebanyak N-1, sehingga model yang akan diduga dalam pemodelan efek tetap adalah sebagai berikut:

$$Y_{it} = \alpha + \alpha_1 D_{1it} + \dots + \alpha_N D_{(N-1)it} \beta X_{it} + u_{it} \quad (2.17)$$

Sedangkan untuk pemodelan efek tetap waktu, variabel boneka yang dibentuk berdasarkan unit waktu, dimana variabel boneka yang terbentuk sebanyak T-1, sehingga model akan diduga dalam pemodelan efek tetap waktu adalah sebagai berikut:

$$Y_{it} = \alpha + \alpha_1 D_{1it} + \dots + \alpha_T D_{(T-1)it} \beta X_{it} + u_{it} \quad (2.18).$$

Model regresi panel dengan *intecept* bervariasi dan *slope* konstan, pemodelan efek tetap tanpa komponen dua arah secara umum dilakukan dengan LSDV dimana model dengan peubah *dummy* sebagai berikut:

$$Y_{it} = \alpha + \gamma_1 D_{1it} + \dots + \gamma_N D_{(N-1)it} + \delta_1 D_{1it} + \dots + \delta_T D_{(T-1)it} + \beta X_{it} + u_{it} \quad (2.19)$$

dengan,

D_{jit} = peubah boneka ke-j unit *cross section* ke-i dan unit waktu ke-t. D_{jit} bernilai satu jika $j = i$ dan bernilai nol jika $j \neq i$.

D_{kit} = peubah boneka ke-k unit *cross section* ke-i dan unit waktu ke-t.

D_{kit} bernilai satu jika $k = i$ dan bernilai nol jika $k \neq i$.

α_j = rata-rata peubah respon jika peubah boneka ke-j bernilai satu dan peubah penjelas bernilai nol

α_k = rata-rata nilai peubah respon jika peubah boneka ke-k bernilai satu dan peubah penjelas bernilai nol.

2.4.5 *Random Effect Model (REM)*

Pendugaan data panel melalui pendekatan FEM, teknik penggunaan variabel *dummy* menunjukkan ketidaktahuan tentang model sebenarnya. Sebagaimana telah diketahui bahwa pada model *fixed effect*, perbedaan karakteristik untuk data *cross section* dan *time series* diakomodasikan lewat *intersept*. Sementara pada model *random effect* karakteristik diakomodasikan lewat galat dari model yang terbentuk. Model *random effect* bertujuan untuk mengestimasi data panel menggunakan variable gangguan (*error*) yang diduga memiliki hubungan antar waktu dan antar individu (Winarno, 2007). Mengingat ada dua komponen yang berkontribusi pada pembentukan galat, yaitu individu dan waktu, maka galat pada *random effect* juga diuraikan menjadi galat untuk komponen waktu dan komponen gabungan (Nachrowi & Usman, 2006). Adapun metode estimasi yang tepat digunakan untuk mengestimasi model data panel *random effect* adalah dengan menggunakan *GLS*. Persamaan model *random Effect* dapat dituliskan sebagai berikut:

$$Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \dots + \beta_k X_{kit} + \varepsilon_{it} \quad (2.20)$$

Variabel gangguan (*error*) dalam model *random effect* mempunyai beberapa karakteristik, antar lain:

$$E(\mu) = 0 \quad (2.21)$$

$$Var(\mu) = \sigma_\mu^2 \quad (2.22)$$

Sehingga asumsi yang digunakan untuk komponen galat dapat dituliskan sebagai berikut:

$$u_i \sim N(0, \sigma_u^2) \quad (2.23)$$

$$v_t \sim N(0, \sigma_v^2) \quad (2.24)$$

$$w_{it} \sim N(0, \sigma_w^2) \quad (2.25)$$

2.4.6 Generalized Least Square (GLS)

Menurut Pangestika (2015), Pendugaan parameter untuk *random effect model* dilakukan dengan menggunakan metode *GLS* jika matriks Ω diketahui, namun jika tidak diketahui maka dilakukan dengan *GLS* yaitu dengan menduga elemen matriks Ω . Pada *random effect model* ketidaklengkapan informasi untuk setiap unit *cross section* dipandang sebagai *error* sehingga μ_i adalah bagian dari unsur gangguan. Model dapat dituliskan sebagai berikut:

$$Y_{it} = \alpha + \sum_{k=1}^K \beta_k X_{kit} + (\mu_i + e_{it}) \quad (2.26)$$

Asumsi:

$$\mu_i \sim N(0, \sigma_\mu^2) E(\mu_i, \mu_j) = 0; i \neq j \quad (2.27)$$

$$e_{it} \sim N(0, \sigma_e^2) E(\mu_i, e_{it}) = 0 \quad (2.28)$$

$$E(e_{it}, e_{is}) = E(e_{it}, e_{jt}) = E(e_{it}, e_{js}) = 0; i \neq j; j \neq s \quad (2.29)$$

Untuk data *cross section* ke- i persamaan dapat ditulis sebagai berikut:

$$Y_i = X_i \beta + (\mu_i 1 + e_i) \quad (2.30)$$

Varian komponen dari unsur $(\mu_i 1 + e_i)$ untuk unit *cross section* ke- i adalah:

$$\Omega = \begin{bmatrix} \sigma_\mu^2 + \sigma_e^2 & \sigma_\mu^2 & \dots & \sigma_\mu^2 \\ \sigma_\mu^2 & \sigma_\mu^2 + \sigma_e^2 & \dots & \sigma_\mu^2 \\ \dots & \dots & \ddots & \dots \\ \sigma_\mu^2 & \sigma_\mu^2 & \dots & \sigma_\mu^2 + \sigma_e^2 \end{bmatrix} \quad (2.31)$$

Varian komponen Ω identik untuk setiap unit *cross section*, sehingga varians komponen untuk seluruh observasi dapat dituliskan sebagai berikut:

$$W = \begin{bmatrix} \Omega & 0 & \dots & 0 \\ 0 & \Omega & \dots & 0 \\ \dots & \dots & \ddots & \dots \\ 0 & 0 & \dots & \Omega \end{bmatrix} \quad (2.32)$$

Jika nilai Ω diketahui maka persamaan dapat diduga menggunakan *GLS* dengan

$$\hat{\beta} = (X'W^{-1}X)^{-1}(X'W^{-1}y). \text{ Jika } \Omega \text{ tidak diketahui maka } \Omega \text{ perlu diduga}$$

dengan menduga $\hat{\sigma}_\mu^2$ dan $\hat{\sigma}_e^2$, sehingga persamaan dapat diduga dengan

$$\hat{\beta} = (X'\hat{W}^{-1}X)^{-1}(X'\hat{W}^{-1}y) \text{ dimana } \hat{\sigma}_e^2 = \frac{e'e}{NT-N-K} \text{ dengan } \hat{e} = y - X\hat{\beta} \text{ adalah}$$

residu dari *LSDV*, sedangkan $\hat{\sigma}_\mu^2 = \frac{\hat{\sigma}_1^2 - \hat{\sigma}_e^2}{T}$.

2.5 Pemilihan Model Analisis Regresi Data Panel

Pemilihan model estimasi analisis regresi data panel merupakan tahapan analisis untuk menentukan model estimasi terbaik dalam analisis menggunakan data panel antara model *common effect*, model *fixed effect* dan model *random effect*.

Pemilihan model terbaik dilakukan menggunakan uji Chow, uji Hausman dan uji Lagrange *Multiplier*.

2.5.1 Uji Chow

Uji Chow adalah salah satu pengujian yang dapat dilakukan untuk menentukan model terbaik pada analisis regresi data panel. Uji chow dilakukan untuk menentukan apakah model *common effect* lebih baik digunakan daripada model *fixed effect*. Adapun hipotesis yang digunakan adalah sebagai berikut:

$H_0 : \alpha_0 = \alpha_1 = \dots = \alpha_n = \alpha = 0$ (model yang sesuai adalah model *common effect*)

$H_1 : \text{minimal ada satu } \alpha_i \neq 0 ; i = 1, 2, \dots, n$ (model yang sesuai adalah model *fixed effect*)

Statistik uji yang digunakan adalah uji F, yaitu :

$$F_{hitung} = \frac{JKR/(k - 1)}{JKG/(n - k)} \quad (2.33)$$

dengan,

JKR = Jumlah Kuadrat Regresi

JKG = Jumlah Kuadrat Galat

k = Jumlah variabel

n = Jumlah pengamatan.

Dasar penolakan dari hipotesis di atas adalah dengan membandingkan perhitungan F-hitung dengan F-tabel. Perbandingan dipakai apabila hasil F-

hitung lebih besar dari F-tabel, maka H_0 ditolak yang berarti model yang lebih tepat digunakan adalah *fixed effect* model. Begitupun sebaliknya, jika F-hitung lebih kecil dari F-tabel, maka H_0 diterima yang berarti model yang lebih tepat digunakan adalah *common effect model*.

Adapun cara lain yang dapat digunakan untuk pengambilan kesimpulan adalah dengan melihat nilai probabilitasnya. Apabila nilai probabilitas *chi-square* kurang dari taraf signifikansi 5% (0,05) artinya model yang lebih digunakan yaitu *fixed effect model* dan jika nilai probabilitas *chi-square* lebih dari taraf signifikan 5% (0,05) maka model yang digunakan yaitu *common effect model* (Gujarati & Porter, 2009).

2.5.2 Uji Hausman

Apabila pada uji Chow telah diperoleh kesimpulan maka langkah selanjutnya yaitu melakukan uji Hausman. Uji Hausman adalah salah satu pengujian yang dilakukan untuk mengetahui model mana yang lebih tepat antara *fixed effect model* dan *random effect model*. Tujuan dari uji ini yaitu untuk menguji apakah terdapat hubungan antara galat pada model dengan satu atau lebih variabel bebas dalam model. Dalam uji ini hipotesisi yang digunakan yaitu sebagai berikut :

$H_0 : corr(X_{it}, \varepsilon_i) = 0$ (model yang sesuai adalah model *random effect*)

$H_1 : corr(X_{it}, \varepsilon_i) \neq 0$ (model yang sesuai adalah model *fixed effect*)

$$W = \hat{q}[\text{var}(\hat{q})^{-1}\hat{q}] \quad (2.34)$$

$$W = (\hat{\beta}_{FEM} - \hat{\beta}_{REM})[\text{var}(\hat{\beta}_{FEM} - \hat{\beta}_{REM})]^{-1}(\hat{\beta}_{FEM} - \hat{\beta}_{REM}) \quad (2.35)$$

dengan,

$\hat{\beta}_{FEM}$ = vektor estimasi *slope* model *fixed effect*

$\hat{\beta}_{REM}$ = vektor estimasi *slope* model *random effect*.

Jika nilai $W > X^2_{(\alpha,K)}$ kurang dari taraf signifikansi yang ditentukan, maka tolak H_0 , sehingga model yang lebih tepat digunakan yaitu model *fixed effect*.

Begitupun sebaliknya apabila nilai $W > X^2_{(\alpha,K)}$ lebih dari taraf signifikansi maka terima H_0 , sehingga model yang lebih tepat digunakan yaitu model *random effect*.

Adapun cara lain yang dapat digunakan untuk pengambilan kesimpulan adalah dengan melihat nilai probabilitasnya. Apabila nilai probabilitas *chi-square* kurang dari taraf signifikansi 5% (0,05) artinya model yang lebih digunakan yaitu *fixed effect model* dan jika nilai probabilitas *chi-square* lebih dari taraf signifikan 5% (0,05) maka model yang digunakan yaitu *random effect model* (Gujarati & Porter, 2009).

2.5.3 Uji Lagrange Multiplier

Menurut Rosadi (2011), uji Lagrange *Multiplier* adalah salah satu pengujian yang dilakukan untuk mengetahui model mana yang lebih tepat antara *common effect* model dan *random effect* model. Tujuan dari uji ini yaitu untuk menguji apakah terdapat hubungan efek waktu, individu atau keduanya dalam model. Dalam uji ini hipotesisi yang digunakan yaitu sebagai berikut :

$H_0: \sigma^2_{\mu} = 0$ (model yang sesuai adalah *common effect model*)

$H_1: \sigma^2_{\mu} \neq 0$ (model yang sesuai adalah *random effect model*)

$$LM = \frac{NT}{2(T-1)} \left[\frac{\sum_{i=1}^N (\sum_{t=1}^T \hat{u}_{it})^2}{\sum_{i=1}^N \sum_{t=1}^T \hat{u}_{it}^2} - 1 \right] \quad (2.36)$$

dengan,

N = jumlah individu

T = jumlah periode waktu

σ^2_{μ} = varians dari residual model

\hat{u}_{it} = estimasi residual model koefisien tetap individu ke-*i* dan periode ke-*t*

Jika nilai $LM > X^2_{(\alpha,1)}$ atau nilai *p-value* lebih kecil dari taraf signifikansi 5% (0,05) maka tolak H_0 artinya model yang tepat untuk digunakan adalah model *random effect*. Begitupun sebaliknya, jika nilai $LM < X^2_{(\alpha,1)}$ atau nilai *p-value* lebih besar dari taraf signifikansi 5% (0,05) maka terima H_0 artinya model yang tepat untuk digunakan adalah model *common effect*.

2.6 Uji Asumsi Klasik

Model analisis regresi data panel dapat dikatakan sebagai model yang baik apabila data yang digunakan memenuhi beberapa uji asumsi klasik yang mencakup uji normalitas, uji heteroskedastisitas dan uji autokorelasi.

2.6.1 Uji Normalitas

Uji normalitas digunakan untuk melihat apakah nilai residual berdistribusi normal atau tidak. Model yang baik memiliki nilai residual yang berdistribusi normal (Draper & Smith, 1992). Tidak terpenuhinya normalitas umumnya disebabkan oleh data yang tidak normal karena terdapat nilai ekstrem pada data yang dianalisis. Nilai ekstrem ini dapat terjadi karena adanya kesalahan dalam pengambilan sampel, bahkan karena kesalahan melakukan input data atau memang karakteristik data yang sangat jauh dari nilai rata-rata.

Pengujian asumsi ini menguji normalitas pada residual yang dihasilkan dari model regresinya. Untuk menguji normalitas digunakan uji Jarque-Bera. Uji Jarque-Bera ini menggunakan perhitungan *skewness* dan kurtosis dengan hipotesis sebagai berikut :

H_0 : residual berdistribusi normal

H_1 : residual tidak berdistribusi normal

Statistik uji :

$$JB = N \left[\frac{S^2_k}{6} + \frac{(K - 3)^2}{24} \right] \quad (2.37)$$

$$K = \frac{\widehat{\mu}_4}{\widehat{\mu}_2^2} = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^4}{\left(\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 \right)^2} \quad (2.38)$$

$$S = \frac{\widehat{\mu}_3}{\widehat{\mu}_2^{3/2}} = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^3}{\left(\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 \right)^{3/2}} \quad (2.39)$$

dengan,

N = banyaknya data

S = *skewness*

K = kurtosis.

Statistik uji JB ini mengikuti distribusi *chi-square* dengan derajat bebas 2 (db-2). Kriteria uji JB adalah pada taraf signifikansi 5%, jika nilai JB-hitung lebih kecil dari nilai X^2 maka H_0 diterima artinya residual berdistribusi normal. Begitupun sebaliknya, jika nilai JB-hitung lebih besar dari nilai X^2 maka H_0 ditolak artinya residual tidak berdistribusi normal (Widarjono, 2007). Cara lain yang dapat dilakukan yaitu dengan melihat nilai probabilitasnya. Apabila nilai probabilitas lebih besar dari pada taraf signifikansi 5% maka residual berdistribusi normal. Begitupun sebaliknya, apabila nilai probabilitas lebih kecil dari pada taraf signifikansi 5% maka residual tidak berdistribusi normal.

2.6.2 Uji Heteroskedastisitas

Salah satu asumsi penting adalah varian dari residual (galat) konstan. Namun dalam kenyataannya sering kali varian dari residual (galat) adalah tidak konstan atau disebut dengan heteroskedastisitas. Heteroskedastisitas dapat mengakibatkan pendugaan parameternya tidak efisien sehingga tidak mempunyai ragam

minimum. Pendugaan parameter dianggap efisien karena memiliki ragam yang minimum, sehingga ragam galat bersifat konstan atau disebut juga bahwa asumsi *non-heteroskedastisitas* terpenuhi. Salah satu usaha untuk mengatasi heteroskedastisitas ini dapat dilakukan dengan mentransformasikan variabel-variabelnya, baik variabel bebas, variabel tak bebas maupun keduanya agar asumsi *non-heteroskedastisitas* terpenuhi.

Deteksi heteroskedastisitas dapat menggunakan uji Glejser. Uji Glejser dilakukan dengan cara meregresikan variabel bebas dengan residual. Jika hasil uji glejser signifikan, maka telah terjadi heteroskedastisitas. Sedangkan jika hasil uji Glejser tidak signifikan, maka tidak terjadi heteroskedastisitas (Hasanah, 2008). Adapun hipotesis dalam uji heteroskedastisitas adalah sebagai berikut :

H_0 : residual tidak mengalami masalah heteroskedastisitas

H_1 : residual mengalami masalah heteroskedastisitas

Dasar pengambilan keputusan uji heteroskedastisitas ini adalah dengan melihat nilai probabilitasnya. Apabila nilai probabilitas lebih besar daripada taraf signifikansi 5% maka terima H_0 , sehingga dapat disimpulkan bahwa data tidak mengalami masalah heteroskedastisitas.

2.6.3. Uji Autokorelasi

Autokorelasi adalah terjadinya korelasi antara satu variabel *error* dengan variabel *error* lainnya. Uji autokorelasi digunakan untuk mengetahui ada atau tidaknya korelasi antara residual pada satu pengamatan dengan pengamatan lainnya pada model regresi (Algifari, 2000). Untuk melihat ada atau tidaknya autokorelasi digunakan nilai Durbin Watson Stat dan membandingkannya dengan nilai pada tabel Durbin Watson. adapun hipotesis yang dilakukan pada uji ini adalah sebagai berikut :

$H_0: \rho = 0$ (data tidak mengalami autokorelasi)

$H_0: \rho \neq 0$ (data mengalami autokorelasi).

Adapun statistik uji yang digunakan adalah :

$$d = \frac{\sum_{i=1}^n (\varepsilon_i - \varepsilon_{i-1})^2}{\sum_{i=1}^n \varepsilon_i^2} \quad (2.40)$$

Dengan tingkat signifikansi (α) sebesar 5% dan dengan kriteria pengambilan keputusan sebagai berikut :

Tabel 1. Kriteria Pengujian Autokorelasi

Hipotesis Nol (H_0)	Keputusan	Kriteria
Ada autokorelasi positif	H_0 ditolak	$0 < d < d_L$
Tidak ada autokorelasi positif	Tidak ada keputusan	$d_L \leq d \leq d_U$
Ada autokorelasi negatif	H_0 ditolak	$4 - d_L \leq d \leq 4$
Tidak ada autokorelasi negatif	Tidak ada keputusan	$4 - d_U \leq d \leq 4 - d_L$
Tidak ada autokorelasi	H_0 diterima	$d_U < d < 4 - d_U$

2.7 Uji Statistik Model

Uji statistik model dilakukan dengan mengetahui hubungan antara variabel tak bebas dan variabel bebas. Terdapat dua pengujian yang harus dilakukan yaitu pengujian secara simultan dan pengujian secara parsial.

2.7.1 Uji F Statistik

Uji F statistik dilakukan untuk mengetahui apakah semua variabel bebas pada model berpengaruh secara signifikan terhadap variabel tak bebas, uji ini dilakukan secara simultan atau keseluruhan. Uji ini sering juga disebut dengan pengujian signifikansi keseluruhan (*Overall significance*) terhadap yang ingin menguji

apakah variabel tak bebas secara linear berhubungan dengan variabel-variabel bebas. Adapun hipotesisi yang digunakan pada pengujian ini adalah sebagai berikut :

$H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$ (variabel bebas secara simultan tidak berpengaruh signifikan terhadap variabel tak bebas)

$H_1 : \beta_1 = \beta_2 = \dots = \beta_k \neq 0$ (variabel bebas secara simultan berpengaruh signifikan terhadap variabel tak bebas).

Dengan statistik uji F yang terdapat pada persamaan (2.33) dengan taraf signifikansi 5%, apabila nilai F-hitung lebih besar daripada F-tabel maka tolak H_0 artinya variabel bebas secara simultan berpengaruh signifikan terhadap variabel tak bebas. Begitupun sebaliknya, apabila nilai F-hitung lebih kecil daripada F-tabel maka terima H_0 artinya variabel bebas secara simultan tidak berpengaruh signifikan terhadap variabel tak bebas. Dasar pengambilan keputusan pada uji F ini juga dapat ditentukan berdasarkan nilai probabilitasnya. Jika nilai probabilitas kurang dari taraf signifikansi 5% maka tolak H_0 , artinya variabel bebas secara simultan berpengaruh signifikan terhadap variabel tak bebas. Begitupun sebaliknya, jika nilai probabilitas lebih besar dari taraf signifikan 5% maka terima H_0 artinya variabel bebas secara simultan tidak berpengaruh signifikan terhadap variabel tak bebas.

2.7.2 Uji T Statistik

Uji t statistik termasuk kelompok uji parametrik, yaitu kelompok uji statistika yang memerlukan persyaratan tertentu agar memberikan hasil yang baik dalam asumsi distribusi data. Uji t statistik dilakukan secara parsial. Uji ini dilakukan untuk mengetahui variabel bebas yang berpengaruh signifikan secara individu terhadap variabel tak bebas. Hipotesis yang digunakan pada uji parsial ini adalah sebagai berikut :

$H_0 : \beta_i = 0 ; i = 1,2,\dots,n$ (suatu variabel bebas secara individu tidak berpengaruh signifikan terhadap variabel tak bebas)

$H_1 : \beta_i \neq 0 ; i = 1,2,\dots,n$ (suatu variabel bebas secara individu berpengaruh signifikan terhadap variabel tak bebas).

Statistik uji t adalah sebagai berikut :

$$t_{hitung} = \frac{b_i}{Sb_i} \quad (2.41)$$

dengan,

b_i = koefisien regresi (*slope*) variabel bebas ke- i ; $i = 1,2,\dots,n$

Sb_i = *standard error* penduga b_i

Dengan taraf signifikan 5% jika nilai T-hitung lebih besar daripada nilai T-tabel maka tolak H_0 artinya suatu variabel bebas secara individu berpengaruh signifikan terhadap variabel tak bebas. Sebaliknya, jika nilai T-hitung lebih kecil daripada nilai T-tabel maka terima H_0 artinya suatu variabel bebas secara individu tidak berpengaruh signifikan terhadap variabel tak bebas. Dasar pengambilan keputusan pada uji ini juga dapat ditentukan berdasarkan nilai probabilitas. Jika nilai probabilitas kurang dari taraf signifikan 5% maka tolak H_0 artinya suatu variabel secara individu berpengaruh signifikan terhadap variabel tak bebas. Sebaliknya jika nilai probabilitas lebih besar dari taraf signifikan 5% maka terima H_0 artinya variabel bebas secara individu tidak berpengaruh signifikan terhadap variabel tak bebas.

2.8 Uji Koefisien Determinasi

Koefisien determinasi pada intinya digunakan untuk mengukur seberapa jauh kemampuan model dalam menjelaskan varian dari variabel tak bebas. Nilai R^2 yang kecil artinya kemampuan variabel-variabel bebas dalam menjelaskan

variabel tak bebas amat terbatas. Nilai R^2 yang mendekati satu artinya variabel bebas memberikan hampir semua informasi yang dibutuhkan untuk memprediksi variabel tak bebas.

Dalam uji ini koefisien determinasi sering digunakan nilai dari *adjusted R-square*. Hal ini dikarenakan adanya kelemahan dalam perhitungan R^2 yaitu nilai R^2 sering terjadi bias yang disebabkan meningkatnya nilai R^2 . Meningkatnya nilai R^2 disebabkan oleh penambahan variabel bebas yang tidak bergantung apakah variabel bebas tambahan tersebut berhubungan dengan variabel tak bebas atau tidak. Berdasarkan hal tersebut mengakibatkan banyaknya peneliti yang menyarankan untuk menggunakan nilai *adjusted R-square* atau Adj- R^2 (Usman & Akbar, 2009). Pada interpretasinya sama dengan R^2 , akan tetapi nilai Adj- R^2 dapat naik dan turun dengan adanya penambahan variabel baru, tergantung dari korelasi variabel bebas tambahan tersebut dengan variabel tak bebasnya. Nilai Adj- R^2 dapat bernilai negatif sehingga jika nilainya negatif maka nilai tersebut dianggap 0, atau variabel bebas sama sekali tidak mampu menjelaskan varian dari variabel tak bebasnya.

Perlu diperhatikan bahwa nilai R^2 yang tinggi tidak selalu menunjukkan kualitas model yang sudah baik. Dalam analisis *time series*, yang biasanya semua variabel mengalami kenaikan seiring berjalannya waktu, maka nilai R^2 akan cenderung tinggi. Namun pada analisis *cross section* nilai R^2 akan cenderung rendah (Winarno, 2007).

III. METODOLOGI PENELITIAN

3.1 Waktu dan Tempat

Penelitian ini dilakukan pada semester genap tahun akademik 2020/2021 di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

3.2 Data Penelitian

Data yang digunakan dalam penelitian ini terdiri atas variabel Y = Indeks Pembangunan Manusia (IPM), X_1 = Angka Harapan Hidup (AHH), X_2 = Rata-rata Lama Sekolah (RLS) dan X_3 = Pengeluaran Perkapita Disesuaikan sebagai Variabel X_3 di Provinsi Lampung pada Tahun 2015-2018 yang di dapat dari buku Provinsi Lampung Dalam Angka tahun 2019 yang merupakan buku publikasi resmi dari Badan Pusat Statistik (BPS) Provinsi Lampung.

3.3 Metode Penelitian

Penelitian dilakukan dengan bantuan program *Eviews 10*. Adapun langkah-langkah yang dilakukan pada penelitian ini adalah sebagai berikut :

1. Membuat model regresi data panel berdasarkan pendekatan *Common Effect Model* dengan metode OLS, *Fixed Effect Model* dengan metode LSDV, dan

Random Effect Model dengan metode *GLS*.

2. Uji pemilihan model terbaik yang meliputi:
 - a. Uji Chow, pengujian antara model *common effect* dan model *fixed effect*.
 - b. Uji Hausman, pengujian antara model *fixed effect* dan model *random effect*.
 Apabila pada uji Hausman telah diperoleh model terbaik maka dilanjutkan dengan melakukan uji asumsi klasik. Dan apabila tidak diperoleh model terbaik maka dilanjutkan dengan uji lagrange *Multiplier*.
 - c. Uji lagrange *Multiplier*, pengujian antar model *common effect* dan model *random effect*.
3. Melakukan uji asumsi klasik yang meliputi:
 - a. Uji normalitas dengan menggunakan statistik uji Jarque Bera
 - b. Uji heteroskedastisitas dengan menggunakan uji Glejser.
 - c. Uji autokorelasi dengan statistik uji Durbin Watson.
4. Apabila semua asumsi dalam tahapan uji asumsi klasik terpenuhi:
 - a. Pengujian parameter dalam model regresi yang meliputi pengujian simultan dengan uji F dan pengujian parsial dengan uji T statistik.
 - b. Melihat kebaikan model dari nilai koefisien determinasi.
5. Apabila terdapat asumsi didalam tahapan uji asumsi klasik yang tidak terpenuhi:
 - a. Melakukan penanganan untuk mengatasi asumsi klasik yang tidak terpenuhi sampai semua uji asumsi klasik terpenuhi.
 - b. Pengujian parameter dalam model regresi yang meliputi pengujian simultan dan pengujian parsial.
 - c. Melihat kebaikan model berdasarkan nilai dari koefisien determinasi.

V. KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan, diperoleh kesimpulan antara lain sebagai berikut:

1. Hasil analisis model *random effect* terhadap data Indeks Pembangunan Manusia, Angka Harapan Hidup, Rata-rata Lama Sekolah dan Pengeluaran Perkapita Disesuaikan di Provinsi Lampung tahun 2015 hingga 2018 sebagai berikut:

$$\hat{Y}_{it} = 7,215355 + 0,512942X_{1it} + 1,573944X_{2it} + 0,001357X_{3it}$$

2. Variabel bebas dalam model yaitu Angka Harapan Hidup, Rata-rata Lama Sekolah dan Pendapatan Perkapita Disesuaikan memberikan pengaruh yang signifikan terhadap variabel tak bebas yaitu Indeks Pembangunan Manusia baik secara parsial maupun secara simultan pada taraf signifikan $\alpha = 5\%$.
3. Sebesar 97,7344% faktor Angka Harapan Hidup, Rata-rata Lama Sekolah, dan Pengeluaran Perkapita Disesuaikan bersama-sama mempengaruhi Indeks Pembangunan Manusia, sedangkan sisanya sebesar 2,2656% dijelaskan oleh faktor lain diluar penelitian yang tidak diikutsertakan ke dalam model.

DAFTAR PUSTAKA

- Algifari. 2000. *Analisis Regresi Teori dan Kasus*. BPEE, Yogyakarta.
- Andini, E.Y. 2017. Faktor-Faktor yang Mempengaruhi Pertumbuhan Ekonomi Sektor Pertanian di Daerah Tertinggal Provinsi Jawa Timur dengan regresi Panel. Tugas Akhir. Departemen Statistika Bisnis Fakultas Vokasi ITS, Surabaya.
- Badan Pusat Statistik Provinsi Lampung. 2019. *Provinsi Lampung Dalam Angka 2019*. Badan Pusat Statistik Provinsi Lampung, Lampung.
- Baltagi, B.H. 2005. *Econometrics Analysis of Panel data*. 3rd Edition. Pearson Education, Inc., New Jersey.
- Draper, N.R & Smith, H. 1992. *Analisis Regresi Terapan*. Terjemahan Bambang Sumantri. Gramedia Pustaka Utama, Jakarta.
- Firdaus, M. 2004. *Ekonometrika Suatu Pendekatan Aplikatif*. Bumi Aksara, Jakarta.
- Gujarati, D.N. & Porter, D.C. 2009. *Basic Econometrics*. 5th Edition. McGraw Hill Company, New York.
- Hasan, I. 2002. *Pokok-Pokok Materi Statistik 2 (Statistik Inferensif)*. Bumi Aksara, Jakarta.
- Hasanah, N.N. 2008. Pengujian Heteroskedastisitas pada Regresi Non Linear dengan Menggunakan Uji Glejset. Skripsi. Jurusan Matematika FST UIN Malang, Malang.

- Hsiao, C. 2003. *Analysis of Panel Data*. 2nd Edition. Cambridge University Press, Cambridge.
- Jaya, I. G. N. M. & Sunengsih, N. 2009. Kajian Analisis Regresi Data Panel. Prosiding Seminar Nasional Penelitian. Yogyakarta, Universitas Negeri Yogyakarta.
- Montgomery, D.C. & Peck, E.A. 1992. *Introduction to Linear Regression Analysis*. 2nd Edition. Willey Intersciences Publication, Canada.
- Nachrowi, D.N. & Usman, H. 2006. *Pendekatan Populer dan Praktis Ekonometrika untuk Analisis Ekonomi dan Keuangan*. Lembaga Penerbit FE UI, Jakarta.
- Pangestika, S. 2015. Analisis Estimasi Model Regresi Data Panel dengan Pendekatan Common Effect Model, Fixed Effect Model, dan random Effect Model. Skripsi. Jurusan Matematika FMIPA UNNES, Semarang.
- Rosadi, D. 2011. *Ekonometrika & Analisis Runtun waktu terapan dengan R*. C. V. Andi Offset, Yogyakarta.
- Singh, S. & Sachdeva, T.K. 2020. Financial Performance of Selected IT Companies in India: A Panel Data Approach. *International Journal of Arts, Science and Humanities*. 7(3): 7-14.
- Turmudi & Hartini, S. 2008. *Metode Statistika*. UIN-Maliki Press, Malang.
- Usman, H. & Akbar. 2009. *Pengantar Statistika*. Bumi Aksara, Jakarta.
- Widarjono, A. 2013. *Ekonometrika Pengantar dan Aplikasinya*. Ed. Ke-4. UPP STIM YKPN, Yogyakarta.
- Winarno, W.W. 2007. *Analisi Ekonometrika dan Statistika dengan Eviews*. UPP STIM YKPN, Yogyakarta.