

**PEMODELAN REGRESI DATA PANEL *RANDOM EFFECT* DENGAN  
METODE *GENERALIZED LEAST SQUARE* (GLS) UNTUK  
MENGANALISIS FAKTOR YANG MEMPENGARUHI EMISI KARBON  
DIOKSIDA (CO<sub>2</sub>)**

**(Skripsi)**

**Oleh**

**MERYAM JOVANKA SOLEH  
2017031070**



**JURUSAN MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMPUNG  
BANDAR LAMPUNG  
2024**

## **ABSTRACT**

### **RANDOM EFFECT PANEL DATA REGRESSION MODELING USING THE GENERALIZED LEAST SQUARE (GLS) METHOD TO ANALYZE FACTORS AFFECTING CARBON DIOXIDE (CO<sub>2</sub>) EMISSIONS**

**By**

**MERYAM JOVANKA SOLEH**

One of the panel data regression analysis models is the Random Effect Model (REM). The random effects model can estimate panel data using nuisance variables that allow them to be correlated over time and between individuals. The appropriate method used in REM is the Generalized Least Square (GLS) method. The GLS method will create an estimator that meets the characteristics of Best Linear Unbiased Estimation (BLUE). The REM model estimates were carried out using the GLS method and applied to CO<sub>2</sub> emission data in 5 ASEAN member countries for 2000-2021. The aim of this research is to determine a random effect panel data regression model using the GLS method and determine the variables that influence CO<sub>2</sub> emissions in the 5 ASEAN member countries. The results of this research obtained a random effect model with the variables economic growth and population having a significant influence on CO<sub>2</sub> emissions while the energy consumption variable did not have a significant influence on CO<sub>2</sub> emissions. With the equation model the estimation results are  $Y_{it} = -62,851 - 0,00109X_{1it} + 0,0009X_{2it} + 5,633X_{3it}$ .

Keywords: panel data, random effect model, GLS

## ABSTRAK

### PEMODELAN REGRESI DATA PANEL *RANDOM EFFECT* DENGAN METODE *GENERALIZED LEAST SQUARE* (GLS) UNTUK MENGANALISIS FAKTOR YANG MEMPENGARUHI EMISI KARBON DIOKSIDA (CO<sub>2</sub>)

Oleh

MERYAM JOVANKA SOLEH

Salah satu model analisis regresi data panel adalah *Random Effect Model* (REM). Model *random effect* dapat mengestimasi data panel menggunakan variabel gangguan yang memungkinkan saling berkorelasi antar waktu dan antar individu. Metode yang tepat dipakai pada REM yaitu metode *Generalized Least Square* (GLS). Pada metode GLS akan menciptakan estimator yang memenuhi karakteristik sebagai *Best Linear Unbiased Estimation* (BLUE). Estimasi model REM yang dilakukan menggunakan metode GLS dan diterapkan pada data emisi CO<sub>2</sub> di 5 negara anggota ASEAN tahun 2000-2021. Tujuan dari penelitian ini yaitu untuk menentukan model regresi data panel *random effect* dengan metode GLS dan mengetahui variabel-variabel yang berpengaruh terhadap emisi CO<sub>2</sub> di 5 negara anggota ASEAN. Hasil penelitian ini diperoleh model *random effect* dengan variabel pertumbuhan ekonomi dan populasi penduduk memiliki pengaruh signifikan pada emisi CO<sub>2</sub> sedangkan variabel konsumsi energi tidak memiliki pengaruh signifikan pada emisi CO<sub>2</sub>. Dengan model persamaan hasil estimasi yaitu  $Y_{it} = -62,851 - 0,00109X_{1it} + 0,0009X_{2it} + 5,633X_{3it}$ .

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

**PEMODELAN REGRESI DATA PANEL *RANDOM EFFECT* DENGAN  
METODE *GENERALIZED LEAST SQUARE* (GLS) UNTUK  
MENGANALISIS FAKTOR YANG MEMPENGARUHI EMISI KARBON  
DIOKSIDA (CO<sub>2</sub>)**

**Oleh**

**MERYAM JOVANKA SOLEH  
2017031070**

**Skripsi**

Sebagai Salah Satu Syarat Untuk Mencapai Gelar  
**SARJANA MATEMATIKA**

Pada

Jurusan Matematika  
Fakultas Matematika Dan Ilmu Pengetahuan Alam  
Universitas Lampung



**JURUSAN MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMPUNG  
BANDAR LAMPUNG  
2024**

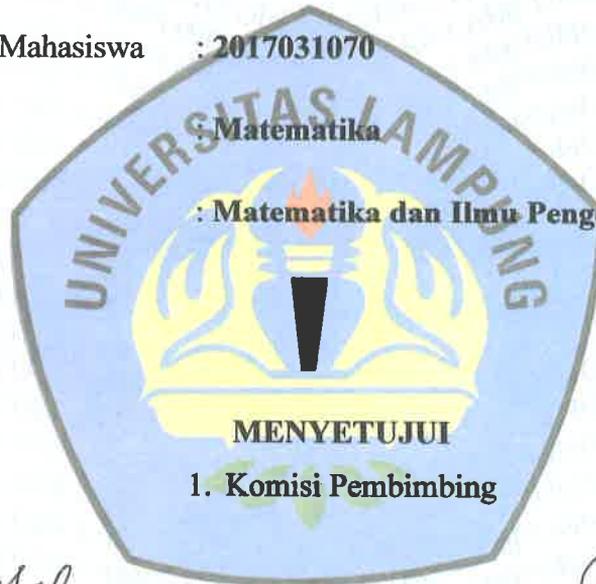
Judul : **PEMODELAN REGRESI DATA PANEL  
RANDOM EFFECT DENGAN METODE  
GENERALIZED LEAST SQUARE (GLS)  
UNTUK MENGANALISIS FAKTOR YANG  
MEMPENGARUHI EMISI KARBON  
DIOKSIDA (CO<sub>2</sub>)**

Nama Mahasiswa : **Meryam Jovanka Soleh**

Nomor Pokok Mahasiswa : **2017031070**

Jurusan : **Matematika**

Fakultas : **Matematika dan Ilmu Pengetahuan Alam**



**Prof. Drs. Mustofa Usman, M.A., Ph.D.**  
NIP. 195701011984031020

**Siti Laelatul Chasanah, S.Pd., M.Si.**  
NIP. 199306012019032021

2. Ketua Jurusan Matematika

**Dr. Aang Nuryaman, S.Si, M.Si.**  
NIP. 197403162005011001

**MENGESAHKAN**

**1. Tim Penguji**

**Ketua**

**: Prof. Drs. Mustofa Usman, M.A., Ph.D.**



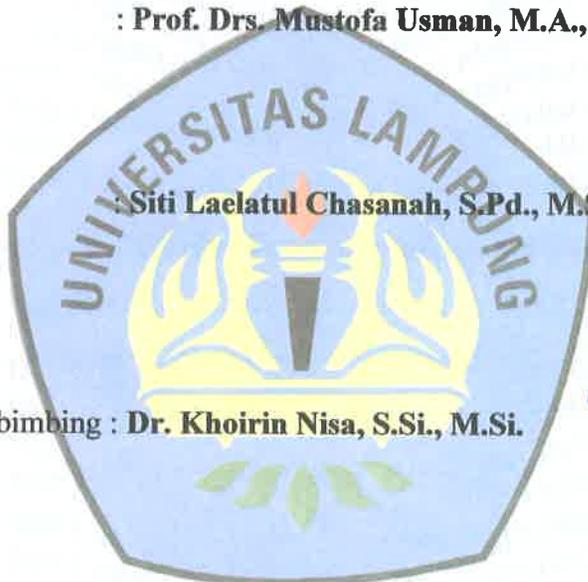
**Sekretaris**

**: Siti Laelatul Chasanah, S.Pd., M.Si.**



**Penguji**

**Bukan Pembimbing : Dr. Khoirin Nisa, S.Si., M.Si.**



**2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam**



**Dr. Eng. Heri Satria, S.Si., M.Si.**  
**NIP. 197110012005011002**

**Tanggal Lulus Ujian Skripsi : 11 Juni 2024**

## PERNYATAAN SKRIPSI MAHASISWA

Saya yang bertanda tangan di bawah ini:

Nama : **Meryam Jovanka Soleh**

Nomor Pokok Mahasiswa : **2017031070**

Jurusan : **Matematika**

Judul skripsi : **Pemodelan Regresi Data Panel *Random Effect* dengan Metode *Generalized Least Square* (GLS) untuk Menganalisis Faktor yang Mempengaruhi Emisi Karbon Dioksida (CO<sub>2</sub>)**

Dengan ini menyatakan bahwa skripsi ini adalah hasil pekerjaan saya sendiri dan semua tulisan yang tertuang dalam skripsi ini telah mengikuti kaidah karya penulisan ilmiah Universitas Lampung.

Bandar Lampung, 11 Juni 2024  
Penulis



**Meryam Jovanka Soleh**  
**NPM. 2017031070**

## **RIWAYAT HIDUP**

Penulis bernama lengkap Meryam Jovanka Soleh, lahir di Garut pada tanggal 25 Juni 2002. Penulis merupakan anak pertama dari dua bersaudara yang lahir dari pasangan Bapak Eri Suyatna dan Ibu Runtah.

Penulis memulai pendidikan di TK Arrusydah 2 pada tahun 2007-2008. Kemudian, melanjutkan sekolah ke sekolah dasar di SD Negeri 1 Pinang Jaya pada tahun 2008-2014, sekolah menengah pertama di SMP Negeri 13 Bandar Lampung pada tahun 2014-2017, dan sekolah menengah kejuruan di SMK SMTI Bandar Lampung pada tahun 2017-2020. Pada tahun 2020, penulis diterima sebagai mahasiswa S1 di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung melalui jalur SBMPTN (Seleksi Bersama Masuk Perguruan Tinggi).

Selama menjadi mahasiswa, penulis pernah mengikuti organisasi HIMATIKA (Himpunan Mahasiswa Matematika) Universitas Lampung selama 2 periode yaitu kepengurusan tahun 2021 dan 2022. Pada bulan Januari-Februari 2023, penulis melakukan Kerja Praktik (KP) di PT PLN (Persero) UID Lampung. Sedangkan sebagai bentuk pengabdian mahasiswa kepada masyarakat dan menjalankan Tri Dharma Perguruan Tinggi, penulis melaksanakan Kuliah Kerja Nyata (KKN) periode kedua di Desa Bina Karya Jaya, Kecamatan Putra Rumbia, Kabupaten Lampung Tengah pada bulan Juni-Agustus 2023.

## KATA INSPIRASI

*“Sesungguhnya bersama kesulitan itu ada kemudahan. Maka apabila engkau telah selesai (dari suatu urusan), tetaplah bekerja keras (untuk urusan yang lain). Dan hanya kepada tuhan mu lah engkau berharap.”*  
(Q.S Al-Insyirah: 6-8)

*“Allah tidak membebani seseorang melainkan sesuai dengan kesanggupannya.”*  
(Q.S Al-Baqarah: 286)

*“Tidak ada mimpi yang terlalu tinggi. Tak ada mimpi yang patut diremehkan. Lambungkan setinggi yang kau inginkan dan gapailah dengan selayaknya yang kau harapkan.”*  
-Maudy Ayunda-

*“Jalan hidup orang beda-beda. Timelinenya juga beda-beda. Definisi “hebat” setiap orang juga beda-beda. Jadi, buat standard kita sendiri, buat goals kita sendiri, lalu hidup menurut itu. Kalau bandingin sama orang lain terus menerus, tidak ada puasnya.”*  
-Jerome Polin-

## **PERSEMBAHAN**

*Dengan mengharap rahmat dan keridhaan Allah SWT, kupersembahkan karya sederhana ini kepada:*

### ***Ayah dan Ibu tercinta***

*Yang 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. Terimakasih atas segala hal yang telah diberikan untukku. Atas doa dan ridho kalian, Allah beri kemudahan dalam menjalankan kehidupan ini.*

### ***Adik dan Keluarga tersayang***

*Yang telah memberikan semangat dukungan dan do'a yang tulus untuk selalu berusaha dan berikhtiar kepada Allah SWT.*

### ***Dosen Pembimbing dan Penguji***

*Terimakasih telah senantiasa meluangkan waktu untuk memberikan bimbingan dan ilmu yang berharga kepada penulis.*

***Almamater Tercinta Universitas Lampung***

## SANWACANA

Puji syukur penulis panjatkan kepada Tuhan Yang Maha Esa karena berkat limpahan rahmat dan izin-Nya kepada penulis sehingga penulis dapat menyelesaikan skripsi dengan judul “Pemodelan Regresi Data Panel *Random Effect* dengan Metode *Generalized Least Square (GLS)* untuk Menganalisis Faktor yang Mempengaruhi Emisi Karbon Dioksida (CO<sub>2</sub>)”.

Dalam penyusunan skripsi ini banyak pihak yang telah membantu penulis, untuk itu penulis mengucapkan terimakasih kepada:

1. Bapak Prof. Drs. Mustofa Usman, M.A., Ph.D. selaku pembimbing satu yang telah memberikan waktu, arahan, serta masukan selama proses penyelesaian penyusunan skripsi.
2. Ibu Siti Laelatul Chasanah, S.Pd., M.Si. selaku pembimbing dua yang telah memberikan waktu, arahan serta masukan selaman proses penyelesaian penyusunan skripsi.
3. Ibu Dr. Khoirin Nisa, S.Si., M.Si. selaku dosen pembimbing akademik dan pembahas pada sidang skripsi.
4. Bapak Dr. Aang Nuryaman, S. Si., M.Si. selaku ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
5. Kedua orang tua, adik, dan seluruh keluarga yang selalu memberi nasihat, doa, dukungan dan motivasinya kepada penulis dalam menyelesaikan skripsi.
6. Teman baik penulis antara lain Cindy, Salsa, Happy, Maya, Nada, dan Gustina yang selalu mendengarkan keluh kesah penulis dalam menyelesaikan pengerjaan skripsi.
7. Pimpinan dan Biro Kesekretariatan HIMATIKA FMIPA Unila Periode 2022 yang telah memberikan semangat kepada penulis.

8. Teman-teman KKN Bina Karya Jaya kelompok 1 yang telah memberikan dukungan kepada penulis.
9. Semua pihak yang telah membantu penulis dalam menyelesaikan skripsi ini.

Semoga Allah senantiasa melimpahkan karunia-Nya atas segala kebaikan semua pihak yang terlibat membantu penulis dalam penyelesaian skripsi ini. Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan, akan tetapi penulis berharap semoga skripsi ini dapat memberikan informasi yang bermanfaat.

Bandar Lampung, 11 Juni 2024  
Penulis,

**Meryam Jovanka Soleh**

## DAFTAR ISI

	Halaman
<b>DAFTAR TABEL</b> .....	vi
<b>DAFTAR GAMBAR</b> .....	vii
<b>I. PENDAHULUAN</b> .....	1
1.1 Latar Belakang .....	1
1.2 Tujuan Penelitian.....	3
1.3 Manfaat Penelitian.....	3
<b>II. TINJAUAN PUSTAKA</b> .....	4
2.1 Analisis Regresi.....	4
2.2 Analisis Regresi Data Panel .....	6
2.3 Estimasi Model Regresi Data Panel .....	7
2.3.1 <i>Common Effect Model</i> (CEM) .....	7
2.3.2 <i>Fixed Effect Model</i> (FEM) .....	9
2.3.3 <i>Random Effect Model</i> (REM).....	10
2.4 Pemilihan Model Analisis Regresi Data Panel.....	13
2.4.1 Uji Chow .....	13
2.4.2 Uji Hausman .....	14
2.4.3 Uji <i>Lagrange Multiplier</i> .....	14
2.4.4 Uji Breusch Pagan.....	15
2.5 Pengujian Asumsi Klasik Model Regresi Data Panel .....	16
2.5.1 Uji Normalitas .....	16
2.5.2 Uji Multikolinearitas .....	17
2.6 Uji Kelayakan ( <i>Goodness of Fit</i> ) Model Regresi Data Panel.....	18
2.6.1 Uji F (Uji Simultan) .....	18
2.6.2 Uji T (Uji Parsial).....	19
2.7 Koefisien Determinasi .....	20

2.8 Emisi Karbon Dioksida (CO <sub>2</sub> ).....	21
<b>III. METODOLOGI PENELITIAN</b> .....	<b>22</b>
3.1 Waktu dan Tempat .....	22
3.2 Data Penelitian .....	22
3.3 Metode Penelitian.....	22
<b>IV. HASIL DAN PEMBAHASAN</b> .....	<b>24</b>
4.1 Analisis Deskriptif.....	24
4.2 Estimasi Model Regresi data Panel .....	25
4.2.1 <i>Common Effect Model</i> (CEM) .....	25
4.2.2 <i>Fixed Effect Model</i> (FEM) .....	26
4.2.3 <i>Random Effect Model</i> (REM).....	27
4.3 Pemilihan Model Estimasi Regresi Data Panel .....	28
4.3.1 Uji Chow .....	28
4.3.2 Uji Hausman .....	28
4.3.3 Uji <i>Lagrange Multiplier</i> .....	29
4.3.4 Uji Breusch Pagan.....	29
4.4 Uji Asumsi Model Regresi Data Panel.....	30
4.4.1 Uji Normalitas .....	31
4.4.2 Uji Multikolinearitas .....	32
4.5 Pengujian Parameter Model Regresi .....	33
4.5.1 Uji F (Uji Simultan) .....	33
4.5.2 Uji T (Uji Parsial).....	33
4.6 Uji Koefisien Determinasi.....	34
4.7 Model Akhir Regresi Data Panel.....	35
<b>V. KESIMPULAN</b> .....	<b>37</b>
<b>DAFTAR PUSTAKA</b> .....	<b>38</b>
<b>LAMPIRAN</b> .....	<b>41</b>

## DAFTAR TABEL

Tabel	Halaman
1. Statistika deskriptif variabel penelitian .....	24
2. Estimasi CEM untuk masing-masing variabel .....	26
3. Estimasi FEM untuk masing-masing variabel .....	26
4. Efek individu terhadap model FEM .....	27
5. Estimasi REM untuk masing-masing variabel .....	27
6. Hasil uji Chow .....	28
7. Hasil uji Hausman .....	29
8. Hasil uji <i>Lagrange Multiplier</i> .....	29
9. Hasil uji Breusch Pagan .....	30
10. Hasil uji normalitas. ....	31
11. Hasil uji multikolinearitas .....	32
12. Hasil uji simultan .....	33
13. Hasil uji parsial .....	34

## DAFTAR GAMBAR

Gambar	Halaman
1. Grafik uji normalitas .....	32

## I. PENDAHULUAN

### 1.1 Latar Belakang

Di era globalisasi ini, pemanasan global menjadi permasalahan lingkungan yang serius di dunia. Terjadinya pemanasan global berakibat dari peningkatan konsentrasi GRK sehingga dampak yang terjadi dari pemanasan global sendiri yaitu berubahnya iklim secara ekstrim. Adapun gas-gas yang terkandung di dalam GRK sendiri meliputi enam komponen gas seperti gas karbon dioksida (CO<sub>2</sub>), N<sub>2</sub>O, CH<sub>4</sub>, HFCs, PFCs dan SF<sub>6</sub>. Emisi CO<sub>2</sub> berkontribusi paling besar mencapai 75% dari total emisi dalam GRK (Sukadri, 2012). Pertumbuhan emisi CO<sub>2</sub> lebih pesat sebesar 5,5% pertahun antara tahun 1990 dan 2010 di negara-negara *Associaton of Southeast Asian Nations* (ASEAN). Sementara itu, pertumbuhan emisi CO<sub>2</sub> di negara-negara OECD (*Organization for Economic Cooperation and Development*) hanya sekitar 0,7% dalam periode yang sama (OECD, 2011).

Peningkatan emisi CO<sub>2</sub> bisa dipengaruhi dari bermacam-macam faktor diantaranya konsumsi energi, luas kawasan hutan, pertumbuhan ekonomi, penanaman modal asing, keterbukaan perdagangan internasional, dan populasi penduduk kota. Pada tahun 2017, Fauzi menganalisis pengaruh dari luas kawasan hutan, konsumsi energi, dan pertumbuhan ekonomi pada emisi karbon dioksida di enam negara ASEAN. Penelitian ini menghasilkan kesimpulan bahwa ketiga faktor tersebut berpengaruh signifikan terhadap emisi CO<sub>2</sub>. Selanjutnya, Candra (2018) menganalisis pengaruh dari pertumbuhan ekonomi dan penanaman modal asing pada emisi CO<sub>2</sub> di 8 negara ASEAN di tahun 2004-2013. Hasil penelitian

tersebut didapat kesimpulan yaitu masing masing variabel yang berpengaruh signifikan positif dan negatif terhadap emisi CO<sub>2</sub> yaitu variabel konsumsi energi dan industri sedangkan variabel penanaman modal asing dan PDB tidak memiliki pengaruh signifikan.

Penelitian lain dilakukan oleh Widyawati, dkk (2021) mengenai pengaruh dari pertumbuhan ekonomi, keterbukaan perdagangan internasional, dan populasi penduduk kota terhadap emisi karbon dioksida di negara ASEAN. Pada penelitian tersebut yang memiliki pengaruh signifikan dan negatif yaitu variabel keterbukaan perdagangan internasional dan pertumbuhan ekonomi. Sedangkan, variabel yang mempunyai pengaruh signifikan dan positif yaitu populasi penduduk.

Untuk mengetahui pengaruh faktor-faktor eksternal terhadap peningkatan emisi CO<sub>2</sub> dapat dilakukan dengan menggunakan analisis regresi. Adapun tujuannya yaitu untuk mengidentifikasi dan mengukur hubungan antara variabel-variabel dengan melibatkan teknik-teknik pemodelan dan analisis terhadap berbagai variabel (Armstrong, 2012). Model regresi yang mempergunakan data panel dinamakan dengan model regresi data panel.

Menurut Gujarati (2003), gabungan antara data silang dan runtun waktu disebut dengan data panel. Kelebihan dari memakai data panel dapat menyediakan pengetahuan yang komprehensif dan utuh untuk melakukan pemodelan dalam analisis regresi data panel. Jaya & Sunengsih (2009), mengatakan bahwa pendekatan analisis statistik dengan menggunakan data panel yang tujuannya guna mengetahui korelasi antara satu atau lebih variabel bebas terhadap satu variabel terikat disebut dengan analisis regresi data panel.

Untuk memutuskan model yang sesuai pada penelitian ini digunakan pendekatan REM (*Random Effect Model*) memakai metode GLS (*Generalized Least Square*). Metode tersebut dapat menciptakan estimator yang menyediakan karakteristik sebagai BLUE (*Best Linear Unbiased Estimation*). Sehingga pendekatan ini dapat mengatasi pelanggaran asumsi homoskedastisitas dan autokorelasi (Gujarati, 2003). Berdasarkan paparan di atas, penulis tertarik untuk melakukan pemodelan

regresi data panel *Random Effect* dengan metode GLS untuk menganalisis faktor yang mempengaruhi emisi karbon dioksida (CO<sub>2</sub>).

## **1.2 Tujuan Penelitian**

Tujuan dari penelitian ini adalah:

1. Menentukan model regresi data panel *random effect* dengan metode GLS pada pengaruh pertumbuhan ekonomi, konsumsi energi dan populasi penduduk pada emisi CO<sub>2</sub> di 5 negara ASEAN.
2. Menganalisis pengaruh pertumbuhan ekonomi, konsumsi energi dan populasi penduduk pada emisi CO<sub>2</sub> di 5 negara ASEAN.

## **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. Sebagai sumber acuan dalam hal pengolahan data menggunakan data panel.
3. Dapat mengaplikasikan estimasi model regresi data panel *random effect* menggunakan metode GLS.

## II. TINJAUAN PUSTAKA

### 2.1 Analisis Regresi

Suatu teknik statistika yang dipergunakan guna memahami keterkaitan linear antara satu variabel bebas dengan satu variabel terikat disebut dengan analisis regresi. Terdapat dua kategori variabel yang ada pada analisis regresi yakni variabel terikat dan variabel bebas. Variabel terikat dianggap sebagai variabel yang terpengaruh, sedangkan variabel bebas dianggap sebagai variabel yang memberikan pengaruh (Sulistiyowati & Astuti, 2017).

Menurut Sumodiningrat (2007) dalam teori ekonomi, hubungan atau persamaan sering kali memiliki spesifikasi hubungan pasti (*exact*) atau deterministik antara variabel-variabelnya. Oleh karena hubungan yang tidak pasti (*non-exact*) jarang terjadi dalam ekonomi, maka unsur-unsur stokastik menjadi penting dalam konteks hubungan ekonomi. Seiring dengan meningkatnya permintaan untuk menguji teori-teori ekonomi, penting untuk menguji keberadaan variabel stokastik dalam hubungan ekonomi. Model regresi linear sederhana merupakan salah satu bentuk paling dasar dari hubungan acak antara dua variabel.

$$Y_i = \beta_0 + \beta_1 X_i + e_i \quad (2.1)$$

dengan

$Y_i$  = nilai variabel terikat

$X_i$  = nilai variabel bebas

$\beta_0$  dan  $\beta_1$  = parameter-parameter regresi

$e_i$  = nilai variabel gangguan atau residual

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

Suatu pendekatan statistik yang dilakukan guna mengidentifikasi hubungan linear antara dua atau lebih variabel bebas dan variabel terikat disebut sebagai analisis regresi linear berganda. Kegunaan dari analisis ini yaitu guna meramalkan nilai variabel terikat ketika nilai variabel bebas mengalami peningkatan atau penurunan serta menentukan hubungan variabel independen apakah memiliki korelasi positif atau negatif dengan variabel dependen (Draper & Smith, 1992). Berikut merupakan model dari persamaan regresi linear berganda:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + u_i \quad (2.2)$$

dengan

$Y_i$	= nilai variabel terikat
$X_{ki}$	= nilai variabel bebas
$\beta_0$	= <i>intercept</i>
$\beta_1, \beta_2, \dots, \beta_k$	= koefisien kemiringan ( <i>slope</i> )
$u_i$	= <i>error</i>
$i$	= $1, 2, 3, \dots, N$

Menurut Widarjono (2009), berikut adalah asumsi-asumsi penting pada regresi linear berganda yaitu:

1. Variabel terikat  $Y$  dan variabel bebas  $X$  memiliki hubungan linear didalam parameter.
2. Dalam tiap variabel independen tidak terdapat multikolinearitas.
3. Variabel gangguan  $u_i$  memiliki rata rata nol.

$$E(u|X_i) = 0 \quad (2.3)$$

4. Tidak ada korelasi antara  $e_i$  dan  $e_j$ .

$$\begin{aligned} cov(u_i, u_j | X_i, X_j) &= E[(u_i - E(u_i|X_i))[(u_j - E(u_j|X_j))] \\ &= E(u_i|X_i)(u_j|X_j) \\ &= 0 \end{aligned} \quad (2.4)$$

5. Variansi dari variabel gangguan atau residual  $u_i$  adalah sama (homoskedastisitas).

$$\begin{aligned} Var(u_i | X_i) &= E[u_i - E(u_i|X_i)]^2 \\ &= E(u_i^2 | X_i) \\ &= \sigma^2 \end{aligned} \quad (2.5)$$

6. Variabel gangguan  $u_i$  berdistribusi normal.

## 2.2 Analisis Regresi Data Panel

Menurut Jaya & Sunengsih (2009), gabungan antara data silang dan runtun waktu dengan unit-unit lintas sektor yang sama diobservasi dalam beberapa periode waktu disebut sebagai data panel. Data panel memiliki karakteristik yang berbeda dari data silang dan runtun waktu. Data panel dapat dibedakan menjadi data panel seimbang atau tidak yang dilihat dari total unit waktu dalam tiap unit individu. Jika setiap individu diamati pada titik waktu yang bersamaan, maka data panel dianggap seimbang (*balanced panel data*). Namun, apabila tidak seluruh unit individu diamati dititik waktu yang bersamaan atau jika terdapat kekurangan data untuk satu atau beberapa individu, dengan demikian data panel dianggap tidak seimbang (*unbalanced panel data*) (Juanda & Junaidi, 2012).

Suatu pendekatan analisis regresi yang menggunakan data panel guna mengidentifikasi keterkaitan antara satu atau lebih variabel bebas pada satu variabel terikat merupakan pengertian dari analisis regresi data panel (Jaya & Sunengsih, 2009). Berikut adalah bentuk umum dari model regresi data panel:

$$Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + u_{it} \quad (2.6)$$

dengan

$Y_{it}$  = variabel dependen dengan individu ke- $i$  dan periode ke- $t$

$X_{kit}$  = variabel independen ke- $k$  dengan individu ke- $i$  dan periode ke- $t$

$\beta$  = parameter yang ditaksir

$u_{it}$  = galat dengan individu ke- $i$  dan waktu ke- $t$

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

$t$  = 1, 2, 3, ...,  $T$

Menurut Gujarati (2003), keunggulan penggunaan data panel baik dari perspektif statistik serta teori ekonomi, antara lain:

1. Data panel memungkinkan perhitungan eksplisit terhadap heterogenitas individu dengan memberikan ruang bagi variabel spesifik individu.
2. Data panel memiliki potensi meningkatkan penelitian empiris menggunakan teknik yang tidak mungkin dicapai hanya dengan memakai data silang atau data runtun waktu.
3. Data panel memanfaatkan pengulangan observasi data silang, sehingga metode ini cocok dipakai sebagai studi mengenai adaptasi dinamis.
4. Banyaknya jumlah observasi pada data panel menghasilkan data yang lebih rinci, memiliki ragam lebih besar, mengurangi kolinearitas antar variabel, dan meningkatkan derajat kebebasan. Hal ini membantu mendapatkan hasil estimasi yang lebih efisien.

### 2.3 Estimasi Model Regresi Data Panel

Dalam estimasi yang ada pada model regresi data panel dibedakan menjadi tiga kategori model regresi data panel yang dapat digunakan. Model-model tersebut antara lain yaitu *Common Effect Model* (CEM), *Fixed Effect Model* (FEM) dan *Random Effect Model* (REM).

#### 2.3.1 *Common Effect Model* (CEM)

CEM adalah model yang tidak mempertimbangkan perbedaan antar unit *cross section* dan *time series* dengan cara menggabungkan (*pooled*) semua data silang dan data runtun waktu lalu memakai metode OLS untuk mengestimasi parameter-parameternya (Baltagi, 2005). Menurut Gujarati (2003), asumsi dalam CEM yaitu intersep dan *slope* konstan dalam semua unit waktu dan individu. Berikut adalah bentuk umum persamaan model CEM.

$$Y_{it} = \beta_0 + \sum_{k=1}^K \beta_k X_{kit} + u_{it} \quad (2.7)$$

dengan

$Y_{it}$  = variabel dependen dengan individu ke- $i$  dan periode ke- $t$

$X_{kit}$  = variabel independen ke- $k$  dengan individu ke- $i$  dan periode ke- $t$

$\beta_k$  = parameter untuk variabel ke- $k$

$\beta_0$  = intersep

$u_{it}$  = galat dalam individu ke- $i$  dan waktu ke- $t$

Metode yang sering dipakai guna memperkirakan nilai parameter pada persamaan regresi linear disebut metode *Ordinary Least Square* (OLS). Metode ini juga memiliki nama lain yaitu Metode Kuadrat Terkecil (MKT). Berikut adalah prinsip dasar MKT:

$$\mathbf{u} = \mathbf{Y} - \mathbf{X}\boldsymbol{\beta} \quad (2.8)$$

Maka diperoleh total kuadrat kesalahan sebagai berikut:

$$\begin{aligned} \mathbf{u}'\mathbf{u} &= (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) \\ &= (\mathbf{Y}' - \boldsymbol{\beta}'\mathbf{X}')(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) \\ &= \mathbf{Y}'\mathbf{Y} - \boldsymbol{\beta}'\mathbf{X}'\mathbf{Y} - \mathbf{Y}'\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta} \\ &= \mathbf{Y}'\mathbf{Y} - 2\boldsymbol{\beta}'\mathbf{X}'\mathbf{Y} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta} \end{aligned} \quad (2.9)$$

Karena matriks *transpose*  $(\mathbf{X}\boldsymbol{\beta})' = \boldsymbol{\beta}'\mathbf{X}'$ , maka *scalar*  $\boldsymbol{\beta}'\mathbf{X}'\mathbf{Y} = \mathbf{Y}'\mathbf{X}\boldsymbol{\beta}$ .

Untuk memperoleh estimasi parameter  $\boldsymbol{\beta}$  yang meminimalkan jumlah kuadrat galat, langkahnya adalah dengan melakukan turunan dari persamaan (2.9) terhadap parameter  $\boldsymbol{\beta}$ , dan kemudian menyeimbangkan hasil turunan tersebut menjadi nol atau  $\frac{\partial(\mathbf{u}'\mathbf{u})}{\partial\boldsymbol{\beta}} = 0$ , dengan demikian diperoleh:

$$\begin{aligned} \frac{\partial(\mathbf{Y}'\mathbf{Y} - 2\boldsymbol{\beta}'\mathbf{X}'\mathbf{Y} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta})}{\partial\boldsymbol{\beta}} &= \mathbf{0} \\ -2\mathbf{X}'\mathbf{Y} + \mathbf{X}'\mathbf{X}\widehat{\boldsymbol{\beta}} + \widehat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{X} &= \mathbf{0} \\ -2\mathbf{X}'\mathbf{Y} + 2\mathbf{X}'\mathbf{X}\widehat{\boldsymbol{\beta}} &= \mathbf{0} \\ 2\mathbf{X}'\mathbf{X}\widehat{\boldsymbol{\beta}} &= 2\mathbf{X}'\mathbf{Y} \\ \mathbf{X}'\mathbf{X}\widehat{\boldsymbol{\beta}} &= \mathbf{X}'\mathbf{Y} \\ (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{X})\widehat{\boldsymbol{\beta}} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} \\ \widehat{\boldsymbol{\beta}} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} \end{aligned} \quad (2.10)$$

### 2.3.2 Fixed Effect Model (FEM)

Istilah *fixed effect* digunakan karena walaupun nilai intersep bervariasi dalam setiap unit *cross section*, dalam model FEM koefisien (*slope*) dari variabel independen tidak mengalami perubahan antara individu maupun seiring waktu (Gujarati, 2003). Diasumsikan pada model FEM bahwa perbedaan nilai tiap individu yang dinyatakan oleh *unobserved factor* dapat berkorelasi dengan variabel independen (Wahyudi, 2020). Persamaan dalam FEM adalah sebagai berikut.

$$Y_{it} = \beta_{0i} + \sum_{k=1}^K \beta_k X_{kit} + u_{it} \quad (2.11)$$

dengan

$Y_{it}$  = variabel dependen dalam individu ke- $i$  dan periode ke- $t$

$X_{kit}$  = variabel independen ke- $k$  dalam individu ke- $i$  dan periode ke- $t$

$\beta_{0i}$  = intersep dalam individu ke- $i$

$\beta_k$  = parameter untuk variabel ke- $k$

$u_{it}$  = *error* dalam individu ke- $i$  dan waktu ke- $t$

$K$  = banyak parameter regresi yang akan ditaksir

Metode yang dipakai dalam melakukan estimasi parameter regresi pada FEM dalam data panel disebut sebagai *Least Square Dummy Variable* (LSDV). Variabel *dummy* yang dimaksud yaitu unit dalam *cross section* menjadi variabel dalam model. Menurut Greene (2012), LSDV adalah teknik yang memanfaatkan MKT untuk memperkirakan parameter dalam regresi linear pada model menggunakan variabel *dummy* yang menjadi salah satu variabel independennya. Variabel tersebut merupakan variabel yang mempunyai dua potensi nilai yang dilakukan dengan memberikan angka 1 pada salah satu golongan, sementara golongan lainnya diberikan angka 0. Variabel *dummy* dipakai guna menjelaskan perbedaan intersep antar unit dalam *cross section* dalam FEM.

Model FEM yang umumnya menggunakan LSDV dapat dirumuskan sebagai berikut.

$$Y_{it} = \beta_{0i} + \sum_{i=1}^N \beta_{0i} D_{ki} + \sum_{k=1}^K \beta_k X_{kit} + u_{it} \quad (2.12)$$

dengan

$Y_{it}$  = variabel dependen dalam individu ke- $i$  dan periode ke- $t$

$X_{it}$  = variabel independen ke- $k$  dalam individu ke- $i$  dan periode ke- $t$

$D_{ki}$  = variabel *dummy* ke- $k$  dalam individu ke- $i$

$\beta_k$  = parameter untuk variabel ke- $k$

$u_{it}$  = nilai *error* untuk individu ke- $i$  dan waktu ke- $t$

### 2.3.3 *Random Effect Model (REM)*

Asumsi yang terdapat pada REM yakni bahwa intersep merupakan variabel *random* atau stokastik. Model ini mengestimasi data panel menggunakan variabel gangguan yang memungkinkan satu sama lain berkorelasi antar waktu dan antar individu. Karena ada dua aspek yang turut serta terhadap pembentukan *error* pada REM maka dapat dibedakan menjadi variabel *error* gabungan serta variabel *error* secara individu yang bersifat *random* dan tidak dapat diamati secara langsung (*unobserved factor*), maka model ini sering dikatakan sebagai *Error Component Model (ECM)* (Widarjono, 2009). Adapun berikut merupakan persamaan REM:

$$Y_{it} = \beta_0 + \sum_{k=1}^K \beta_k X_{kit} + w_{it} \quad (2.13)$$

dengan

$$w_{it} = \varepsilon_i + u_{it} \quad (2.14)$$

keterangan:

$Y_{it}$  = variabel dependen dalam individu ke- $i$  dan periode ke- $t$ ,

$X_{kit}$  = variabel independen ke- $k$  dalam individu ke- $i$  dan periode ke- $t$

$\beta_k$  = parameter untuk variabel ke- $k$

$w_{it}$  = komponen galat dalam individu ke- $i$  dan periode ke- $t$

$\varepsilon_i$  = komponen galat *cross section*

$u_{it}$  = komponen galat gabungan *time series & cross section*

$K$  = banyak parameter regresi yang akan ditaksir

Adapun asumsi yang diterapkan pada komponen *error* tersebut adalah:

$$\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$$

$$u_{it} \sim N(0, \sigma_u^2)$$

$$E(\varepsilon_i u_{it}) = 0 \quad (2.15)$$

$$E(\varepsilon_i \varepsilon_j) = 0, i \neq j \quad (2.16)$$

$$E(u_{it} u_{is}) = E(u_{it} u_{jt}) = E(u_{it} u_{js}) = 0, i \neq j; t \neq s \quad (2.17)$$

Dengan mempertimbangkan persamaan (2.13) dan (2.14), dapat dijelaskan bahwa dalam REM efek rata-rata dari semua data intersep (*cross section*) diwakili oleh *intercept* bersama ( $\beta_0$ ). Sementara  $\varepsilon_i$  merupakan *error random* dari masing-masing intersep individu terhadap nilai rata-rata tersebut. Maka dari itu *varians* dari *error* tersebut dapat ditulis dengan:

$$E(w_{it}) = 0 \quad (2.18)$$

$$Var(w_{it}) = \sigma_\varepsilon^2 + \sigma_u^2 \quad (2.19)$$

Jika  $\sigma_\varepsilon^2 = 0$  maka variansi komponen *error*  $w_{it}$  bersifat homoskedastik, akan tetapi terdapat korelasi antara  $w_{it}$  dan  $w_{is}$  yaitu:

$$corr(w_{it}, w_{is}) = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_u^2}; t \neq s \quad (2.20)$$

Oleh karena itu, model efek *random* tidak dapat diestimasi menggunakan metode OLS. Metode yang sesuai untuk mengestimasi REM adalah metode GLS dengan asumsi bahwa homoskedastisitas dan tidak terdapat *cross-sectional correlation*. Menurut Greene (2012), metode *Generalized Least Square* atau dapat disebut sebagai kuadrat terkecil yang diberlakukan secara umum adalah metode estimasi melalui pembobotan (*weighted*) untuk menanggulangi kasus heteroskedastisitas dan autokorelasi.

Gujarati (2003) mengatakan bahwa metode GLS memberikan hasil yang lebih efektif dan stabil daripada metode OLS. Metode GLS memanfaatkan informasi tentang struktur varian kovarian dalam data secara langsung dan oleh karenanya dapat menghasilkan BLUE. Dengan matriks kovarians  $\Omega$  dapat dinyatakan sebagai hasil kali  $C, D$  dan transpos  $C$  sebagai berikut:

$$\Omega = CDC' \quad (2.21)$$

Dengan bentuk matriks  $D$  yang dapat ditulis pada persamaan (2.22).

$$\mathbf{D} = \begin{bmatrix} d_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_{nn} \end{bmatrix} \quad (2.22)$$

Didefinisikan matriks  $\mathbf{G}$  sehingga transposenya adalah perkalian matriks  $\mathbf{C}$  dan  $\mathbf{D}$  sebagai berikut:

$$\mathbf{G}' = \mathbf{C}\mathbf{D}^{\left(-\frac{1}{2}\right)} \quad (2.23)$$

Diberikan model regresi linear umum:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \quad (2.24)$$

Persamaan (2.24) dikalikan dengan matriks  $\mathbf{G}$  sehingga diperoleh:

$$\mathbf{G}\mathbf{Y} = \mathbf{G}\mathbf{X}\boldsymbol{\beta} + \mathbf{G}\mathbf{u} \quad (2.25)$$

Dengan  $\mathbf{E}[(\mathbf{G}\mathbf{u})(\mathbf{G}\mathbf{u})'] = \mathbf{I}$  atau ekuivalen dengan  $\mathbf{G}\boldsymbol{\Omega}\mathbf{G}' = \mathbf{I}$  dan  $\mathbf{G}'(\mathbf{G}\boldsymbol{\Omega}\mathbf{G}')\mathbf{G} = \mathbf{G}'\mathbf{G}$  sehingga  $\mathbf{G}'\mathbf{G} = \boldsymbol{\Omega}^{-1}$ . Dari persamaan (2.25) dapat dibentuk persamaan baru untuk galat yaitu:

$$\mathbf{G}\mathbf{u} = \mathbf{G}\mathbf{Y} - \mathbf{G}\mathbf{X}\boldsymbol{\beta} \quad (2.26)$$

Turunan pertama dari jumlah kuadrat  $\mathbf{G}\mathbf{u}$  disamakan dengan nol sehingga diperoleh:

$$\begin{aligned} \hat{\boldsymbol{\beta}} &= [(\mathbf{G}\mathbf{X})'(\mathbf{G}\mathbf{X})]^{-1}(\mathbf{G}\mathbf{X})'(\mathbf{G}\mathbf{Y}) \\ &= (\mathbf{X}'\mathbf{G}'\mathbf{G}\mathbf{X})^{-1}(\mathbf{X}'\mathbf{G}'\mathbf{G}\mathbf{Y}) \\ &= (\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X})^{-1}(\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{Y}) \end{aligned} \quad (2.27)$$

Metode GLS merupakan *Best Linear Unbiased Estimator* (BLUE). Salah satu teknik untuk memperkirakan  $\boldsymbol{\Omega}$  yaitu dengan memodelkannya sebagai berikut:

$$\boldsymbol{\Omega} = \begin{bmatrix} w_{11}^2 & \rho & \rho^2 & \cdots & \rho^{n-1} \\ \rho & w_{22}^2 & \rho & \cdots & \rho^{n-2} \\ \rho^2 & \rho & w_{33}^2 & \cdots & \rho^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \cdots & \cdots & w_{nn}^2 \end{bmatrix} \quad (2.28)$$

dengan

$w_{ii}^2$  = varians ke- $i$  dari  $\varepsilon_i$

$\rho$  = korelasi antara  $\varepsilon_i$  dan  $\varepsilon_{i+1}$

## 2.4 Pemilihan Model Analisis Regresi Data Panel

Dalam analisis regresi data panel pemilihan model estimasi merupakan langkah penting dalam menentukan model estimasi terbaik. Untuk melakukan pemilihan model terbaik digunakan 4 uji antara lain yakni uji Chow, uji Hausman, uji *Lagrange Multiplier* dan uji Breusch Pagan.

### 2.4.1 Uji Chow

Uji Chow atau bisa disebut sebagai Uji *Likelihood Ratio* dapat dilakukan guna membandingkan antara FEM atau CEM yang merupakan model yang lebih baik digunakan. Dalam hal ini, pengujian signifikansi model FEM dapat dilakukan menggunakan uji statistik  $F$  (Juanda & Junaidi, 2012).

Hipotesis statistik:

$H_0: \alpha_1 = \alpha_2 = \dots = \alpha_n = 0$  (model yang dipakai yakni CEM)

$H_1$ : minimal ada satu  $\alpha_i \neq 0; i = 1, 2, 3, \dots, n$  (model yang dipakai yakni FEM)

Statistik uji yang digunakan yaitu statistik uji  $F$  dengan menggunakan rumus berikut:

$$F_{hitung} = \frac{(RSS_1 - RSS_2)/(n-1)}{(RSS_2)/(nT-n-K)} \quad (2.29)$$

dengan

$T$  = jumlah periode waktu

$n$  = jumlah individu

$K$  = parameter yang ditaksir

$RSS_1$  = *residual sum of squares* dalam CEM

$RSS_2$  = *residual sum of squares* dalam FEM

Dasar penolakan pada hipotesis tersebut yaitu jika nilai  $F_{hitung} > F_{tabel}$  atau  $p\text{-value} < \alpha$  maka tolak  $H_0$ . Akibatnya, cukup bukti bahwa model FEM yang lebih tepat digunakan. Sebaliknya, apabila nilai  $F_{hitung} < F_{tabel}$  atau  $p\text{-value} > \alpha$  maka

tidak tolak  $H_0$ . Akibatnya, cukup bukti untuk model CEM yang lebih tepat digunakan.

#### 2.4.2 Uji Hausman

Menurut Madany, dkk. (2022), Uji Hausman bertujuan guna mengetahui apakah antara kesalahan model dengan satu atau lebih variabel bebas memiliki korelasi. Uji ini juga dapat dipakai untuk menentukan antara REM atau FEM yang merupakan model yang lebih baik digunakan.

Hipotesis statistik:

$H_0: \rho_{\mu_i x_i} = 0$  (model yang dipakai yakni REM)

$H_1: \rho_{\mu_i x_i} \neq 0$  (model yang dipakai yakni FEM)

Statistik uji yang dipakai yaitu mengikuti *chi-squared's distribution* mengacu pada kriteria Wald dengan menggunakan rumus berikut:

$$W = X^2(K) = (\mathbf{b} - \hat{\boldsymbol{\beta}})' [\text{var}(\mathbf{b}) - \text{var}(\hat{\boldsymbol{\beta}})]^{-1} (\mathbf{b} - \hat{\boldsymbol{\beta}}) \quad (2.30)$$

dengan

$\mathbf{b}$  = vektor estimasi parameter FEM

$\hat{\boldsymbol{\beta}}$  = vektor estimasi parameter REM

Dasar penolakan pada hipotesis tersebut yakni apabila nilai  $W > X^2_{(\alpha, k)}$  atau *p-value*  $< \alpha$ , maka tolak  $H_0$ . Akibatnya, cukup bukti bahwa model FEM yang lebih tepat digunakan. Sebaliknya, jika nilai  $W < X^2_{(\alpha, k)}$  atau *p-value*  $> \alpha$ , maka tidak tolak  $H_0$ . Akibatnya, cukup bukti untuk model REM yang lebih tepat digunakan.

#### 2.4.3 Uji Lagrange Multiplier

Uji LM atau bisa disebut dengan uji *Lagrange Multiplier* dapat dilakukan guna membandingkan antara REM atau CEM yang merupakan model yang lebih tepat

digunakan. Uji ini mempertimbangkan nilai residu yang ada dalam model CEM untuk dasar pengujian (Juanda & Junaidi, 2012).

Hipotesis statistik:

$H_0 : \sigma_\mu^2 = 0$  (model yang dipakai adalah CEM)

$H_1 : \sigma_\mu^2 \neq 0$  (model yang dipakai adalah REM)

Statistik uji yang dipakai yakni statistik uji  $LM$  dengan menggunakan rumus berikut:

$$LM = \frac{nT}{2(T-1)} \left[ \frac{\sum_{i=1}^n [\sum_{t=1}^T e_{it}]^2}{\sum_{i=1}^n \sum_{t=1}^T e_{it}^2} - 1 \right]^2 \quad (2.31)$$

dengan

$n$  = jumlah individu

$T$  = jumlah periode waktu

$e_{it}$  = *residual* pada model CEM

Dasar penolakan pada hipotesis tersebut yakni apabila nilai statistik  $LM > X^2_{(\alpha,1)}$  atau  $p\text{-value} < \alpha$ , maka tolak  $H_0$ . Akibatnya, cukup bukti untuk model REM yang lebih tepat digunakan. Sebaliknya, apabila nilai statistik  $LM < X^2_{(\alpha,1)}$  atau nilai  $p\text{-value} > \alpha$ , sehingga terima  $H_0$ . Akibatnya, cukup bukti untuk model CEM yang lebih tepat digunakan.

#### 2.4.4 Uji Breusch Pagan

Menurut Rosadi (2011), guna mengetahui adakah efek kali silang, waktu atau dua arah maka dilakukan uji Breusch Pagan.

Hipotesis:

- Uji efek individu maupun waktu

$H_0^{cd} : c_i = 0, d_t = 0$  atau tidak ada efek individu maupun waktu

$H_1^c : c_i \neq 0, d_t \neq 0$  atau ada efek individu maupun waktu

- Uji efek individu  
 $H_0^c: c_i = 0, d_t \sim iid, N(0, \sigma_d^2)$  atau tidak ada efek individu  
 $H_1^c: c_i \neq 0, d_t \sim iid, N(0, \sigma_d^2)$  atau ada efek individu
- Uji efek waktu  
 $H_0^d: d_t = 0, c_i \sim iid, N(0, \sigma_c^2)$  atau tidak ada efek waktu  
 $H_1^d: d_t \neq 0, c_i \sim iid, N(0, \sigma_c^2)$  atau ada efek waktu

Tingkat signifikansi:

$$\alpha = 5\%$$

Dasar penolakan pada hipotesis tersebut yakni apabila nilai *p-value* kurang dari  $\alpha = 0,05$  maka tolak hipotesis nol.

## 2.5 Pengujian Asumsi Klasik Model Regresi Data Panel

Greene (2012) menjelaskan bahwa dalam menggunakan metode estimasi OLS pada regresi, asumsi klasik mencakup normalitas, tidak adanya multikolinearitas, homoskedastisitas, dan tidak ada autokorelasi harus terpenuhi. Namun pada REM hanya diperlukan uji asumsi klasik yaitu normalitas dan nonmultikolinearitas karena model REM memakai metode estimasi GLS yang dapat menanggulangi autokorelasi dan heteroskedastisitas (Halim, dkk., 2020).

### 2.5.1 Uji Normalitas

Uji normalitas digunakan guna memeriksa residual yang ada pada model memiliki distribusi normal atau tidak. Asumsi pada uji ini mewajibkan residual memiliki distribusi normal dengan *mean* 0 dan varians tertentu. Untuk melakukan uji asumsi ini digunakan uji Jarque-Bera. Uji ini dipakai guna melihat data terdistribusi normal dan ideal untuk sampel besar (Hamid, dkk., 2020).

Hipotesis statistik:

$H_0 = error$  terdistribusi secara normal

$H_1 = error$  tidak terdistribusi secara normal

Statistik uji yang dipakai yakni statistik uji  $JB$  dengan menggunakan rumus berikut:

$$JB = \frac{n}{6} \left[ S^2 + \frac{(K-3)^2}{4} \right] \quad (2.32)$$

$n$  adalah jumlah observasi dalam sampel,  $S$  mewakili tingkat kemiringan (*skewness*), dan  $K$  mewakili tingkat peruncingan (*kurtosis*). Dengan:

$$K = \frac{\hat{\mu}_4}{\hat{\mu}_2} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{\left( \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^2} \quad (2.33)$$

$$S = \frac{\hat{\mu}_3}{\hat{\mu}_2^{\frac{3}{2}}} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3}{\left( \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^{\frac{3}{2}}} \quad (2.34)$$

Dasar penolakan pada hipotesis tersebut yaitu jika nilai uji  $JB > X^2_{(\alpha,2)}$  atau  $p\text{-value} < \alpha$ , maka tolak  $H_0$ . Akibatnya, cukup bukti untuk menyimpulkan bahwa residual tidak berdistribusi normal. Sebaliknya, apabila nilai uji  $JB < X^2_{(\alpha,2)}$  atau  $p\text{-value} > \alpha$ , maka tidak ditolak  $H_0$ . Akibatnya, cukup bukti untuk menyimpulkan bahwa residual terdistribusi normal.

### 2.5.2 Uji Multikolinearitas

Uji multikolinearitas dilakukan guna mengidentifikasi adakah korelasi antar variabel independen pada model regresi (Ghozali, 2016). Menurut Gujarati (2003), penyebab terjadinya multikolinearitas dalam model regresi dikarenakan terdapat keterkaitan yang kuat antara satu atau lebih variabel bebas. Untuk mengidentifikasi multikolinearitas maka dapat diukur melalui nilai *Variance Inflation Factor* (VIF). Ketika nilai VIF  $> 10$  akibatnya dapat diambil keputusan bahwa terjadi multikolinearitas.

$$VIF = \frac{1}{1 - R_k^2} \quad (2.35)$$

dengan  $R_k^2$  merupakan nilai koefisien determinasi variabel independen  $X_k$  yang diregresikan dengan variabel independen lainnya. Apabila variabel independen  $X_k$  tidak memiliki korelasi terhadap variabel independen lainnya, maka nilai  $R_k^2$  memiliki nilai yang kecil dan nilai VIF akan mendekati 1.

## 2.6 Uji Kelayakan Model Regresi Data Panel

Uji kelayakan model diperlukan guna mengidentifikasi hubungan diantara variabel bebas dan variabel terikat. Untuk melakukan uji kelayakan model terdapat dua kategori uji yang digunakan yaitu uji simultan dan uji parsial.

### 2.6.1 Uji F (Uji Simultan)

Uji pengaruh simultan atau dikenal dengan uji F digunakan sebagai penentu variabel-variabel bebas secara serentak mempengaruhi variabel terikat atau tidak. Tujuan dilakukannya uji F yaitu guna mengevaluasi pengaruh simultan dari variabel-variabel bebas pada variabel terikat dan juga dikenal sebagai uji model atau uji anova (*Analysis of Variance*) (Ghozali, 2016).

Hipotesis statistik:

$H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$  (variabel independen secara simultan tidak memberikan pengaruh signifikan dalam variabel dependen)

$H_1$  : paling tidak ada satu *slope*  $\beta_k \neq 0, k = 1, 2, \dots, K$  (variabel independen secara simultan memberikan pengaruh signifikan dalam variabel dependen)

Statistik uji yang digunakan adalah statistik uji  $F$  dengan menggunakan rumus berikut

$$F_{hitung} = \frac{R^2/(k-1)}{(1-R^2)/(n-k)} \quad (2.36)$$

dengan

$R^2$  = koefisien determinasi

$n$  = banyak observasi

$k$  = banyak parameter (termasuk intersep)

Dasar penolakan pada hipotesis tersebut yaitu apabila nilai  $F_{hitung} > F_{tabel}$  atau nilai  $p\text{-value} < \alpha$ , maka tolak  $H_0$ . Akibatnya, cukup bukti untuk menyatakan bahwa variabel independen secara serentak mempunyai pengaruh yang signifikan pada variabel dependen. Sebaliknya, apabila nilai  $F_{hitung} < F_{tabel}$  atau nilai  $p\text{-value} > \alpha$ , maka hipotesis nol tidak ditolak. Akibatnya, cukup bukti untuk menyatakan bahwa variabel independen secara serentak tidak mempunyai pengaruh signifikan pada variabel dependen.

### 2.6.2 Uji T (Uji Parsial)

Pada umumnya tujuan dilakukannya uji T atau yang bisa dikatakan dengan uji parsial yakni guna mengidentifikasi adakah pengaruh signifikan secara individual antara suatu variabel independen pada variabel dependen (Gujarati, 2003).

Hipotesis statistik:

$H_0 : \beta_k = 0$  (variabel bebas secara individual tidak mempunyai pengaruh signifikan pada variabel terikat)

$H_1 : \beta_k \neq 0, k = 0, 1, 2, \dots, K$  (variabel bebas secara individual mempunyai pengaruh signifikan pada variabel terikat)

Statistik uji yang digunakan yakni statistik uji  $t$  dengan menggunakan rumus berikut

$$t = \frac{\hat{\beta}_k}{SE(\hat{\beta}_k)} \quad (2.37)$$

dengan

$\hat{\beta}_k$  = koefisien regresi variabel independen ke- $k$ ;  $k = 1, 2, \dots, K$

$SE(\hat{\beta}_k) = \text{Standard Error } \hat{\beta}_k$

Dasar penolakan pada hipotesis tersebut yaitu jika nilai  $t$  yang dihitung  $|t_{hitung}| > t_{tabel}$  atau nilai  $p\text{-value} < \alpha$ , maka tolak  $H_0$ . Akibatnya, dapat diambil kesimpulan bahwa variabel bebas secara individual mempunyai pengaruh signifikan pada variabel terikat. Sebaliknya, apabila nilai  $|t_{hitung}| < t_{tabel}$  atau nilai  $p\text{-value} > \alpha$ , maka terima  $H_0$ . Akibatnya, dapat diambil kesimpulan variabel bebas secara individual tidak mempunyai pengaruh signifikan terhadap variabel terikat.

## 2.7 Koefisien Determinasi

Koefisien determinasi adalah parameter yang sangat relevan dalam analisis regresi. Koefisien ini, disimbolkan dengan  $R^2$  yang memberikan indikasi tentang seberapa baik model regresi yang diestimasi sesuai dengan data sebenarnya. Dengan kata lain,  $R^2$  mengukur tingkat kesesuaian antara garis regresi yang diestimasi terhadap data aktual (Nachrowi & Usman, 2006). Adapun rumus untuk nilai  $R^2$  adalah:

$$R^2 = \frac{JKR}{JKT} \quad (2.38)$$

dengan

$JKR$  = jumlah kuadrat regresi

$JKT$  = jumlah kuadrat total

Jika koefisien determinasi ( $R^2$ ) memiliki nilai 0 maka ini berarti variasi pada variabel terikat ( $Y$ ) tidak bisa diidentifikasi sedikit pun oleh variabel bebas ( $X$ ). Sebaliknya, jika ( $R^2$ ) memiliki nilai 1 maka ini berarti seluruh variasi pada variabel terikat ( $Y$ ) dapat diidentifikasi sepenuhnya oleh variabel bebas ( $X$ ). Dengan kata lain  $R^2$  merupakan ukuran seberapa baik variabilitas dalam data dapat dijelaskan oleh model regresi. Oleh karena itu, kualitas suatu persamaan regresi dapat dinilai berdasarkan nilai yang berkisar antara 0 dan 1 (Alwi, dkk., 2018).

## 2.8 Emisi Karbon Dioksida (CO<sub>2</sub>)

Gas yang dilepaskan melalui kegiatan pembakaran senyawa yang mengandung karbon antara lain CO<sub>2</sub>, solar, LPG, dan bahan bakar lainnya disebut sebagai emisi karbon. Emisi CO<sub>2</sub> menyumbang kontribusi tertinggi dalam GRK yaitu sebesar 75% (Sukadri, 2012). Jika emisi gas ini berlebihan hal ini dapat mengakibatkan pemanasan global yang signifikan menyebabkan peningkatan suhu di bumi. Pemanasan global disebabkan dari peningkatan suhu akibat tingginya konsentrasi GRK yang terletak pada atmosfer bumi. Adapun faktor-faktor pendorong emisi CO<sub>2</sub> diantaranya sebagai berikut:

1. Pertumbuhan ekonomi

Pertumbuhan ekonomi bisa dijelaskan sebagai peningkatan produksi barang dan jasa dalam suatu masyarakat. Tingkat pertumbuhan ekonomi adalah indikator yang sering digunakan untuk mengukur kinerja makroekonomi dan dapat dihitung berdasarkan nilai Produk Domestik Bruto (PDB) atau *Gross Domestic Product* (GDP).

2. Konsumsi energi

Konsumsi energi primer digunakan untuk mengukur total permintaan energi suatu negara. Hal ini mencakup konsumsi sektor energi itu sendiri, kerugian selama transformasi (misalnya, dari minyak atau gas menjadi listrik) dan distribusi energi, serta konsumsi akhir oleh pengguna akhir. Hal ini tidak termasuk pembawa energi yang digunakan untuk tujuan non-energi (seperti minyak bumi yang tidak digunakan untuk pembakaran tetapi untuk memproduksi plastik) (Eurostat, 2018).

3. Populasi penduduk kota

Definisi populasi penduduk secara de facto merujuk pada penghitungan semua individu dalam suatu wilayah tanpa memperhatikan status hukum atau kewarganegaraan mereka.

### **III. METODOLOGI PENELITIAN**

#### **3.1 Waktu dan Tempat**

Penelitian ini dilakukan pada semester ganjil tahun akademik 2023/2024 di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

#### **3.2 Data Penelitian**

Penelitian ini menggunakan data sekunder dari *world in data* dan *world bank* yang meliputi data emisi CO<sub>2</sub> ( $Y$ ), data GDP (*Gross Domestic Product*) ( $X_1$ ), data konsumsi energi primer ( $X_2$ ), dan data populasi penduduk kota ( $X_3$ ) dari tahun 2000-2021. Objek yang digunakan sebagai unit *cross section*-nya terdiri dari 5 negara ASEAN antara lain Indonesia, Malaysia, Singapura, Thailand dan Philipina.

#### **3.3 Metode Penelitian**

Adapun metode yang dilakukan pada analisis data penelitian ini yaitu sebagai berikut:

1. Melakukan analisis data panel dengan memperkirakan tiga model regresi data panel yaitu CEM menggunakan metode OLS, FEM menggunakan metode LSDV, dan REM dengan metode GLS.

2. Memilih model regresi yang paling sesuai dengan data menggunakan empat uji yaitu uji Chow, uji Hausman, uji *Lagrange Multiplier* dan uji Breusch Pagan.
3. Menguji asumsi-asumsi model regresi seperti uji normalitas dan uji multikolinearitas.
4. Melakukan penilaian model regresi menggunakan uji serentak, uji parsial dan koefisien determinasi.
5. Menginterpretasi hasil analisis dan temuan dari model regresi yang telah dipilih.
6. Membuat kesimpulan berdasarkan hasil analisis untuk menjawab masalah penelitian.

## V. KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan, didapat kesimpulan sebagai berikut:

1. Model estimasi terbaik pada penelitian ini menggunakan *Random Effect Model* (REM) dengan terdapat 2 variabel bebas yang berpengaruh signifikan pada variabel terikat emisi CO<sub>2</sub> yaitu GDP ( $X_1$ ) dan populasi penduduk kota ( $X_3$ ) sedangkan variabel konsumsi energi primer ( $X_2$ ) tidak berpengaruh signifikan dengan bentuk model yaitu:

$$Y_{it} = -62,851 - 0,00109X_{1it} + 0,0009X_{2it} + 5,633X_{3it}$$

2. Variabel independen  $X_1$  (GDP) memiliki korelasi negatif terhadap emisi CO<sub>2</sub> menandakan bahwa negara yang memiliki GDP tinggi kemungkinan lebih mampu mengelola emisi CO<sub>2</sub>. Sedangkan variabel independen  $X_2$  (konsumsi energi primer) dan variabel independen  $X_3$  (populasi penduduk kota) memiliki korelasi positif terhadap emisi CO<sub>2</sub> artinya semakin tinggi konsumsi energi primer dan jumlah populasi penduduk kota maka menyebabkan peningkatan emisi CO<sub>2</sub>.
3. Sebesar 91,12% variabel GDP dan populasi penduduk kota bersama-sama mempengaruhi emisi CO<sub>2</sub> di 5 negara anggota ASEAN, sedangkan 8,88% dipaparkan oleh faktor lain yang tidak termasuk dalam model.

## DAFTAR PUSTAKA

- Alwi, W., Rayyani, I., & Nurfadilah. 2018. Analisis Regresi Data Panel pada Faktor-Faktor yang Mempengaruhi Tingkat Kemiskinan Provinsi Sulawesi Selatan Tahun 2011-2015. *Jurnal Matematika dan Statistika Serta Aplikasinya*. **6**(2): 30-44.
- Armstrong, J. S. 2012. Illusions in Regression Analysis. *Internasional Journal of Forecasting*. **28**: 689-694.
- Baltagi, B. H. 2005. *Econometric Analysis of Panel Data (3rd ed.)*. England: John Wiley & Sons.
- Candra, K. A. 2018. Analisis Pengaruh Pertumbuhan Ekonomi dan Penanaman Modal Asing terhadap Emisi Karbondioksida di Delapan Negara ASEAN Periode 2004-2013. *Jurnal Ilmiah Mahasiswa Universitas Surabaya*. **7**(1): 2464-2661.
- Draper, N., & Smith, H. 1992. *Analisis Regresi Terapan*. Jakarta: Gramedia Pustaka Utama
- Eurostat. 2018. Glossary: Primary Energy Consumption. [https://ec.europa.eu/eurostat/statisticsexplained/index.php?title=Glossary:Primary\\_energy\\_consumption](https://ec.europa.eu/eurostat/statisticsexplained/index.php?title=Glossary:Primary_energy_consumption). Diakses pada 1 November 2023.
- Fauzi, R. 2017. Pengaruh Konsumsi Energi, Luas Kawasan Hutan dan Pertumbuhan Ekonomi terhadap Emisi CO<sub>2</sub> di 6 Negara ASEAN: Pendekatan Analisis Data Panel. *ECOLAB*. **11**(1): 14-26.
- Ghozali, I. 2016. *Aplikasi Analisis Multivariate SPSS 23*. Semarang: Badan Penerbit Universitas Diponegoro.

- Greene, W. H. 2012. *Econometrics Analysis (7th ed.)*. New York: Pearson.
- Gujarati, D. N. 2003. *Basic Econometrics (4th ed.)*. New York: McGraw-Hill.
- Halim, I. T., Ramadhanty, A. P., & Oscarini, D. R. 2020. Faktor-Faktor yang Mempengaruhi Persentase Penduduk Miskin di Indonesia Tahun 2015-2018 menggunakan Regresi Data Panel. *JURNAL EMACS*. 2(2): 55-63.
- Hamid, R. S., Bachri, S., Salju & Iqbal, M. 2020. *Ekonometrika Konsep Dasar dan Penerapan menggunakan Eviews 10*. Banten: CV. AA. Rizky.
- Jaya, I. G., & Sunengsih, N. 2009. Kajian Analisis Regresi dengan Data Panel. *Prosiding Seminar Nasional Penelitian, Pendidikan dan Penerapan MIPA* (hal. 51-58). Yogyakarta: Universitas Negeri Yogyakarta.
- Juanda, B., & Junaidi. 2012. *Ekonometrika Deret Waktu: Teori dan Aplikasi*. Bogor: IPB Press.
- Madany, N., Ruliana, & Rais, Z. 2022. Regresi Data Panel dan Aplikasinya dalam Kinerja Keuangan terhadap Pertumbuhan Laba Perusahaan Ixq45 Bursa Efek Indonesia. *Journal of Statistics and Its Application on Teaching and Research*. 4(2): 79-94.
- Nachrowi, D., & Usman, H. 2006. *Pendekatan Populer dan Praktis Ekonometrika untuk Analisis Ekonomi dan Keuangan*. Jakarta: Lembaga Penerbit FE UI.
- OECD. 2011. *Asia Tenggara: Pertumbuhan Tetap Kokoh dalam Jangka Menengah 5,6% pada 2012-2016*. <http://www.oecd.org/dev/49136551>. diakses pada 15 Oktober 2023.
- Rosadi, D. 2011. *Ekonometrika & Analisis Runtun Waktu Terapan dengan R*. Yogyakarta: C.V. Andi Offset.
- Sukadri, D. S. 2012. *REDD+ dan LULUCF: Panduan untuk Negosiator*. Jakarta: UNEP.

- Sulistiyowati, W., & Astuti, C. C. 2017. *Statistika Dasar Konsep dan Aplikasinya*. Sidoarjo: UMSIDA .
- Sumodiningrat, G. 2007. *Ekonometrika Pengantar (1st ed)*. Yogyakarta: BPFE Fakultas Ekonomi UGM.
- Wahyudi, S. T. 2020. *Konsep dan Penerapan Ekonometrika Menggunakan E-Views Edisi Kedua*. Depok: Rajawali Pers.
- Widyawati, R. F., Hariani, E., Ginting, A. L., & nainggolan, E. 2021. Pengaruh Pertumbuhan Ekonomi, Populasi Penduduk Kota, Keterbukaan Perdagangan Internasional terhadap Emis Karbon Dioksida (CO<sub>2</sub>) di Negara ASEAN. *Jambura Agribusiness Journal*. **3**(1): 37-47.
- Widarjono, A. 2009. *Ekonometrika Pengantar dan Aplikasinya Edisi Ketiga*. Yogyakarta: EKONISIA.