

METODE REGRESI DATA PANEL UNTUK PERAMALAN KONSUMSI ENERGI DI INDONESIA

Mariska Srihardianti¹, Mustafid², Alan Prahutama³

¹Mahasiswa Jurusan Statistika FSM Universitas Diponegoro

^{2,3}Staff Pengajar Jurusan Statistika FSM Universitas Diponegoro

ABSTRACT

Panel data regression is a method that aims to model the effect of one or more predictor variables on the response variable, observed in some sectors of an object of research for a specific time period. To estimate the panel data regression model, there are three approaches, namely Common Effect Model (CEM), Fixed Effects Model (FEM) and Random Effects Model (REM). In estimating the parameters for each model, there are several methods that can be used based on the assumption of the structure residual variance-covariance matrix, that is Ordinary Least Square/Least Square Dummy Variable (OLS/LSDV), Weighted Least Square (WLS) dan Seemingly Unrelated Regression (SUR). This research aims to implement the panel data regression to analyze the effect of GDP on energy consumption in Indonesia for each sector. Panel data regression model that has been obtained then is used to predict the amount of energy consumption in Indonesia for each sector in 2015 and 2016 using trend analysis. The analysis showed that the panel data regression model corresponding to the data of energy consumption in Indonesia in 1990-2014 is Fixed Effect Model (FEM) with Cross-section SUR, with R^2 value is 0.975943. Forecasting results show energy consumption in Indonesia in 2015 and 2016 will increase to the household sector and transport. Whereas for industrial, commercial and others sectors will decline in 2015 and then increase in 2016.

Keywords : Panel Data, Fixed Effect Model, SUR, Trend Analysis, Energy Consumption

1. PENDAHULUAN

Dalam menjalankan setiap aktivitas perekonomian di berbagai sektor, tentunya energi merupakan hal penting yang sangat dibutuhkan. Terlihat dari total energi yang dikonsumsi di Indonesia pada kurun waktu 2000-2012 meningkat rata-rata 2,91% per tahun^[7]. Namun, disisi lain jumlah sumber daya energi yang tersedia terbatas sehingga apabila tidak dikelola dengan baik bisa terjadi krisis energi di Indonesia. Untuk itu perlu dilakukan peramalan dan pemodelan konsumsi energi di Indonesia sebagai bahan acuan untuk menerapkan kebijakan mengenai energi.

Permasalahan yang dihadapi adalah data konsumsi energi merupakan suatu variabel yang tersusun atas beberapa sub-variabel atau sektor-sektor, yaitu sektor rumah tangga, industri, komersial, transportasi dan lainnya. Sehingga, metode yang tepat untuk digunakan dalam memodelkan konsumsi energi pada masing-masing sektor secara bersama-sama adalah metode regresi data panel.

Analisis regresi data panel merupakan suatu metode yang digunakan untuk memodelkan pengaruh variabel prediktor terhadap variabel respon dalam beberapa sektor yang diamati dari suatu objek penelitian selama periode waktu tertentu. Selain itu, regresi data panel juga digunakan untuk melakukan peramalan variabel respon pada setiap sektor yang ada. Namun, untuk meramalkannya, perlu dilakukan peramalan terlebih dahulu untuk variabel prediktornya pada masing-masing sektor.

Penelitian ini bertujuan untuk menerapkan regresi data panel dalam menganalisis pengaruh dari pertumbuhan ekonomi terhadap konsumsi energi di Indonesia untuk masing-masing sektor.

2. TINJAUAN PUSTAKA

2.1 Regresi Data Panel

Model regresi data panel dinyatakan dalam bentuk persamaan (1)^[5].

$$y_{it} = \mathbf{x}'_{it} \boldsymbol{\beta} + \mathbf{Z}'_i \boldsymbol{\alpha} + \varepsilon_{it} \quad (1)$$

$$i = 1, \dots, K; t = 1, \dots, T$$

dimana i menunjukkan unit *cross-section* sejumlah K , sementara t menunjukkan waktu sejumlah T . Terdapat p variabel bebas pada \mathbf{x}_{it} , tidak termasuk dengan konstan. Efek spesifik individual adalah $\mathbf{Z}'_i \boldsymbol{\alpha}$ dimana \mathbf{Z}_i terdiri dari konstan dan efek spesifik individual, baik yang dapat diobservasi maupun tidak terobservasi. $\boldsymbol{\beta}$ adalah matriks slope berukuran $p \times 1$.

2.2 Struktur Umum Model

Dalam mengestimasi model regresi panel, metode yang akan digunakan sangat bergantung pada asumsi yang dibuat mengenai intersep, slope koefisien dan error^[6]. Ditinjau dari berbagai asumsi dan faktor pembentukannya, struktur model dibagi menjadi 3, yaitu *Pooled Regression*, *Fixed Effect* dan *Random Effect*^[5].

2.2.1 Pooled Regression

Model ini mengasumsikan bahwa \mathbf{Z}_i hanya terdiri dari konstan saja atau dapat diartikan bahwa tidak terdapat efek spesifik individual. Struktur model ini sering juga disebut dengan *Common Effect Model* (CEM) dan dinyatakan kedalam bentuk persamaan (2).

$$y_{it} = \alpha + \mathbf{x}'_{it} \boldsymbol{\beta} + \varepsilon_{it} \quad (2)$$

$$i = 1, \dots, K; t = 1, \dots, T$$

dimana α adalah koefisien intersep (konstan) yang mana merupakan bilangan skalar, $\boldsymbol{\beta}$ adalah matriks slope berukuran $p \times 1$ dan \mathbf{x}_{it} merupakan observasi ke- i dan waktu ke- t pada variabel penjelas p .

Dalam mengestimasi parameter model CEM, terdapat 4 metode estimasi yang dapat digunakan, yaitu *Ordinary Least Square* (OLS), jika bersifat homoskedastik dan tidak ada *cross-sectional correlation*; *Weighted Least Square* (WLS), jika bersifat heteroskedastik dan tidak ada *cross-sectional correlation*; *Seemingly Uncorrelated Regression* (SUR), jika bersifat heteroskedastik dan ada *cross-sectional correlation*; dan *Feasible Generalized Least Square* (FGLS) dengan proses *autoregressive* (AR), jika bersifat heteroskedastik dan ada korelasi antar waktu pada residualnya.

2.2.2 Fixed Effect

Struktur model *fixed effect* merupakan model yang memperhatikan adanya keberagaman dari variabel independen menurut individu. Model *Fixed Effect* dinyatakan kedalam bentuk persamaan (3).

$$y_{it} = \mathbf{x}'_{it} \boldsymbol{\beta} + \alpha_i + \varepsilon_{it} \quad (3)$$

dimana $\alpha_i = \mathbf{Z}'_i \boldsymbol{\alpha}$, mewujudkan semua efek yang diamati dan menspesifikasikan rata-rata kondisional yang dapat diestimasi. α_i diperlakukan sebagai parameter tetap yang tidak diketahui dan akan di estimasi. \mathbf{Z}_i diasumsikan tidak terobservasi dan memiliki korelasi dengan variabel independen. ε_{it} adalah error yang bersifat stokastik dan terdistribusi secara independen dan identik dengan rata-rata 0 dan varian σ_ε^2 . Variabel independen \mathbf{x}_{it} diasumsikan independen dengan error ε_{it} untuk semua i dan t .

Dalam mengestimasi parameter model *fixed effects* terdapat 3 metode estimasi yang dapat digunakan berdasarkan pada asumsi struktur matriks varians-kovarians residualnya. Metode estimasi parameter tersebut sama dengan yang digunakan dalam model CEM,

namun tidak termasuk dengan *Feasible Generalized Least Square* (FGLS) dengan proses *autoregressive* (AR).

2.2.3 Random Effect

Apabila efek individual Z_i tidak memiliki korelasi dengan variabel independen, maka struktur model ini dikenal dengan *Random Effect Model* yang modelnya dinyatakan kedalam bentuk persamaan (4).

$$y_{it} = \mathbf{x}'_{it} \boldsymbol{\beta} + \alpha + \eta_{it} \quad (4)$$

dimana

$$\begin{aligned} \eta_{it} &= u_i + \varepsilon_{it}; \\ \alpha &= E[\mathbf{Z}'_i \boldsymbol{\alpha}]; \\ u_i &= \{\mathbf{Z}'_i \boldsymbol{\alpha} - E[\mathbf{Z}'_i \boldsymbol{\alpha}]\}; \end{aligned}$$

Terdapat sejumlah P variabel independen termasuk dengan konstan. α merupakan rata-rata dari efek individual yang tidak terobservasi. u_i merupakan efek random spesifik untuk observasi ke-i. Dalam model ini u_i diasumsikan independen dengan ε_{it} , selain itu diasumsikan pula bahwa variabel independen x_{it} independen dengan u_i dan ε_{it} .

Apabila dilihat dari struktur matriks varian kovarian residualnya, terlihat bahwa komponen error dari model bersifat homokedastik dan terdapat korelasi antar waktu antara error dengan sektor yang sama (*equicorrelated*)^[1]. Dengan demikian, metode yang tepat untuk mengestimasi model *random effects* adalah *Generalized Least Squares (GLS)* dengan asumsi homokedastik dan tidak ada *cross-sectional correlation*^[4].

2.3 Penyeleksian Model

2.3.1 Uji Chow

Uji Chow digunakan untuk mengetahui apakah model FEM lebih baik dari model CEM. Berikut merupakan statistik uji yang digunakan:

$$F \text{ hitung} = \frac{(RSS_1 - RSS_2)/(K - 1)}{RSS_2/(KT - K - P)} \sim F_{(\alpha, (K-1), (KT-K-P))}$$

dimana K adalah jumlah sektor; T adalah periode waktu observasi; P adalah jumlah parameter dalam model *fixed effects*; RSS_1 merupakan *residual sum of squares common effects model*, sedangkan RSS_2 merupakan *residual sum of squares fixed effects model*.

Jika nilai statistik F hitung lebih besar daripada F tabel ($F_{(\alpha, (K-1), (KT-K-P))}$) pada α tertentu, maka model yang terpilih adalah model FEM.

2.3.2 Uji Lagrange Multiplier

Uji *Lagrange Multiplier* digunakan untuk pengujian *random effect* yang didasarkan pada nilai residual dari model *common effect*^[2]. Adapun nilai statistik uji LM dapat dihitung menggunakan rumus sebagai berikut:

$$LM = \frac{KT}{2(T-1)} \left[\frac{\sum_{i=1}^K [\sum_{t=1}^T e_{it}]^2}{\sum_{i=1}^K \sum_{t=1}^T e_{it}^2} - 1 \right]^2 \sim \chi^2_{\alpha, 1}$$

dimana K adalah jumlah sektor, T adalah jumlah periode waktu dan e_{it} adalah residual model *common effects*. Apabila nilai $LM > \chi^2_{\alpha, 1}$, maka model yang terpilih adalah model REM.

2.3.3 Uji Hausman

Uji *Hausman* bertujuan untuk memilih antara model FEM dan model REM. Dengan mengikuti kriteria Wald, nilai statistik Hausman dapat dihitung dengan rumusan sebagai berikut:

$$W = \chi^2_{(P)} = [\mathbf{b} - \boldsymbol{\beta}]' \boldsymbol{\psi}^{-1} [\mathbf{b} - \boldsymbol{\beta}]$$

dimana

$$\boldsymbol{\psi} = \text{Var}[\mathbf{b}] - \text{Var}[\boldsymbol{\beta}]$$

b adalah parameter (tanpa intersep) *random effect* dan β adalah parameter *fixed effect* menggunakan LSDV. $\text{Var}[b]$ merupakan matriks kovarian parameter (tanpa intersep) *random effect* dan $\text{Var}[\beta]$ adalah matriks kovarian parameter *fixed effect*.

Apabila nilai $W > \chi^2_{(\alpha,P)}$, maka model yang terpilih adalah model FEM. P adalah jumlah variabel independen.

2.4 Pengujian Asumsi Struktur Varians-kovarians Residual

Pengujian ini bertujuan untuk mengetahui apakah struktur varians-kovarians residual memenuhi asumsi struktur homoskedastik atau heteroskedastik. Nilai statistik uji LM dirumuskan sebagai berikut:

$$LM = \frac{T}{2} \sum_{i=1}^K \left[\frac{\sigma_i^2}{\sigma^2} - 1 \right]^2 \sim \chi^2_{\alpha;K-1}$$

dengan T adalah jumlah periode waktu, K adalah jumlah unit *cross-section*, σ_i^2 adalah varian residual persamaan ke-i pada kondisi homoskedastik, dan σ^2 adalah *Mean Square Error* persamaan sistem pada kondisi homoskedastik. Apabila nilai statistik $LM > \chi^2_{\alpha;K-1}$, maka struktur varians-kovarians residual heteroskedastik.

2.5 Pengujian Asumsi Adanya *Cross-Sectional Correlation*

Apabila didapat kesimpulan bahwa struktur varians-kovarians residual bersifat heteroskedastik, maka selanjutnya dilakukan uji LM untuk mengetahui apakah terdapat *cross sectional correlation* atau tidak pada struktur varians-covarians tersebut. Statistik uji yang digunakan dirumuskan sebagai berikut:

$$\lambda_{LM} = T \sum_{i=2}^K \sum_{j=1}^{i-1} r_{ij}^2$$

dimana K adalah jumlah unit *cross-section*, T adalah jumlah periode waktu, r_{ij}^2 adalah *residual correlation coefficient* antara persamaan ke-i dan ke-j.

Apabila nilai statistik $\lambda_{LM} > \chi^2_{\alpha;K(K-1)/2}$, maka struktur varians-kovarians residual bersifat heteroskedastik dan ada *cross-sectional correlation*.

2.6 *Seemingly Uncorrelated Regression*

Sebuah sistem *Seemingly Unrelated Regression* (SUR) terdiri dari beberapa persamaan yang errornya saling berkorelasi. Dimisalkan bahwa terdapat sebuah data set yang terdiri dari K unit *cross-section* dengan observasi sejumlah T periode waktu untuk setiap unit. Model SUR adalah sebuah sistem dari persamaan regresi linier yang dapat dituliskan menjadi:

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_K \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 & 0 & \dots & 0 \\ 0 & \mathbf{X}_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & \mathbf{X}_K \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \vdots \\ \boldsymbol{\beta}_K \end{bmatrix} + \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \vdots \\ \mathbf{u}_K \end{bmatrix}$$

atau

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$$

Dimana \mathbf{y}_i dan \mathbf{u}_i adalah vektor berdimensi T. \mathbf{X}_i adalah matriks berukuran T x P_i . $\boldsymbol{\beta}_i$ adalah vektor berdimensi P_i . $\boldsymbol{\beta}$ adalah vektor berdimensi P dari parameter yang belum diketahui dan harus diestimasi dan $P = \sum_{i=1}^K P_i$. \mathbf{u} adalah vektor berukuran $KT \times 1$ dari error. Matriks varian kovarian dari \mathbf{u} diberikan sebagai berikut:

$$\boldsymbol{\Omega} = \boldsymbol{\Sigma} \otimes \mathbf{I}_T = \begin{bmatrix} \sigma_{11}\mathbf{I}_T & \sigma_{12}\mathbf{I}_T & \dots & \sigma_{1K}\mathbf{I}_T \\ \sigma_{21}\mathbf{I}_T & \sigma_{22}\mathbf{I}_T & \dots & \sigma_{2K}\mathbf{I}_T \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{K1}\mathbf{I}_T & \sigma_{K2}\mathbf{I}_T & \dots & \sigma_{KK}\mathbf{I}_T \end{bmatrix}$$

Metode *Feasible Generalized Least Square* (FGLS) digunakan untuk mengestimasi parameter yang ada pada model SUR. Berikut adalah estimasi parameter β dengan menggunakan FGLS:

$$\hat{\beta} = [X'\Omega^{-1}X]^{-1}X'\Omega^{-1}y$$

dengan matriks kovarian sebagai berikut:

$$\text{var}[\beta] = [X'\Omega^{-1}X]^{-1}$$

2.7 Uji Asumsi

Pada regresi data panel, estimator yang digunakan adalah *Feasible Generalized Least Square* (FGLS) yang memiliki 2 tahap estimasi yang berbeda dengan OLS. Sehingga pengujian asumsi klasik, seperti homoskedastisitas dan non-autokorelasi, tidak dapat digunakan dalam estimator ini karena tidak lagi relevan dengan konsep pengujian asumsi klasik pada regresi OLS. Namun, pengujian multikolinieritas dan normalitas akan tetap dilakukan^[3]. Pemeriksaan asumsi normalitas ini dapat dilakukan menggunakan uji *Jarque-Bera* dengan statistik uji sebagai berikut:

$$JB = N \left[\frac{S_k^2}{6} + \frac{(K_u - 3)^2}{24} \right]$$

dimana N adalah banyaknya data, S_k adalah *skewness* (kemencengan) dan K_u adalah *kurtosis* (keruncingan). Residual akan berdistribusi normal apabila nilai statistik uji $JB < \chi^2_{(\alpha,2)}$.

2.8 Uji Signifikansi Parameter

2.8.1 Uji Serentak (Uji F)

Uji F dilakukan untuk menguji hasil estimasi model regresi apakah variabel independen secara bersama-sama memiliki pengaruh terhadap variabel dependen. Statistik uji F dirumuskan sebagai berikut:

$$F_{\text{hitung}} = \frac{\text{Mean Square Regresi}}{\text{Mean Square Residual}}$$

Variabel independen secara bersama-sama berpengaruh signifikan terhadap variabel dependen jika nilai $F_{\text{hitung}} > F_{\alpha;(K+P-1,KT-K-P)}$. Dimana K adalah jumlah unit *Cross-Section*, T adalah jumlah periode waktu dan P adalah jumlah variabel independen.

2.8.2 Uji Parsial (Uji t)

Uji t bertujuan untuk melihat signifikansi dari pengaruh variabel independen secara individu terhadap variabel dependen dengan menganggap variabel lain bersifat konstan. Statistik uji yang digunakan dirumuskan sebagai berikut:

$$t_{\text{hitung}} = \frac{\hat{\beta}_k}{SE(\hat{\beta}_k)}$$

$\hat{\beta}_k$ adalah penduga parameter ke-i dan $SE(\hat{\beta}_k)$ adalah simpangan baku dari nilai penduga dari parameter ke-i. Variabel independen memiliki pengaruh yang signifikan terhadap variabel dependen jika nilai $|t_{\text{hitung}}| > t_{(\frac{\alpha}{2},KT-K-P)}$.

2.9 Peramalan Menggunakan Model Regresi Data Panel

Model regresi data panel juga dapat digunakan untuk melakukan peramalan variabel respon. Namun, peramalan variabel respon untuk setiap sektor dalam beberapa tahun mendatang hanya dapat dilakukan apabila nilai dari variabel prediktor untuk setiap sektor selama tahun tersebut diketahui. Dengan demikian, akan dilakukan peramalan variabel prediktor terlebih dahulu untuk beberapa tahun kedepan di setiap sektornya menggunakan Analisis Trend Linier, Trend Kuadratik dan Trend Eksponensial. Hasil peramalan yang terbaik adalah peramalan dengan metode yang memiliki nilai MAPE, MAD dan MSD yang paling kecil. Setelah peramalan dilakukan, hasil peramalan variabel

prediktor disubstitusikan ke dalam persamaan regresi data panel sehingga diperoleh nilai ramalan untuk variabel respon.

2.10 Pemilihan Hasil Peramalan

2.10.1 MAPE (*Mean Absolute Percent Error*)

MAPE dinyatakan dalam bentuk persentase dan dihitung menggunakan rumus sebagai berikut:

$$\text{MAPE} = \frac{\sum_{i=1}^K \sum_{t=1}^T \left| \frac{Y_{it} - \widehat{Y}_{it}}{Y_{it}} \right|}{n} \times 100$$

dimana Y_{it} adalah nilai data asli pada sektor ke- i waktu ke- t , \widehat{Y}_{it} adalah nilai prediksi pada sektor ke- i waktu ke- t dan n adalah jumlah observasi atau KT.

2.10.2 MAD (*Mean Absolute Deviation*)

MAD dirumuskan sebagai berikut:

$$\text{MAD} = \frac{\sum_{i=1}^K \sum_{t=1}^T |Y_{it} - \widehat{Y}_{it}|}{n}$$

dimana Y_{it} adalah nilai data asli pada sektor ke- i waktu ke- t , \widehat{Y}_{it} adalah nilai prediksi pada sektor ke- i waktu ke- t dan n adalah jumlah observasi atau KT.

2.10.3 MSD (*Mean Squared Deviation*)

MSD adalah ukuran yang lebih sensitif dari kesalahan perkiraan yang luar biasa besar dari MAD. MSD dapat dihitung menggunakan rumus sebagai berikut:

$$\text{MSD} = \frac{\sum_{i=1}^K \sum_{t=1}^T |Y_{it} - \widehat{Y}_{it}|^2}{n}$$

dimana Y_{it} adalah nilai data asli pada sektor ke- i waktu ke- t , \widehat{Y}_{it} adalah nilai prediksi pada sektor ke- i waktu ke- t dan n adalah jumlah observasi atau KT.

3. METODE PENELITIAN

3.1 Jenis dan Sumber Data

Data yang digunakan dalam penelitian ini adalah Total Konsumsi Energi Akhir di Indonesia dan Produk Domestik Bruto (PDB) di Indonesia atas dasar harga konstan menurut lapangan usaha. Kedua data ini diambil dari tahun 1990 hingga 2014 dan disusun kedalam 5 sektor, yaitu rumah tangga, industri, transportasi, komersial dan lainnya. Data konsumsi energi diperoleh dari buku Handbook of Energy and Economic Statistics of Indonesia 2012 dan 2015. Sedangkan untuk data PDB diperoleh dari Badan Pusat Statistik Indonesia.

3.2 Variabel Penelitian

Variabel yang digunakan dalam penelitian ini adalah sebagai berikut:

- a. Total konsumsi energi Indonesia (ribuan SBM) sebagai variabel respon (Y)
- b. Produk Domestik Bruto atas dasar harga konstan (miliar rupiah) sebagai variabel prediktor (X)

3.3 Langkah-langkah Analisis Data

Metode analisis data yang digunakan dalam penelitian ini adalah peramalan dengan menggunakan analisis regresi data panel. Berikut langkah-langkah yang dilakukan untuk menganalisis data dalam penelitian ini:

1. Menentukan model regresi data panel yang terbaik untuk memodelkan pengaruh pertumbuhan ekonomi (PDB) terhadap konsumsi energi di Indonesia.
 - a. Mengestimasi model *Common Effect*, model *Fixed Effect* dan model *Random Effect*.
 - b. Menentukan model terbaik melalui Uji Chow, uji *Langrange Multiplier* (LM) dan uji *Hausman*.

- c. Menentukan metode estimasi parameter yang tepat dengan melihat bagaimana struktur varians-covarian dan korelasi antar sektornya dari residualnya menggunakan uji LM (*Langrange Multiplier*). Apabila model yang terpilih adalah model REM, maka tidak perlu dilakukan pengujian untuk menentukan metode estimasi yang terbaik.
 - d. Melakukan uji asumsi normalitas residual dengan menggunakan uji Jarque-Bera.
 - e. Melakukan uji signifikansi parameter regresi data panel yang meliputi Uji Serentak (Uji F) dan Uji Parsial (Uji t).
2. Melakukan peramalan untuk variabel respon berdasarkan pada model regresi data panel yang telah diperoleh untuk beberapa periode mendatang.
 - a. Melakukan peramalan untuk masing-masing variabel prediktor dengan menggunakan metode analisis trend linear, kuadratik dan eksponensial.
 - b. Menentukan hasil peramalan yang terbaik dari ketiga metode berdasarkan nilai MAPE, MAD dan MSD yang paling kecil.
 - c. Mensubstitusikan hasil peramalan variabel prediktor pada masing-masing model regresi data panel yang telah diperoleh untuk mendapatkan hasil peramalan dari variabel respon.

4. HASIL DAN PEMBAHASAN

4.1 Estimasi Model Regresi Data Panel

Pemodelan yang dilakukan dalam penelitian ini menggunakan data Konsumsi Energi di Indonesia sebagai variabel dependen dan data Produk Domestik Bruto sebagai variabel independen.

4.1.1 Model *Common Effect*

Pemodelan menggunakan *Common Effect Model* dilakukan dengan menggunakan bantuan program Eviews 5.1 dengan hasil estimasi sebagai berikut:

$$\hat{y}_{it} = 124.536,4 + 0,058051 x_{it} \quad (1)$$

4.1.2 Model *Fixed Effect*

Pemodelan menggunakan *Fixed Effect Model* dilakukan dengan menggunakan bantuan program Eviews 5.1 dengan hasil estimasi model sebagai berikut:

$$\hat{y}_{it} = \hat{\alpha}_i + 0,501887 x_{it} \quad (2)$$

dimana besaran nilai intersep $\hat{\alpha}_i$ berbeda-beda untuk setiap sektor yang tersaji dalam Tabel 1.

Tabel 1. Estimasi intersep $\hat{\alpha}_i$ Model FEM

Indeks (i)	Sektor	$\hat{\alpha}_i$
1	Industri	12.434,44
2	Rumah Tangga	153.506,01
3	Komersial	-208.011,69
4	Transportasi	98.994,51
5	Lainnya	-215.390,79

4.1.3 Model *Random Effect*

Pemodelan menggunakan *Random Effect Model* dilakukan dengan menggunakan bantuan program Eviews 5.1 dengan hasil estimasi model sebagai berikut:

$$\hat{y}_{it} = \hat{\alpha}_i + 0,455791x_{it} \quad (3)$$

dimana besaran nilai intersep $\hat{\alpha}_i$ berbeda-beda untuk setiap sektor yang tersaji dalam Tabel 2.

Tabel 2. Estimasi intersep $\hat{\alpha}_i$ Model REM

Indeks (i)	Sektor	$\hat{\alpha}_i$
1	Industri	37.080,4
2	Rumah Tangga	150.888,03
3	Komersial	-172.197,77
4	Transportasi	99.214,23
5	Lainnya	-174.866,67

4.2 Pemilihan Model Regresi Data Panel

4.2.1 Pemilihan Model CEM dan FEM dengan Uji Chow

Perhitungan Uji Chow dilakukan dengan bantuan program Eviews 5.1 dan diperoleh nilai $F_{hitung} = 49,101073 > F_{(0,05;4;119)} = 2,44788$, sehingga H_0 ditolak. Dengan demikian, dapat disimpulkan bahwa terdapat efek individu pada model persamaan konsumsi energi di Indonesia, sehingga model yang sesuai adalah model *Fixed Effect* (FEM).

Hasil dari pengujian ini menunjukkan bahwa pada pemodelan konsumsi energi terdapat efek spesifik individu yang tidak terobservasi dan mempengaruhi model. Ada tidaknya efek spesifik individu dapat dilakukan dengan melihat heterogenitas antar sektor pada data konsumsi energi dan PDB. Pengujian ini dilakukan menggunakan Uji Levene. Berdasarkan hasil olahan Minitab 16 diperoleh nilai statistik uji W untuk konsumsi energi adalah 3,9 dengan p-value 0,005 dan nilai statistik uji W sebesar 12,81 dengan p-value 0,000 untuk PDB. Dari nilai-nilai tersebut dapat diambil kesimpulan bahwa baik variabel konsumsi energi maupun PDB memiliki varian antar sektor yang berbeda-beda, karena memiliki nilai p-value $< 0,05$.

Menurut peneliti, terjadinya heterogenitas antar sektor pada data konsumsi energi dan PDB mengindikasikan bahwa masing-masing sektor memiliki pengaruh tersendiri yang tidak terlihat dan mempengaruhi model regresi. Untuk itu, pemodelan menggunakan CEM yang mengasumsikan tidak adanya efek spesifik dari masing-masing sektor dirasa kurang tepat karena asumsi tersebut telah dilanggar. Hal ini sesuai dengan hasil Uji Chow yang mengatakan bahwa model *Fixed Effect* adalah model yang lebih baik dari pada model *Common Effect*. Karena model estimasi yang terpilih adalah model FEM, maka pengujian selanjutnya adalah Uji Hausman. Uji *Lagrange Multiplier* tidak perlu untuk dilakukan.

4.2.2 Pemilihan Model FEM dan REM

Dalam menentukan model yang lebih sesuai antara model FEM atau REM, pengujian yang digunakan adalah uji Hausman. Perhitungan Uji Hausman dilakukan dengan program Eviews 5.1 dan diperoleh nilai W sebesar 13,783173. Nilai $W > \chi^2_{(0,05;1)} = 3,841$, sehingga dapat disimpulkan bahwa model yang sesuai adalah model *Fixed Effect* (FEM).

Model FEM yang lebih baik dapat disebabkan karena adanya korelasi antara residual model REM dengan variabel PDB. Dengan menggunakan program Minitab 16, diperoleh besaran korelasi pearson antara error dari model REM dengan variabel PDB sebesar -0,668 dengan P-Value 0,000. Nilai P-Value $< \alpha = 0,05$ menunjukkan bahwa terdapat korelasi antara residual model REM dengan variabel PDB. Dengan demikian, asumsi untuk model REM telah dilanggar dengan adanya korelasi ini yang menjadikan model REM tidak tepat untuk digunakan. Sehingga, estimasi model regresi yang tepat untuk data konsumsi energi di Indonesia tahun 1990-2014 adalah menggunakan *Fixed Effect Model*.

4.3 Pengujian Struktur Varian-Covarian Residual

4.3.1 Heteroskedastisitas antar Sektor

Pengujian yang digunakan untuk mengetahui apakah struktur varians-kovarians residual dari model *Fixed Effect* mengalami homoskedastik atau heteroskedastik adalah Uji *Lagrange Multiplier* (Uji LM). Berdasarkan hasil perhitungan diperoleh nilai statistik uji

LM sebesar 57,0308447. Terlihat bahwa nilai $LM > \chi_{0,05;4}^2 = 9,487$, sehingga dapat disimpulkan bahwa struktur varian-covarian residual dari model FEM adalah bersifat heterokedastik.

4.3.2 Cross-Sectional Correlation

Pengujian yang digunakan untuk mengetahui adanya *cross-sectional correlation* atau tidak pada struktur varians-covarians residual model FEM yang bersifat heterokedastik adalah Uji Lagrange Multiplier (Uji LM). Diperoleh nilai statistik uji $\lambda_{LM} = 92,1467 > \chi_{0,05;10}^2 = 18,307$, sehingga struktur varian-covarian residual dari model FEM adalah bersifat heterokedastik dan terdapat *Cross-Sectional Correlation*. Dengan struktur varian-covarian residual model FEM yang seperti ini, maka metode yang paling tepat untuk digunakan dalam mengestimasi parameter model *fixed effect* adalah *Seemingly Unrelated Regression* (SUR).

4.4 Estimasi Model Akhir Regresi Data Panel

Pemodelan menggunakan model FEM dengan *Cross-section* SUR akan dilakukan dengan menggunakan bantuan program Eviews 5.1 dengan hasil estimasi model sebagai berikut:

$$\hat{y}_{it} = \hat{\alpha}_i + 0,501887 x_{it} \quad (4)$$

Besaran nilai intersep $\hat{\alpha}_i$ berbeda-beda untuk setiap sektor seperti yang disajikan dalam Tabel 3.

Tabel 3. Estimasi intersep $\hat{\alpha}_i$ Model FEM *Cross-Section* SUR

Indeks (i)	Sektor	$\hat{\alpha}_i$
1	Industri	17.650,2156
2	Rumah Tangga	154.135,4899
3	Komersial	-202.062,5764
4	Transportasi	99.820,8949
5	Lainnya	-208.529,7316

4.5 Uji Asumsi Normalitas Residual

Pengujian asumsi normalitas dilakukan dengan Uji Jarque-Bera. Berdasarkan hasil pengolahan menggunakan Eviews 5.1 diperoleh nilai statistik uji JB sebesar 0,992542. Terlihat bahwa nilai $JB < \chi_{(0,05;2)}^2 = 5,9914$, sehingga dapat disimpulkan bahwa residual dari model FEM *cross-section* SUR mengikuti distribusi normal.

4.6 Uji Signifikansi Parameter

4.6.1 Uji Serentak (Uji F)

Pengujian ini dilakukan untuk menguji estimasi dari model FEM *cross-section* SUR apakah variabel independen secara bersama-sama memiliki pengaruh terhadap variabel dependen. Berdasarkan hasil pengolahan menggunakan program Eviews 5.1 diperoleh nilai F_{hitung} sebesar 965,5314 dengan $Prob(F\text{-statistic}) < 0,000$. Terlihat bahwa nilai $F_{hitung} > F_{0,05;5;119} = 2,2905$ dan $Prob(F\text{-statistic}) < 0,05$, sehingga dapat disimpulkan bahwa variabel PDB secara bersama-sama berpengaruh signifikan terhadap konsumsi energi.

4.6.2 Uji Parsial (Uji t)

Uji t bertujuan untuk melihat signifikansi dari pengaruh variabel independen secara individu terhadap variabel dependen dengan menganggap variabel lain bersifat konstan. Berdasarkan hasil pengolahan menggunakan Eviews 5.1, diperoleh nilai $|t_{hitung}|$ untuk variabel PDB dan intersep lebih besar dari nilai $t_{(0,025;119)}$, yaitu 1,9801. Sehingga dapat disimpulkan bahwa intersep model dan variabel PDB memiliki pengaruh yang signifikan terhadap variabel konsumsi energi di Indonesia.

4.7 Peramalan Konsumsi Energi di Indonesia

4.7.1 Peramalan PDB dengan Analisis Trend

Peramalan variabel PDB untuk masing-masing sektor akan dilakukan dengan menggunakan Analisis Trend Linier, Trend Kuadratik dan Trend Eksponensial. Pemilihan metode yang paling baik akan didasarkan pada model dengan nilai MAPE, MAD dan MSD yang paling kecil. Berdasarkan hasil pengolahan dengan program Minitab 16, diperoleh bahwa untuk semua sektor yang ada, analisis trend kuadratik merupakan metode yang paling baik karena memiliki nilai MAPE, MAD dan MSD yang paling kecil. Dengan demikian, kemudian dilakukan *forecasting* atau peramalan untuk variabel PDB tahun 2015 dan 2016 masing-masing sektor menggunakan analisis trend kuadratik. Model persamaan trend kuadratik yang akan digunakan untuk peramalan dan hasil peramalan variabel PDB tahun 2015 dan 2016 untuk setiap sektornya disajikan dalam Tabel 5.

Tabel 5. Model Trend Kuadratik dan Hasil Peramalan Variabel PDB Tahun 2015 dan 2016

Sektor	Model Trend Kuadratik	PDB	
		2015	2016
Rumah Tangga	$X_t = 29.179 - 719 t + 161,2 t^2$	119.429	127.252
Industri	$X_t = 197.462 + 12.837 t + 430,3 t^2$	822.124	857.768
Komersial	$X_t = 336.590 - 2.573 t + 1.000 t^2$	945.377	995.779
Transportasi	$X_t = 69.201 - 2.064 t + 137,6 t^2$	108.590	113.822
Lainnya	$X_t = 481.486 - 10.543 t + 1.176 t^2$	1.002.513	1.054.311

4.7.2 Peramalan Konsumsi Energi dengan Regresi Data Panel

Untuk mendapatkan nilai prediksi dari variabel konsumsi energi di Indonesia pada tahun 2015 dan 2016, langkah selanjutnya yang dilakukan adalah mensubstitusi hasil peramalan variabel PDB pada tabel 5 kedalam model FEM dengan *Cross-section* SUR yang telah diperoleh.

Tabel 6. Peramalan Konsumsi Energi

Sektor	Model Regresi Data Panel	Konsumsi Energi	
		2015	2016
Rumah Tangga	$\hat{y}_t = 154.135,4899 + 0,501887 x_t$	214.075,3043	218.001,5631
Industri	$\hat{y}_t = 17.650,2156 + 0,501887 x_t$	430.263,2321	448.152,4780
Komersial	$\hat{y}_t = -202.062,5764 + 0,501887 x_t$	272.409,4688	297.705,5571
Transportasi	$\hat{y}_t = 99.820,8949 + 0,501887 x_t$	154.320,7605	156.946,6312
Lainnya	$\hat{y}_t = -208.529,7316 + 0,501887 x_t$	294.618,1062	320.614,8282

Berdasarkan pada hasil peramalan yang telah dilakukan, terlihat bahwa konsumsi energi di Indonesia pada tahun 2015 dan 2016 akan semakin meningkat untuk sektor rumah tangga dan transportasi. Sedangkan untuk sektor industri, komersial dan lainnya akan mengalami penurunan pada tahun 2015 dan kembali meningkat pada tahun 2016.

5. KESIMPULAN

Dalam melakukan pemodelan menggunakan regresi data panel, terdapat tiga metode estimasi model yang dapat digunakan, yaitu Model *Common Effect*, Model *Fixed Effect* dan Model *Random Effect*. Ketiga model ini memiliki asumsi-asumsi tersendiri yang harus terpenuhi agar diperoleh estimasi model yang tepat.

Dalam menentukan model estimasi regresi yang paling tepat dari ketiga model yang tersedia, dilakukan serangkaian pengujian yaitu Uji Chow yang digunakan untuk memilih antara model CEM dan FEM; uji *Langrange Multiplier* (LM) yang digunakan untuk memilih antara model CEM dan REM; dan uji *Hausman* yang digunakan untuk memilih antara model FEM dan REM. Selain pengujian diatas, dapat dilakukan pengujian tambahan untuk menguji asumsi-asumsi dari ketiga model. Model yang terbaik adalah model yang memenuhi semua asumsi yang disyaratkan pada masing-masing modelnya, memiliki nilai R^2 yang paling tinggi dan modelnya valid untuk digunakan berdasarkan pada Uji F.

Model regresi data panel yang terbaik untuk data Konsumsi Energi di Indonesia sebagai variabel dependen dan data Produk Domestik Bruto sebagai variabel independen adalah model *Fixed Effect Model* dengan *Cross-section SUR*. Pengolahan model ini dilakukan dengan menggunakan program Eviews 5.1 dengan hasil estimasi model sebagai berikut:

$$\hat{y}_{it} = \hat{\alpha}_i + 0,501887 x_{it}$$

Dimana besaran nilai intersep $\hat{\alpha}_i$ berbeda-beda untuk setiap sektor yang tercantum pada Tabel 3. Model ini memiliki nilai R^2 sebesar 0,975943, yang memberikan arti bahwa 97,5943% variasi dari variabel konsumsi energi dapat dijelaskan oleh variabel PDB. Berdasarkan pada model regresi diatas, dapat diinterpretasikan bahwa setiap bertambahnya PDB sebesar 1 milyar rupiah akan menaikkan total konsumsi energi akhir sebesar 501,887 SBM.

6. DAFTAR PUSTAKA

- [1] Baltagi, B.H. 2003. *A Companion to Theoretical Econometrics*. Blackwell Publishing Ltd.
- [2] Breusch, T.S., Pagan, A.R. 1980. *The Lagrange Multiplier Test and Its Applications to Model Specification in Econometrics*. Review of Economic Studies Vol. 47, No. 1 : Hal. 239-253.
- [3] Ekananda, M. 2016. *Analisis Ekonometrika Data Panel*. Edisi ke-2. Jakarta: Mitra Wacana Media.
- [4] Fadly, F. 2011. *Peran Pertumbuhan Ekonomi dan Intervensi Pemerintah di Bidang Fiskal terhadap Kemiskinan, Pengangguran dan Ketimpangan Distribusi Pendapatan di Indonesia Periode 2005 – 2008*. Skripsi Sarjana, Sekolah Tinggi Ilmu Statistik, Jakarta.
- [5] Greene, W.H. 2002. *Econometric Analysis*. Edisi ke-5. New York: Macmillan Publishing Company.
- [6] Hsiao, C. 2003. *Analysis of Panel Data*. Edisi ke-3. United State of America: Cambridge University Press.
- [7] [PTPSP] Pusat Teknologi Pengembangan Sumberdaya. 2014. *Outlook Energi Indonesia 2014: Pengembangan Energi dalam Mendukung Program Substitusi BBM*. Jakarta: PTPSP.