

## Estimasi Model *Fixed Effect* Pada Analisis Regresi Data Panel Dengan Metode *Least Square Dummy Variable* (LSDV)

Junia Rahma Nur Imani<sup>1\*</sup>, Khoirin Nisa<sup>2</sup>, Dorrah Aziz<sup>3</sup>, Nusyirwan<sup>4</sup>

<sup>1,2,3,4</sup>Jurusan Matematika, Fakultas MIPA, Universitas Lampung, Indonesia

\*corresponding author: [juniarahma017@gmail.com](mailto:juniarahma017@gmail.com)

Received November 26, 2024; Received in revised form January 05, 2025; Accepted January 09, 2025

**Abstrak.** Data panel merupakan gabungan antara data *cross section* dan data *time series*. Salah satu model analisis regresi data panel adalah model *fixed effect*. Model *fixed effect* mempunyai asumsi bahwa intersep berbeda untuk setiap individu, tetapi koefisien *slope* konstan. Estimasi dilakukan dengan menggunakan variabel *dummy* untuk menjelaskan adanya perbedaan intersep antar individu. Penelitian ini bertujuan untuk mengestimasi model *fixed effect* pada analisis regresi data panel dengan metode *least square dummy variable* dan menerapkannya pada data upah minimum provinsi di Indonesia tahun 2014-2017. Berdasarkan hasil penelitian yang telah dilakukan dengan menggunakan estimasi parameter  $\hat{Y}_{LSDV} = (\tilde{X}^T \tilde{X})^{-1} \tilde{X}^T Y$  untuk model *fixed effect* pada analisis regresi data panel upah minimum provinsi di Indonesia diperoleh model sebagai berikut,  $\hat{Y}_{it} = 5.248452 + \sum_{i=2}^N \beta_{0i} D_{ki} + 0.007415X_{1it} + 0.002882X_{2it} + 1.63E-07X_{3it}$ , dengan,  $\hat{Y}_{it}$  = upah minimum provinsi,  $X_1$  = indeks harga konsumen,  $X_2$  = tingkat partisipasi angkatan kerja,  $X_3$  = produk domestik regional bruto dan  $D_{ki}$  = variabel *dummy*,  $k = 1, 2, \dots, 33$  (provinsi).

**Kata kunci:** analisis regresi data panel; model *fixed effect*; variabel *dummy*

**Abstract.** Panel data is a combination of cross section data and time series data. One of panel data regression analysis model is the fixed effect model. The fixed effect model has the assumption that intercepts are different for each individual, but the slope coefficient is constant. Estimation is done by using dummy variables to explain the existence of intercept differences between individuals. This study aims to estimate the fixed effect model in panel data regression analysis using the least square dummy variable method and apply it to the provincial minimum wage data in Indonesia in 2014-2017. Based on the results of the research that has been done by using parameter estimator  $\hat{Y}_{LSDV} = (\tilde{X}^T \tilde{X})^{-1} \tilde{X}^T Y$  for fixed effect model in the panel regression analysis on provincial minimum wage data in Indonesia, we obtained as follows,  $\hat{Y}_{it} = 5.248452 + \sum_{i=2}^N \beta_{0i} D_{ki} + 0.007415X_{1it} + 0.002882X_{2it} + 1.63E-07X_{3it}$ . with,  $\hat{Y}_{it}$  = provincial minimum wage,  $X_1$  = consumer price index,  $X_2$  = labor force participation rate,  $X_3$  = regional gross domestic product,  $D_{ki}$  = dummy variable,  $k = 1, 2, \dots, 33$  (province).

**Keywords:** dummy variable; fixed effect models; panel data regression analysis



This is an open access article under the [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

## PENDAHULUAN

Analisis regresi adalah salah satu analisis dalam statistika yang digunakan untuk melihat ketergantungan antara satu variabel dependen pada satu atau lebih variabel independen dengan maksud mengestimasi dan memprediksi nilai variabel dependen berdasarkan nilai variabel independen yang diketahui. Data panel adalah data gabungan antara data runtun waktu (*time series*) dan data seksi silang (*cross section*). Data *time series* adalah data yang terdiri dari satu objek namun meliputi beberapa periode waktu. Data *cross section* adalah data yang terdiri atas beberapa atau banyak objek namun pada suatu periode waktu. Data panel adalah gabungan dari kedua jenis data tersebut menyebabkan data panel memiliki karakteristik gabungan dari keduanya yaitu terdiri atas beberapa objek dan meliputi beberapa periode waktu (Syukron & Fahri, 2018).

Analisis regresi data panel adalah analisis regresi dengan struktur data merupakan data panel (Athey et al., 2021). Secara umum dengan menggunakan data panel maka intersep dan koefisien *slope* yang diperoleh akan berbeda untuk setiap objek dan periode waktu. Hal itu menyebabkan di dalam analisis regresi data panel perlu menggunakan asumsi yang dibuat tentang intersep, koefisien *slope*, dan variabel gangguannya atau variabel *error* (Zyphur et al., 2020). Metode pendekatan digunakan dalam analisis regresi data panel untuk memperoleh model regresi data panel yang sesuai. Salah satu metode dalam mengestimasi model regresi dengan data panel adalah menggunakan pendekatan model *fixed effect*. Model *fixed effect* menggunakan variabel *dummy* untuk menunjukkan adanya perbedaan antara unit *cross section* dan *time series* (Hill et al., 2020; Kropko & Kubinec, 2020). Penelitian ini akan membahas lebih lanjut mengenai estimasi model *fixed effect* dalam analisis regresi data panel dengan metode *Least Square Dummy Variable* (LSDV).

Penggunaan data panel dalam penelitian memberikan beberapa keuntungan. Pertama, memungkinkan jumlah data meningkat dan memberikan sejumlah besar titik data yang dapat meningkatkan derajat bebas sehingga mengurangi terjadinya multikolinearitas diantara variabel-variabel independen yang terdapat dalam penelitian. Kedua, lebih bervariasi dan dapat mengurangi masalah yang muncul apabila ada variabel yang dihilangkan. Ketiga, dapat mengontrol heterogenitas individu (Barros et al., 2020). Secara umum model persamaan regresi data panel dapat dituliskan sebagai berikut:

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \varepsilon_{it}; \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T \quad (1)$$

dengan,

$Y_{it}$  = variabel dependen untuk pengamatan pada unit *cross section* ke- $i$  dan *time series* ke- $t$

$\beta_0$  = intersep

$\beta_1$  = koefisien kemiringan (*slope*) untuk semua unit

$X_{it}$  = variabel independen pada unit *cross section* ke- $i$  dan *time series* ke- $t$

$\varepsilon_{it}$  = nilai galat pada unit *cross section* ke- $i$  dan *time series* ke- $t$ .

Estimasi dengan model regresi data panel dapat dilakukan dengan tiga pendekatan yang sering digunakan, antara lain *Common Effect Model* (CEM), *Fixed Effect Model* (FEM), dan *Random Effect Model* (REM). Model *common effect* mengasumsikan bahwa intersep ( $\beta_0$ ) konstan atau sama untuk setiap individu maupun setiap waktu. Persamaan model CEM dapat dituliskan sebagai berikut (Nguyen Thai et al., 2021):

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

Pendekatan yang sesuai untuk mengestimasi parameter model CEM adalah *Ordinary Least Square* (OLS) atau sering pula disebut dengan Metode Kuadrat Terkecil (MKT) (Novais & Faria, 2021). Jika persamaan CEM ditulis dalam bentuk matriks, maka persamaannya menjadi seperti berikut (Draper, 1998):

$$Y = \beta X + e \tag{3}$$

$$e = Y - X\beta \tag{4}$$

dengan cara meminimumkan fungsi kuadrat total residual (*error*), sehingga didapatkan  $\hat{\beta}_{OLS}$  sebagai berikut:

$$\hat{\beta}_{OLS} = (X'X)^{-1} X'Y. \tag{5}$$

Model *fixed effect* adalah model yang mengasumsikan bahwa intersep berbeda antar unit *cross section*, sedangkan koefisien kemiringan (*slope*) nya tetap sama antar unit *cross section* (Basak & Das, 2017). Metode yang digunakan adalah suatu metode pengestimasi menggunakan MKT dengan adanya penambahan variabel *dummy*, disebut dengan metode *Least Square Dummy Variable* (LSDV). Variabel *dummy* adalah sebuah variabel yang hanya memiliki dua kemungkinan. Variabel *dummy* yang dibentuk berjumlah  $N-1$ , dengan  $\beta_0$  sebagai intersep untuk unit *cross section* yang pertama (Singh & Sachdeva, 2020). Model *fixed effect* dengan variabel *dummy* dapat dituliskan sebagai berikut:

$$Y_{it} = \beta_0 + \sum_{i=2}^N \beta_{0i} D_{ki} + \sum_{k=1}^K \beta_k X_{kit} + \varepsilon_{it}. \tag{6}$$

dengan,

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

Model *random effect* bertujuan untuk mengestimasi data panel menggunakan variabel gangguan (*error*) yang diduga memiliki hubungan antar waktu dan antar individu (Bell et al., 2019). Mengingat ada dua komponen yang berkontribusi pada pembentukan galat, yaitu individu dan waktu, maka galat pada REM juga diuraikan menjadi galat untuk komponen waktu dan komponen gabungan (Panichi et al., 2017). Estimasi yang tepat digunakan adalah dengan menggunakan *Generalized Least Square* (GLS). Persamaan model *random effect* dapat dituliskan sebagai berikut:

$$Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \dots + \beta_k X_{kit} + \varepsilon_{it} \tag{7}$$

$$\varepsilon_{it} = u_i + v_t + w_{it}. \tag{8}$$

Sebelum model diestimasi, maka dilakukan terlebih dahulu uji spesifikasi model untuk mengetahui model yang akan dipakai, dengan menggunakan uji Chow dan uji Hausman. Uji Chow digunakan untuk memilih model terbaik antara model *fixed effect* dengan model *common effect*.

Statistik uji:

$$F_{hitung} = \frac{[RRSS - URSS]/(n-1)}{URSS/(nT - n - K)} \tag{9}$$

dengan,

$RRSS$  = *restricted residual sums of squares* yang berasal dari model koefisien tetap

$URSS$  = *unrestricted residual sums of squares* yang berasal dari model efek tetap

$K$  = jumlah variabel independen

$n$  = jumlah individu (*unit cross section*)

$T$  = jumlah periode waktu (*unit time series*).

Uji Hausman digunakan untuk memilih model terbaik antara model *fixed effect* dengan model *random effect*. Statistik uji yang digunakan mengikuti distribusi *chi-squared* berdasarkan kriteria *Wald* (Moore, 1977):

$$W = (\hat{\beta}_{FEM} - \hat{\beta}_{REM})' [var(\hat{\beta}_{FEM} - \hat{\beta}_{REM})]^{-1} (\hat{\beta}_{FEM} - \hat{\beta}_{REM}) \tag{10}$$

dengan,

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

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

Pengujian asumsi klasik digunakan untuk mengetahui apakah asumsi dari pengujian sudah terpenuhi atau belum meliputi uji normalitas, uji multikolinearitas, uji heteroskedastisitas, dan uji autokorelasi. Uji normalitas berguna untuk membuktikan data dari sampel yang dimiliki berasal dari populasi berdistribusi normal. Uji normalitas dapat dideteksi dari metode yang dikembangkan oleh Jarque-Bera (JB). Metode JB ini didasarkan pada sampel besar yang diasumsikan bersifat *asymptotic* dengan menggunakan perhitungan *skewness* dan *kurtosis* (Siraj-Ud-Doulah, 2021). Statistik uji :

$$JB = N \left[ \frac{S^2}{6} + \frac{(K-3)^2}{24} \right]; \quad K = \frac{\hat{\mu}_4}{\hat{\mu}_2^2} = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^4}{\left( \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 \right)^2}; \quad S = \frac{\hat{\mu}_3}{\hat{\mu}_2^{3/2}} = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^3}{\left( \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 \right)^{3/2}} \quad (11)$$

Setelah terpilihnya model yang sesuai untuk digunakan pada analisis regresi data panel, maka yang selanjutnya akan dilakukan adalah uji terhadap parameter dalam model regresi. Terdapat dua pengujian yang harus dilakukan yaitu pengujian secara serentak dan pengujian secara parsial. Pengujian simultan dilakukan untuk mengetahui apakah semua variabel independen pada model berpengaruh secara signifikan terhadap variabel dependen, pengujian ini dilakukan secara simultan atau serentak. Statistik uji yang digunakan yaitu uji F, dengan keputusan jika nilai *p-value* <  $\alpha$  maka hipotesis nol ditolak yang berarti bahwa variabel independen secara simultan berpengaruh signifikan terhadap variabel dependen. Pengujian parsial digunakan untuk mengetahui variabel independen yang berpengaruh signifikan secara individu terhadap variabel dependen. Jika nilai *p-value* <  $\alpha$  maka hipotesis nol ditolak, yang berarti bahwa suatu variabel independen secara individual berpengaruh signifikan terhadap variabel dependen. Uji ini menggunakan statistik uji t, yaitu :

$$t_{hitung} = \frac{b_j}{Sb_j} \quad (13)$$

dengan,

$b_j$  = koefisien regresi (*slope*) variabel bebas ke- $j$  ( $j=1,2,\dots,k$ )

$Sb_j$  = *standard error* penduga  $b_j$ .

Koefisien determinasi adalah suatu nilai yang bertujuan untuk mengukur seberapa jauh kemampuan model dalam menerangkan variasi nilai dependen. Nilai  $R^2$  berada diantara 0 sampai 1. Nilai  $R^2$  yang kecil berarti kemampuan variabel-variabel independen dalam menjelaskan varian variabel dependen amat terbatas. Nilai  $R^2$  yang mendekati satu berarti variabel independen memberikan hampir semua informasi yang dibutuhkan untuk memprediksi variasi variabel dependen.

## METODE PENELITIAN

Data yang digunakan dalam penelitian ini adalah data Upah Minimum Provinsi (UMP), Indeks Harga Konsumen (IHK), Tingkat Partisipasi Angkatan Kerja (TPAK), dan Produk Domestik Regional Bruto (PDRB) di 34 Provinsi pada tahun 2014-2017 yang didapat dari buku Statistik Indonesia tahun 2016 dan 2018 yang merupakan buku publikasi resmi dari Badan Pusat Statistik (BPS).



Model *fixed effect* untuk  $i = N$  dan  $t = 1, 2, \dots, T$  serta  $k = i = N$  dalam persamaan (13) dapat dituliskan dalam bentuk matriks dengan cara yang sama seperti pada persamaan (14) sehingga menjadi,

$$\begin{bmatrix} Y_{N1} \\ Y_{N2} \\ \vdots \\ Y_{NT} \end{bmatrix} = \begin{bmatrix} 1 & X_{1N1} & X_{2N1} & \dots & X_{KN1} \\ 1 & X_{1N2} & X_{2N2} & \dots & X_{KN2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1NT} & X_{2NT} & \dots & X_{KNT} \end{bmatrix} \begin{bmatrix} \beta_{01} \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_P \end{bmatrix} + \begin{bmatrix} \varepsilon_{N1} \\ \varepsilon_{N2} \\ \vdots \\ \varepsilon_{NT} \end{bmatrix} \quad (17)$$

Model *fixed effect* terdiri dari  $N \times T$  pengamatan, berdasarkan uraian bentuk matriks yang sudah dijelaskan sebelumnya untuk masing-masing individu (*cross section*) maka secara keseluruhan  $N \times T$  pengamatan dapat dituliskan dalam bentuk matriks sebagai berikut:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{bmatrix} = \begin{bmatrix} \mathbf{1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{1} \end{bmatrix} \begin{bmatrix} \beta_{01} \\ \beta_{02} \\ \vdots \\ \beta_{0N} \end{bmatrix} + \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix} \beta + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_N \end{bmatrix} \quad (18)$$

(NT x 1)      (NT x N)      (N x 1)      (NT x P)(P x 1)      (NT x 1)

dimana,

$$Y_i = \begin{bmatrix} Y_{i1} \\ Y_{i2} \\ \vdots \\ Y_{iT} \end{bmatrix} \quad X_i = \begin{bmatrix} X_{1i1} & X_{2i1} & \dots & X_{ki1} \\ X_{1i2} & X_{2i2} & \dots & X_{ki2} \\ \vdots & \vdots & \ddots & \vdots \\ X_{1iT} & X_{2iT} & \dots & X_{kiT} \end{bmatrix} \quad \varepsilon_i = \begin{bmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \\ \vdots \\ \varepsilon_{iT} \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_T \end{bmatrix} \quad \mathbf{1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \quad \mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$D = \begin{bmatrix} \mathbf{1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{1} \end{bmatrix}$$

dengan  $\mathbf{1}$  dan  $\mathbf{0}$  adalah vektor berukuran  $T \times 1$ , maka matriks (4.8) dapat dituliskan sebagai berikut:

$$Y = D\beta_0 + X\beta + \varepsilon = [D \quad X] \begin{bmatrix} \beta_0 \\ \beta \end{bmatrix} + \varepsilon \quad (19)$$

Misal  $[D \quad X] = \tilde{X}$  dan  $\begin{bmatrix} \beta_0 \\ \beta \end{bmatrix} = \gamma$ , maka persamaan (18) dapat dituliskan menjadi,

$$Y = \tilde{X}\gamma + \varepsilon. \quad (20)$$

Persamaan (19) jika dituliskan dalam bentuk matriks lengkap menjadi,

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{bmatrix} = \begin{bmatrix} \mathbf{1} & D_{k1} & D_{k1} & \dots & D_{k1} & X_{11} & X_{21} & \dots & X_{k1} \\ \mathbf{1} & D_{k2} & D_{k2} & \dots & D_{k2} & X_{12} & X_{22} & \dots & X_{k2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{1} & D_{kN} & D_{kN} & \dots & D_{kN} & X_{1T} & X_{2T} & \dots & X_{kT} \end{bmatrix} \begin{bmatrix} \beta_{01} \\ \beta_{02} \\ \vdots \\ \beta_{0N} \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_N \end{bmatrix} \quad (21)$$

Model *fixed effect* yang sudah didapatkan ialah model awal, selanjutnya akan dicari estimasi parameter  $\gamma$ . Estimasi parameter  $\gamma$  menggunakan metode kuadrat terkecil atau *ordinary least square*, dengan cara meminimumkan fungsi jumlah kuadrat galat.

$$\begin{aligned}
 S &= \boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} \\
 &= \boldsymbol{\varepsilon}_1 \boldsymbol{\varepsilon}_1 + \boldsymbol{\varepsilon}_2 \boldsymbol{\varepsilon}_2 + \dots + \boldsymbol{\varepsilon}_N \boldsymbol{\varepsilon}_N \\
 &= \sum_{i=1}^N \boldsymbol{\varepsilon}_i^2 \\
 &= (\mathbf{Y} - \tilde{\mathbf{X}}\boldsymbol{\gamma})^T (\mathbf{Y} - \tilde{\mathbf{X}}\boldsymbol{\gamma})
 \end{aligned} \tag{22}$$

Meminimumkan suatu fungsi dapat dilakukan dengan melakukan turunan pertama  $S$  terhadap  $\boldsymbol{\gamma}$  dan kemudian hasil turunan tersebut disamakan dengan nol.

$$\begin{aligned}
 \frac{dS}{d\boldsymbol{\gamma}} &= \frac{d\left((\mathbf{Y} - \tilde{\mathbf{X}}\boldsymbol{\gamma})^T (\mathbf{Y} - \tilde{\mathbf{X}}\boldsymbol{\gamma})\right)}{d\boldsymbol{\gamma}} \\
 &= \frac{d(\mathbf{Y}^T \mathbf{Y} - \mathbf{Y}^T \tilde{\mathbf{X}}\boldsymbol{\gamma} - \tilde{\mathbf{X}}^T \boldsymbol{\gamma}^T \mathbf{Y} + \tilde{\mathbf{X}}^T \boldsymbol{\gamma}^T \tilde{\mathbf{X}}\boldsymbol{\gamma})}{d\boldsymbol{\gamma}} \\
 &= \frac{d(\mathbf{Y}^T \mathbf{Y} - \mathbf{Y}^T \tilde{\mathbf{X}}\boldsymbol{\gamma} - \tilde{\mathbf{X}}^T \boldsymbol{\gamma}^T \mathbf{Y} + \tilde{\mathbf{X}}^T \boldsymbol{\gamma}^T \tilde{\mathbf{X}}\boldsymbol{\gamma})}{d\boldsymbol{\gamma}} \\
 &= 0 - \mathbf{Y}^T \tilde{\mathbf{X}} - \mathbf{Y}^T \tilde{\mathbf{X}} + 2\boldsymbol{\gamma}^T \tilde{\mathbf{X}}^T \tilde{\mathbf{X}} \\
 &\quad - 2\mathbf{Y}^T \tilde{\mathbf{X}} + 2\boldsymbol{\gamma}^T \tilde{\mathbf{X}}^T \tilde{\mathbf{X}} = 0 \\
 &\quad \tilde{\mathbf{X}}^T \tilde{\mathbf{X}}\hat{\boldsymbol{\gamma}} = \tilde{\mathbf{X}}^T \mathbf{Y}
 \end{aligned}$$

Sehingga didapatkan  $\hat{\boldsymbol{\gamma}}_{LSDV}$  sebagai berikut:

$$\hat{\boldsymbol{\gamma}}_{LSDV} = (\tilde{\mathbf{X}}^T \tilde{\mathbf{X}})^{-1} \tilde{\mathbf{X}}^T \mathbf{Y} \tag{23}$$

Penduga parameter  $\boldsymbol{\gamma}$  dengan metode LSDV pada model *fixed effect* mempunyai beberapa karakteristik, yaitu:

1. Bias penduga LSDV  
 $\hat{\boldsymbol{\gamma}}_{LSDV}$  adalah penduga tak bias karena  $E(\hat{\boldsymbol{\gamma}}_{LSDV}) = \boldsymbol{\gamma}$
2. Variansi penduga LSDV

$$\text{var}(\hat{\boldsymbol{\gamma}}_{LSDV}) = \boldsymbol{\sigma}^2 (\tilde{\mathbf{X}}^T \tilde{\mathbf{X}})^{-1} \tag{24}$$

dimana  $\boldsymbol{\sigma}^2 (\tilde{\mathbf{X}}^T \tilde{\mathbf{X}})^{-1}$  merupakan variansi terkecil dari semua penaksir linear tak bias (Sembiring, 1995).

Penerapan estimasi model *fixed effect* pada analisis regresi data panel dilakukan dengan menggunakan data Upah Minimum Provinsi (UMP) sebagai variabel dependen (Y), Indeks Harga Saham (IHK) sebagai variabel independen pertama (X1), Tingkat Partisipasi Angkatan Kerja (TPAK) sebagai variabel independen kedua (X2), dan Produk Domestik Regional Bruto (PDRB) sebagai variabel independen ketiga (X3).

Model *common effect* adalah model dalam analisis regresi data panel yang tidak memperhatikan pengaruh yang diberikan oleh dimensi individu/unit maupun waktu. Berdasarkan analisis yang telah dilakukan maka diperoleh hasil estimasi parameter untuk CEM yang dapat dilihat pada Tabel 1.

Tabel 1. Estimasi parameter CEM

Variabel	Koefisien	Standar Error	t-hitung	P-value
C	-649396.5	855819.9	-	0.4493
IHK (X1)	25411.91	4969.189	5.113894	0.0000
TPAK (X2)	-9528.135	9011.445	-	0.2923
			1.057337	

PDRB (X3)	-0.074513	0.093698	-	0.4279
			0.795252	

Berdasarkan Tabel 1. diperoleh persamaan umum untuk model dengan pendekatan *common effect model* sebagai berikut:

$$\hat{Y}_{it} = -649396.5 + 25411.91X_{1it} - 9528.135X_{2it} - 0.074513X_{3it}. \tag{25}$$

Model *fixed effect* mengasumsikan bahwa terdapat pengaruh yang berbeda antar unit *cross section*. Berdasarkan analisis yang telah dilakukan maka diperoleh hasil estimasi parameter untuk FEM yang dapat dilihat pada Tabel 2.

Tabel 2. Estimasi parameter FEM

Variabel	Koefisien	Standar Error	t-hitung	P-value
C	-1784270.	359142.4	-	0.0000
			4.968078	
IHK (X1)	30945.35	1486.143	20.82259	0.0000
TPAK (X2)	3113.362	5523.135	0.563695	0.5742
PDRB (X3)	0.525557	0.318619	1.649483	0.1022

Nilai  $\sum_{i=2}^N \beta_{0i} D_{ki}$  adalah nilai variabel *dummy* yang berbeda-beda untuk tiap unit *cross section*, sehingga diperoleh persamaan umum untuk model dengan pendekatan *fixed effect model* sebagai berikut:

$$\hat{Y}_{it} = -1784270. + \sum_{i=2}^N \beta_{0i} D_{ki} + 30945.35X_{1it} + 3113.362X_{2it} + 0.525557X_{3it}. \tag{26}$$

Model *random effect* mengasumsikan bahwa variabel gangguan dalam data panel mungkin saling berhubungan antar waktu dan individu. Berdasarkan analisis yang telah dilakukan maka diperoleh hasil estimasi parameter untuk REM yang dapat dilihat pada Tabel 3.

Tabel 3. Estimasi parameter REM

Variabel	Koefisien	Standar Error	t-hitung	P-value
C	-2221110.	371978.5	-5.971071	0.0000
IHK (X1)	31931.14	1321.383	24.16494	0.0000
TPAK (X2)	1362.012	5287.258	0.257603	0.7971
PDRB (X3)	0.088843	0.159414	0.557307	0.5783

Berdasarkan Tabel 3. diperoleh persamaan umum untuk model dengan pendekatan *random effect model* sebagai berikut:

$$\hat{Y}_{it} = -2221110. + 31931.14X_{1it} + 1362.012X_{2it} + 0.088843X_{3it}. \tag{27}$$

Pemilihan model estimasi analisis regresi data panel merupakan tahapan analisis untuk menentukan model estimasi terbaik dalam analisis data panel dengan menggunakan uji Chow dan uji Hausman. Berdasarkan uji Chow yang telah dilakukan didapatkan hasil analisis pada Tabel 4.

Tabel 4. Hasil uji Chow

Keterangan	Nilai
Uji Efek Statistik	<i>Cross-section Chi-square</i> 447.363398
db	33
P-value	0.0000



Berdasarkan Tabel 4. nilai *p-value* dari *cross-section chi-square* adalah sebesar  $0,0000 < \alpha$  (0,05) maka tolak  $H_0$ , yang berarti bahwa model dengan metode FEM lebih baik digunakan daripada model dengan metode CEM. Sedangkan berdasarkan uji Hausman yang telah dilakukan didapatkan hasil analisis pada Tabel 5.

Tabel 5. Hasil uji Hausman

Keterangan	Nilai
Uji Efek Statistik	<i>Cross-section random</i> 8.583450
db	3
<i>P-value</i>	0.0337

Nilai *p-value* dari *cross-section random* adalah sebesar  $0,0337 < \alpha$  (0,05) maka tolak  $H_0$ , yang berarti bahwa metode FEM lebih baik digunakan daripada REM. Sehingga dapat disimpulkan bahwa berdasarkan uji Chow dan uji Hausman metode terbaik adalah FEM. Pemilihan metode terbaik juga dapat dilakukan dengan melihat perbandingan antara nilai statistik pada masing-masing metode untuk melihat model terbaik.

Tabel 6. Perbandingan nilai statistik FEM dan REM

Statistik	<i>Fixed Effect</i>	<i>Random Effect</i>
<i>R-squared</i>	0.969129	0.832047
<i>Adjusted R-squared</i>	0.957903	0.828230
Standar Error	91717.67	92610.85
<i>P-value</i>	0.000000	0.000000

Berdasarkan Tabel 6. model *fixed effect* memberikan interpretasi yang lebih baik untuk semua nilai statistiknya dibandingkan dengan model *random effect*, sehingga dapat disimpulkan bahwa model *fixed effect* adalah model yang sesuai untuk pemodelan data panel pada penelitian ini. 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 multikolinearitas, uji heteroskedastisitas, dan uji autokorelasi.

Berdasarkan uji asumsi klasik yang telah dilakukan terdapat masalah autokorelasi pada data penelitian, sehingga harus dilakukan perbaikan agar semua tahapan dalam uji asumsi klasik terpenuhi. Transformasi adalah salah satu cara yang dapat dilakukan untuk memperbaiki masalah uji asumsi klasik yang tidak terpenuhi. Penelitian ini dilakukan transformasi terhadap variabel dependen. Estimasi model *fixed effect* dengan transformasi variabel dependen ditunjukkan pada Tabel 7.

Tabel 7. Estimasi parameter FEM dengan transformasi log(Y)

Variabel	Koefisie n	Standar Error	t-hitung	<i>P-value</i>
C	5.248452	0.087271	60.13873	0.0000
IHK (X1)	0.007415	0.000361	20.52731	0.0000
TPAK (X2)	0.002882	0.001342	2.146986	0.0342
PDRB (X3)	1.63E-07	7.74E-08	2.102899	0.0380

Berdasarkan Tabel 7. nilai  $\sum_{i=2}^N \beta_{0i} D_{ki}$  adalah nilai variabel *dummy* yang berbeda-beda untuk tiap unit *cross section*, dapat dilihat pada Tabel 8.

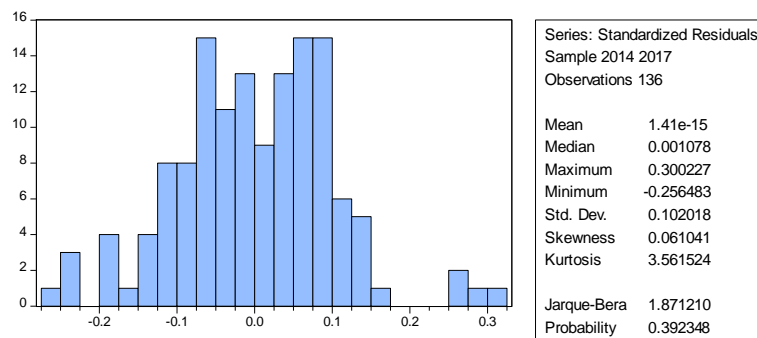
Tabel 8. Nilai variabel *dummy* dengan metode *fixed effect*

Provinsi	Nilai Variabel Dummy	Provinsi	Nilai Variabel Dummy
Aceh	0.000000	Nusa Tenggara Barat	-0.204246
Sumatera Utara	-0.204750	Nusa Tenggara Timur	-0.245069
Sumatera Barat	-0.165088	Kalimantan Barat	-0.209884
Riau	-0.118234	Kalimantan Tengah	-0.054405
Jambi	-0.056603	Kalimantan Selatan	-0.071709
Sumatera Selatan	-0.216932	Kalimantan Timur	-0.105781
Bengkulu	-0.043239	Kalimantan Utara	-0.085648
Lampung	-0.061917	Sulawesi Utara	-0.020994
Kep. Bangka Belitung	-0.166415	Sulawesi Tengah	-0.181776
Kep. Riau	-0.116350	Sulawesi Selatan	-0.043667
DKI Jakarta	-0.127534	Sulawesi Tenggara	-0.102005
Jawa Barat	-0.472440	Gorontalo	-0.087369
Jawa Tengah	-0.324751	Sulawesi Barat	-0.112660
DI Yogyakarta	-0.531287	Maluku	-0.101353
Jawa Timur	-0.219727	Maluku Utara	-0.147263
Banten	-0.455774	Papua Barat	-0.012147
Bali	-0.133849	Papua	-0.046565

Berdasarkan Tabel 8. diperoleh persamaan umum untuk model dengan pendekatan *fixed effect model* sebagai berikut:

$$\hat{Y}_{it} = 5.248452 + \sum_{i=2}^N \beta_{0i} D_{ki} + 0.007415X_{1it} + 0.002882X_{2it} + 1.63E-07X_{3it} \quad (28)$$

Uji asumsi klasik dilakukan kembali untuk memastikan bahwa semua uji asumsi klasik dengan transformasi variabel dependen ke dalam bentuk log terpenuhi. Berikut hasil uji normalitas yang disajikan pada Gambar 2.



Gambar 2. Hasil uji normalitas dengan transformasi log(Y)

Berdasarkan uji normalitas Gambar 2. dengan uji Jarque-Bera yang telah dilakukan didapat nilai probabilitas adalah sebesar 0,392348. Karena nilai  $prob > \alpha$  maka tidak tolak  $H_0$ . Sehingga dapat disimpulkan bahwa residual dari data penelitian berdistribusi normal. Kemudian hasil dari uji multikolinearitas ditunjukkan pada Tabel 9.

Tabel 9. Hasil uji multikolinearitas dengan VIF

Variabel Bebas	Koefisien Ragam	VIF
IHK (X1)	24692843	1.000762
TPAK (X2)	81206149	1.035521
PDRB (X3)	0.008779	1.036115

Berdasarkan Tabel 9. nilai VIF dari ketiga variabel independen di dalam penelitian  $< 10$ , yang berarti dapat disimpulkan bahwa tidak terjadi multikolinearitas antar variabel independen dalam penelitian. Hasil uji heteroskedastisitas disajikan pada Tabel 10.

Tabel 10. Hasil uji heteroskedastisitas dengan uji Glejser

Variabel	t-hitung	P-value
IHK(X1)	-0.342040	0.7330
TPAK(X2)	-1.162663	0.2478
PDRB(X3)	-0.025827	0.9794

Berdasarkan Tabel 10. nilai *p-value* untuk semua variabel independen dalam penelitian  $> 0.05$ , sehingga dapat disimpulkan bahwa tidak ada gejala heteroskedastisitas dalam model regresi. Pengujian autokorelasi dilakukan dengan melihat nilai statistik Durbin Watson (*d*), berdasarkan analisis yang telah dilakukan didapat nilai sebesar 1.937402. Nilai tersebut selanjutnya dibandingkan dengan nilai Durbin-Watson pada tabel, dengan jumlah *n* (pengamatan) sebanyak 136 dan *k* (variabel independen) dalam penelitian ini sebanyak 3 diperoleh nilai tabel DW yaitu  $d_L = 1.6751$  dan  $d_U = 1.7652$ . Berdasarkan kriteria pengujian autokorelasi nilai tersebut berada pada kriteria  $d_U < d < (4 - d_U)$  maka dapat disimpulkan bahwa tolak  $H_0$  yang berarti bahwa tidak terdapat masalah autokorelasi pada model regresi.

Pengujian parameter model regresi adalah pengujian yang dilakukan untuk mengetahui hubungan antara variabel dependen dengan variabel independen dalam model regresi. Pengujian simultan dilakukan dengan melihat nilai F-hitung dan probabilitas statistik dari F-hitung. Hasil pengujian bernilai 0.000000 maka  $< \alpha$  sehingga tolak  $H_0$ . Dengan tingkat kepercayaan sebesar 95% didapat disimpulkan bahwa terdapat pengaruh yang signifikan pada data penelitian antara variabel independen (IHK, TPAK, dan PDRB) terhadap variabel dependen (UMP).

Pengujian parsial dilakukan untuk mengetahui pengaruh signifikansi setiap variabel independen terhadap variabel dependen dengan menggunakan uji t. Pengujian ini akan dilakukan dengan dua tahap uji untuk masing-masing variabel independen pada model regresi dalam penelitian, yaitu uji signifikansi dan uji arah atas nilai koefisien masing-masing variabel independennya.

1. Indeks Harga Konsumen (IHK)

Variabel IHK memiliki nilai *p-value* sebesar 0.0000. Nilai tersebut  $< 0.05$  maka tolak  $H_0$ . Sehingga dapat disimpulkan bahwa IHK secara signifikan mempengaruhi UMP.

- Tabel 14 menunjukkan bahwa koefisien IHK(X1) bernilai 0.007415 artinya hubungan yang terjadi antara UMP dan IHK adalah hubungan yang searah atau positif.
2. Tingkat Partisipasi Angkatan Kerja (TPAK)  
Variabel TPAK memiliki nilai *p-value* sebesar 0.0342. Nilai tersebut < 0.05 maka tolak  $H_0$ . Sehingga dapat disimpulkan bahwa TPAK secara signifikan mempengaruhi UMP. Tabel 14 menunjukkan bahwa koefisien TPAK(X2) bernilai 0.002882 artinya hubungan yang terjadi antara UMP dan TPAK adalah hubungan yang searah atau positif.
  3. Produk Domestik Regional Bruto (PDRB)  
Variabel PDRB memiliki nilai *p-value* sebesar 0.0380. Nilai tersebut < 0.05 maka tolak  $H_0$ . Sehingga dapat disimpulkan bahwa PDRB secara signifikan mempengaruhi UMP. Tabel 14 menunjukkan bahwa koefisien PDRB(X3) bernilai 1.63E-07 artinya hubungan yang terjadi antara UMP dan PDRB adalah hubungan yang searah atau positif.

Koefisien determinasi digunakan dengan tujuan untuk melihat seberapa besar variabel independen mampu menjelaskan variabel dependen di dalam model. Berdasarkan hasil analisis nilai koefisien determinasi didapat adalah 0.961436, yang berarti bahwa sebesar 96.14% variabel independen (IHK, TPAK, dan PDRB) dapat menjelaskan variabel dependen (UMP), sedangkan sisanya yaitu sebesar 3.86% dijelaskan oleh faktor lain yang tidak dimasukkan dalam model penelitian.

Penelitian ini sejalan dengan temuan Rahardja dan Kusuma (2020), yang menemukan bahwa PDRB memiliki pengaruh signifikan terhadap UMP, meskipun koefisien determinasi dalam penelitian tersebut lebih rendah, yaitu 85%. Selain itu, hasil penelitian ini juga konsisten dengan Hastuti (2019), yang menunjukkan bahwa IHK memiliki hubungan positif dengan UMP, meskipun nilai koefisien determinasi hanya mencapai 70%. Penelitian terdahulu oleh Santoso dan Yuliawati (2018) juga mendukung temuan ini, di mana TPAK berpengaruh signifikan terhadap UMP dengan hubungan yang positif, tetapi penelitian mereka hanya berfokus pada lima provinsi tertentu dan menggunakan tingkat kepercayaan 90%. Dengan demikian, penelitian ini memberikan kontribusi yang lebih komprehensif dibandingkan penelitian terdahulu karena mencakup tiga variabel independen sekaligus, dengan koefisien determinasi yang lebih tinggi, menunjukkan kemampuan model yang lebih baik dalam menjelaskan hubungan antarvariabel.

## KESIMPULAN DAN SARAN

Bentuk estimasi penduga parameter model *fixed effect* dalam analisis regresi data panel dengan metode *Least Square Dummy Variable* (LSDV) adalah sebagai berikut:  $\hat{\boldsymbol{\gamma}}_{LSDV} =$

$(\tilde{\mathbf{X}}^T \tilde{\mathbf{X}})^{-1} \tilde{\mathbf{X}}^T \mathbf{Y}$  dengan,  $\boldsymbol{\gamma} = \begin{bmatrix} \beta_0 \\ \beta \end{bmatrix}$  dan  $\tilde{\mathbf{X}} = [\mathbf{D} \quad \mathbf{X}]$ , serta karakteristik penduga  $\hat{\boldsymbol{\gamma}}_{LSDV}$  yaitu: Tak

bias ( $E(\hat{\boldsymbol{\gamma}}_{LSDV}) = \boldsymbol{\gamma}$ ) dan Ragam minimum ( $var(\hat{\boldsymbol{\gamma}}_{LSDV}) = \sigma^2(\tilde{\mathbf{X}}^T \tilde{\mathbf{X}})^{-1}$ ). Hasil analisis model *fixed effect* terhadap data Upah Minimum Provinsi (UMP), Indeks Harga Konsumen (IHK), Tingkat Partisipasi Angkatan Kerja (TPAK), dan Produk Domestik Regional Bruto (PDRB) diperoleh kesimpulan model *fixed effect* dalam analisis regresi data panel dengan metode *Least Square Dummy Variable* (LSDV) untuk pemodelan Upah Minimum Provinsi (UMP) di Indonesia tahun 2014 hingga 2017 adalah sebagai berikut,  $\hat{Y}_{it} = 5.248452 + \sum_{i=2}^N \beta_{0i} D_{ki} + 0.007415X_{1it} + 0.002882X_{2it} + 1.63E-07X_{3it}$ . Variabel independen dalam model (IHK, TPAK, dan PDRB) memberikan pengaruh yang signifikan terhadap UMP baik secara parsial maupun simultan pada taraf signifikansi  $\alpha = 5\%$ . Sebesar 96.14% faktor Indeks Harga Konsumen (IHK), Tingkat Partisipasi Angkatan Kerja (TPAK), dan Produk Domestik Regional Bruto (PDRB) bersama-sama mempengaruhi Upah Minimum Provinsi (UMP),

sedangkan sisanya sebesar 3.86% dijelaskan oleh faktor lain di luar penelitian yang tidak diikutsertakan ke dalam model.

#### DAFTAR PUSTAKA

- Athey, S., Bayati, M., Doudchenko, N., Imbens, G., & Khosravi, K. (2021). Matrix completion methods for causal panel data models. *Journal of the American Statistical Association*, 116(536), 1716–1730. <https://doi.org/10.1080/01621459.2021.1891924>
- Barros, L. A. B. C., Bergmann, D. R., Castro, F. H., & Silveira, A. D. M. da. (2020). Endogeneity in panel data regressions: methodological guidance for corporate finance researchers. *Revista Brasileira de Gestão de Negócios*, 22, 437–461. <https://doi.org/10.7819/rbgn.v22i0.4059>
- Basak, G. K., & Das, S. (2017). Intercept homogeneity test for fixed effect models under cross-sectional dependence: some insights. *Journal of Econometric Methods*, 6(1), 20150004. <https://doi.org/10.1515/jem-2015-0004>
- Bell, A., Fairbrother, M., & Jones, K. (2019). Fixed and random effects models: making an informed choice. *Quality & Quantity*, 53, 1051–1074. <https://doi.org/10.1007/s11135-018-0802-x>
- Draper, N. R. (1998). *Applied regression analysis*. McGraw-Hill. Inc. <https://doi.org/10.1002/9781118625590>
- Hastuti, E. (2019). Pengaruh Indeks Harga Konsumen terhadap Upah Minimum Provinsi di Jawa Tengah. *Jurnal Ekonomi dan Pembangunan Daerah*, 12(1), 45-58.
- Hill, T. D., Davis, A. P., Roos, J. M., & French, M. T. (2020). Limitations of fixed-effects models for panel data. *Sociological Perspectives*, 63(3), 357–369. <https://doi.org/10.1177/0731121419863785>
- Kropko, J., & Kubinec, R. (2020). Interpretation and identification of within-unit and cross-sectional variation in panel data models. *PloS One*, 15(4), e0231349. <https://doi.org/10.1371/journal.pone.0231349>
- Moore, D. S. (1977). Generalized inverses, Wald's method, and the construction of chi-squared tests of fit. *Journal of the American Statistical Association*, 72(357), 131–137. <https://doi.org/10.1080/01621459.1977.10479921>
- Nguyen Thai, D., Minh, P. Van, Phan Hoang, C., Ta Duc, T., Nguyen Thi Cam, N., & Nguyen Thi, D. (2021). Bending of symmetric sandwich FGM beams with shear connectors. *Mathematical Problems in Engineering*, 2021(1), 7596300. <https://doi.org/10.1155/2021/7596300>
- Novais, L., & Faria, S. (2021). Comparison of the EM, CEM and SEM algorithms in the estimation of finite mixtures of linear mixed models: a simulation study. *Computational Statistics*, 36(4), 2507–2533. <https://doi.org/10.1007/s00180-021-01088-1>
- Panichi, F., Goździewski, K., & Turchetti, G. (2017). The reversibility error method (REM): a new, dynamical fast indicator for planetary dynamics. *Monthly Notices of the Royal Astronomical Society*, 468(1), 469–491. <https://doi.org/10.1093/mnras/stx374>
- Rahardja, D., & Kusuma, A. (2020). Analisis PDRB terhadap Upah Minimum Provinsi di Indonesia Tahun 2015-2019. *Jurnal Ekonomi Regional*, 15(2), 67-80.
- Santoso, B., & Yuliawati, L. (2018). Pengaruh Tingkat Partisipasi Angkatan Kerja terhadap Upah Minimum di Indonesia. *Jurnal Ilmu Ekonomi dan Ketenagakerjaan*, 8(3), 123-130.
- 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. <https://doi.org/10.34293/sijash.v7i3.411>
- Siraj-Ud-Doulah, M. (2021). An alternative measures of moments skewness kurtosis and jb test of normality. *Journal of Statistical Theory and Applications*, 20(2), 219–227. <https://doi.org/10.2991/jsta.d.210525.002>

- Syukron, M., & Fahri, H. M. (2018). Pendekatan Regresi Data Panel untuk Pemodelan Jumlah Angkatan Kerja dan Penanaman Modal Luar Negeri terhadap PDRB Provinsi di Indonesia. *Indonesian Journal of Applied Statistics*, 1(2), 100–116. <https://doi.org/10.13057/ijas.v1i2.26172>
- Zyphur, M. J., Voelkle, M. C., Tay, L., Allison, P. D., Preacher, K. J., Zhang, Z., Hamaker, E. L., Shamsollahi, A., Pierides, D. C., & Koval, P. (2020). From data to causes II: Comparing approaches to panel data analysis. *Organizational Research Methods*, 23(4), 688–716. <https://doi.org/10.1177/1094428119847280>