

Penerapan Regresi Data Panel Komponen Satu Arah untuk Menentukan Faktor-Faktor yang Mempengaruhi Indeks Pembangunan Manusia

Bayu Sutikno¹, Alfensi Faruk², Oki Dwipurwani³

^{1,2,3}Jurusan Matematika, Fakultas MIPA, Universitas Sriwijaya
Jl. Raya Palembang-Prabumulih KM 32, Indralaya 30662

Email: bayu.sutikno21@gmail.com, alfensifaruk@unsri.ac.id, okidwip@yahoo.com

ABSTRAK

Efek spesifikasi wilayah dan waktu dari data panel dapat menjelaskan perbedaan antar nilai Indeks Pembangunan Manusia (IPM) di setiap wilayah dan periode waktu. Tujuan dari penelitian ini adalah (1) menentukan model regresi data panel terbaik dengan efek komponen satu arah dalam menjelaskan tingkat keberagaman dari nilai IPM, dan (2) menganalisis faktor-faktor yang berpengaruh signifikan terhadap perubahan nilai IPM. Data yang digunakan dalam penelitian ini adalah nilai IPM dari seluruh kabupaten dan kota di Provinsi Sumatera Selatan mulai tahun 2007 hingga 2014 yang diperoleh dari Badan Pusat Statistik Provinsi Sumatera Selatan (BPSPSS). Berdasarkan uji Chow, uji Hausman, dan pemilihan efek komponen satu arah, hasil yang diperoleh memperlihatkan bahwa model regresi data panel terbaik dari nilai IPM di Provinsi Sumatera Selatan mulai tahun 2007 hingga 2014 adalah model efek tetap komponen waktu yang diboboti dengan *seemingly unrelated regression* (SUR). Adapun, faktor-faktor yang berpengaruh signifikan terhadap perubahan nilai IPM di Provinsi Sumatera Selatan mulai tahun 2007 hingga 2014 adalah banyaknya rumah tangga yang dapat mengakses air bersih, angka partisipasi Sekolah Menengah Atas (SMA), angka melek huruf, dan tingkat partisipasi angkatan kerja.

Kata kunci: regresi data panel, IPM, uji Chow, uji Hausman, efek komponen satu arah.

ABSTRACT

*Region and time effects specification of the panel data can explain the difference among the values of Human Development Index (HDI) in each region and period of time. The objectives of this study are (1) to determine the best panel data regression model with one way component effect in explaining the diversity of HDI values, and (2) to analyze the factors that significantly affect the changes of the HDI values. The data used in this research are the HDI values of all districts and cities in South Sumatra Province from year 2007 to 2014 which obtained from the Statistics of Sumatera Selatan Province. By using the Chow test, Hausman test, and the selection of the one way component effects, the results show that the best panel data regression model for HDI values in South Sumatra Province from year 2007 to 2014 is time component fixed effect model with *seemingly unrelated regression* weighted. Meanwhile, the factors that significantly affect the changing of the HDI values in South Sumatra Province from year 2007 to 2014 are the number of households with clean water access, senior high school participation rate, the literacy rate, and the labor force participation rate.*

Keywords: *panel data regression, HDI, Chow test, Hausman test, one way component effects.*

1. Pendahuluan

Sumber daya manusia yang berkualitas adalah aset terbaik dalam menunjang pembangunan di suatu negara untuk jangka pendek maupun panjang. Sehubungan dengan hal tersebut, kualitas dari sumber daya manusia di suatu daerah dapat diketahui berdasarkan nilai Indeks Pembangunan Manusia (IPM). Capaian nilai IPM didasarkan pada tiga dimensi pokok kehidupan yaitu bidang kesehatan, pendidikan, dan ekonomi [1]. Perbedaan nilai IPM di masing-masing wilayah dan periode waktu diduga akibat perbedaan karakteristik di masing-masing wilayah dan periode waktu, sehingga untuk mengetahui perbedaan tersebut digunakan data panel yang dapat memuat informasi mengenai efek spesifikasi wilayah dan efek spesifikasi waktu. Adapun, model regresi untuk data yang memuat baik komponen wilayah maupun komponen waktu disebut sebagai model regresi data panel.

Model regresi data panel banyak digunakan di berbagai bidang. Sebagai contoh, Sugiharso dan Ester [12] mengaplikasikan Model Efek Umum (MEU) dengan metode estimasi parameter *Ordinary Least Square* (OLS) untuk menentukan determinan investasi portofolio. Dalam MEU, diasumsikan bahwa efek spesifikasi wilayah dan waktu memiliki pengaruh yang sama, sehingga perbedaan karakteristik dari masing-masing wilayah maupun periode waktu tidak dapat teridentifikasi. Fajriyah dan Santi [5] menggunakan Model Efek Tetap (MET) komponen wilayah dengan metode estimasi parameter *Least Square Dummy Variable* (LSDV) untuk menentukan faktor-faktor yang mempengaruhi kemiskinan kabupaten/kota di Provinsi Jawa Timur. Dalam MET, diasumsikan bahwa efek spesifikasi wilayah dan waktu memiliki pengaruh yang tidak sama. Perbedaan karakteristik dari masing-masing wilayah maupun periode waktu dalam MET dapat teridentifikasi dengan adanya penambahan variabel *dummy* ke dalam model. Sementara itu, Setiawan dan Krismanti [11] menerapkan Model Efek Acak (MEA) komponen wilayah dengan metode estimasi parameter *Generalized Least Square* (GLS) untuk mengkaji konsumsi rokok di Indonesia. Dalam MEA, perbedaan karakteristik dari masing-masing wilayah maupun periode waktu dapat teridentifikasi. Hal ini dikarenakan adanya efek spesifikasi wilayah dan waktu yang tidak berkorelasi dengan variabel prediktor sebagai bagian dari komponen residual model, sehingga kekurangan dari bentuk model yang sebenarnya dapat diatasi.

Penerapan regresi data panel dalam penentuan faktor-faktor yang mempengaruhi nilai IPM pernah dilakukan oleh Melliana dan Ismaini [9]. Akan tetapi, model yang digunakan dalam penelitian tersebut hanya satu saja yaitu MET *cross section weight*. Oleh karena itu, penelitian ini bertujuan untuk memilih model regresi data panel terbaik dalam memodelkan nilai IPM dan menentukan faktor-faktor yang berpengaruh signifikan pada nilai IPM menggunakan data nilai IPM Provinsi Sumatera Selatan yang diperoleh dari [2]. Model regresi data panel yang dibahas dalam penelitian ini adalah MEU, MET, dan MEA yang tidak hanya melibatkan komponen wilayah saja namun juga melibatkan komponen waktu. Adapun, variabel-variabel prediktor yang digunakan dalam penelitian ini adalah tingkat keluhan kesehatan masyarakat, jumlah rumah tangga dengan akses air bersih, angka partisipasi SMA, angka melek huruf, tingkat partisipasi angkatan kerja, dan nilai produk domestik regional bruto.

2. Metode Penelitian

Data IPM dan faktor-faktor yang diduga berpengaruh terhadap nilai IPM dalam penelitian ini bersumber dari [2] dengan cakupan 11 kabupaten dan 4 kota (15 wilayah) pada tahun 2007 hingga 2014. Langkah-langkah pengerjaan dalam penelitian ini adalah sebagai berikut :

1. Melakukan deskriptif statistik terhadap variabel respon dan prediktor.
2. Melakukan uji multikolinearitas untuk semua variabel prediktor menggunakan metode koefisien korelasi sampel. Rumus koefisien korelasi sampel yaitu [7]

$$r_{12} = \frac{NT \sum_{i=1}^N \sum_{t=1}^T X_{1it} X_{2it} - (\sum_{i=1}^N \sum_{t=1}^T X_{1it}) (\sum_{i=1}^N \sum_{t=1}^T X_{2it})}{\sqrt{[NT \sum_{i=1}^N \sum_{t=1}^T X_{1it}^2 - (\sum_{i=1}^N \sum_{t=1}^T X_{1it})^2] [NT \sum_{i=1}^N \sum_{t=1}^T X_{2it}^2 - (\sum_{i=1}^N \sum_{t=1}^T X_{2it})^2]}} \quad (1)$$

Pada Persamaan (1), $i = 1, 2, \dots, N$ banyaknya wilayah dan $t = 1, 2, \dots, T$ banyaknya periode waktu.

3. Melakukan uji Chow untuk memilih model regresi data panel antara MEU dan MET. Adapun, bentuk umum dari MEU adalah sebagai berikut [3]

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

dengan Y adalah variabel respon, β_0 merupakan nilai konstanta, β_K menyatakan nilai koefisien variabel prediktor ke $-K$, dan nilai residual dilambangkan oleh ε_{it} . Selanjutnya, dalam MET perbedaan karakteristik dari masing-masing wilayah maupun periode waktu diinterpretasikan pada nilai konstanta yang berbeda. Bentuk umum MET adalah [3]

$$Y_{it} = \beta_{0i} + \beta_1 X_{1it} + \beta_2 X_{2it} + \dots + \beta_K X_{Kit} + \varepsilon_{it} \quad (3)$$

dengan nilai $\beta_{0i} = \pi_0 + \mu_2 D_2 + \dots + \mu_N D_N$, π_0 adalah konstanta wilayah pembeda, μ_N nilai efek spesifikasi wilayah ke $-N$, dan D_N adalah variabel *dummy*.

Hipotesis penelitian dalam uji Chow dapat dituliskan sebagai berikut

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_N = 0 \quad (\text{MEU})$$

$$H_1 : \text{minimal ada satu } \mu_i \neq 0 \quad (\text{MET}).$$

Statistik uji Chow dapat dituliskan sebagai

$$F_{hitung} = \frac{\frac{JKR1 - JKR2}{(N - 1)}}{\frac{JKR2}{NT - N - K}} \quad (4)$$

dengan $JKR1$ merupakan jumlah kuadrat residual MEU, sedangkan $JKR2$ merupakan jumlah kuadrat residual MET. Kriteria penolakannya adalah jika nilai $F_{hitung} \geq F_{tabel} = F_{(N-1, NT-N-K)}$ atau $p - value \leq 0.05$, maka hipotesis H_0 ditolak, artinya MET merupakan model yang paling sesuai dengan kondisi data dikarenakan adanya unsur perbedaan karakteristik model dari masing-masing wilayah maupun waktu. Namun, jika $F_{hitung} < F_{tabel} = F_{(N-1, NT-N-K)}$ atau $p - value > 0.05$, maka hipotesis H_0 diterima. Pada MET komponen waktu, indeks i diubah menjadi t dan $F_{tabel} = F_{(T-1, NT-T-K)}$.

4. Setelah uji Chow, langkah selanjutnya adalah melakukan Uji Hausman untuk memilih model regresi data panel yang paling sesuai antara MEA dan MET. Prinsip uji Hausman adalah untuk mengetahui ada atau tidak korelasi antara efek spesifikasi wilayah maupun efek spesifikasi waktu terhadap variabel prediktor. Dalam MEA, perbedaan karakteristik dari masing-masing wilayah maupun periode waktu diakibatkan oleh efek spesifikasi wilayah dan efek spesifikasi waktu yang dinyatakan sebagai komponen residual model. Bentuk umum MEA yaitu [3]

$$Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \dots + \beta_K X_{Kit} + (\mu_i + \varepsilon_{it}) \quad (5)$$

Hipotesis penelitian dari uji Hausman adalah

H_0 : korelasi $(\mu_i, X_{it}) = 0$ (MEA)

H_1 : korelasi $(\mu_i, X_{it}) \neq 0$ (MET)

Statistik uji yang digunakan dalam uji Hausman menggunakan kriteria Wald sebagai berikut [3]

$$\widehat{W}_{hitung} = (\widehat{\beta}_{MET} - \widehat{\beta}_{MEA})' [var(\widehat{\beta}_{MET} - \widehat{\beta}_{MEA})]^{-1} (\widehat{\beta}_{MET} - \widehat{\beta}_{MEA}). \quad (6)$$

Pada Persamaan (6), $\widehat{\beta}_{MET}$ adalah vektor kolom nilai koefisien MET, sedangkan $\widehat{\beta}_{MEA}$ ialah vektor kolom nilai koefisien MEA. Kriteria penolakannya yaitu jika $\widehat{W}_{hitung} \geq \chi^2_{tabel} = \chi^2_{(\alpha, K)}$ atau $p - value \leq 0.05$, maka tolak hipotesis H_0 , yang artinya MET merupakan model yang paling sesuai digunakan dikarenakan efek spesifikasi wilayah atau waktu berkorelasi dengan variabel prediktor (X_{it}). Sebaliknya, jika $\widehat{W}_{hitung} < \chi^2_{tabel} = \chi^2_{(\alpha, K)}$ atau $p - value > 0.05$, maka terima hipotesis H_0 . Sama halnya dengan MET komponen waktu, dalam MEA komponen waktu, indeks i diubah menjadi t .

5. Melakukan pemilihan efek komponen satu arah.

Pada bagian ini, akan dipilih efek komponen satu arah antara komponen wilayah dan komponen waktu yang paling baik dalam menjelaskan keberagaman dari variabel respon. Hal ini dilakukan dengan menggunakan nilai koefisien determinasi majemuk sebagai berikut [7]

$$R^2 = 1 - \frac{\sum_{i=1}^N \sum_{t=1}^T (Y_{it} - \widehat{Y}_{it})^2}{\sum_{i=1}^N \sum_{t=1}^T (Y_{it} - \bar{Y}_{it})^2}, \quad (7)$$

Komponen satu arah yang paling baik dalam menjelaskan tingkat keragaman dari variabel respon dipilih berdasarkan nilai R^2 yang paling besar antara model regresi data panel terpilih untuk komponen wilayah dan komponen waktu.

6. Setelah model regresi data panel dengan efek komponen satu arahnya terpilih, maka selanjutnya melakukan uji asumsi klasik yang terdiri dari :

- a. Uji Normalitas.

Data-data pengamatan dari sampel yang diambil perlu diuji kembali apakah berasal dari populasi yang berdistribusi normal atau tidak. Hipotesis penelitiannya adalah sebagai berikut :

H_0 : $E(\varepsilon_{it}) = 0$ (data berdistribusi normal).

H_1 : $E(\varepsilon_{it}) \neq 0$ (data tidak berdistribusi normal).

Menurut [7], statistik uji normalitas menggunakan rumus Jarque-Bera yaitu

$$JB = NT \left(\frac{S_k^2}{6} + \frac{(Kr - 3)^2}{24} \right) \quad (8)$$

Dari Persamaan (8), S_k (*skewness*) merupakan kemiringan, sedangkan Kr (*kurtosis*) merupakan ketinggian. Kriteria penolakannya adalah jika nilai $JB < \chi^2_{tabel} = \chi^2_{(\alpha, 2)}$ atau $p - value > 0,05$, maka terima hipotesis H_0 , artinya data-data pengamatan dari sampel berasal dari populasi yang berdistribusi normal. Namun, jika $JB \geq \chi^2_{tabel} = \chi^2_{(\alpha, 2)}$ atau $p - value \leq 0,05$, maka tolak hipotesis H_0 . Salah satu cara untuk mengatasi data yang tidak berdistribusi normal yaitu dengan melakukan transformasi logaritma natural (ln).

- b. Uji Heteroskedastisitas.

Varians residual yang tidak sama ($var(\varepsilon_{it}) \neq \sigma^2$) antar pengamatan akan berpengaruh pada nilai standar residual yang dihasilkan.. Bentuk hipotesis uji heteroskedastisitas yaitu

H_0 : $Var(\varepsilon_{it}) = \sigma^2$ (homoskedastisitas).

H_1 : $Var(\varepsilon_{it}) \neq \sigma^2$ (heteroskedastisitas).

Statistik uji yang digunakan adalah Lagrange *Multiplier* (LM) sebagai berikut [6]

$$LM = \frac{T}{2} \sum_{i=1}^N \left(\frac{\sigma_i^2}{\sigma^2} - 1 \right)^2 \quad (9)$$

Pada Persamaan (9), T banyaknya periode waktu, σ_i^2 varians residual persamaan ke $-i$, dan σ^2 varians residual sistem. Kriteria penolakannya adalah tolak hipotesis H_0 jika nilai $LM_{hitung} \geq \chi^2_{(\alpha; N-1)}$ atau $p\text{-value} \leq 0,05$. Sebaliknya, jika $LM_{hitung} < \chi^2_{(\alpha; N-1)}$ atau $p\text{-value} > 0,05$, maka terima hipotesis H_0 yang artinya varians residual konstan. Pembobotan *cross section weight* (komponen wilayah) atau *period weight* (komponen waktu) digunakan untuk mengatasi heteroskedastisitas.

c. Uji Autokorelasi.

Uji asumsi klasik selanjutnya adalah uji autokorelasi yang bertujuan untuk mengetahui apakah residual model bersifat saling bebas (tidak saling mempengaruhi). Hipotesis penelitian dalam uji autokorelasi yaitu

$H_0: Cov(\varepsilon_{it}, \varepsilon_{i,t-1}) = 0$ (tidak ada autokorelasi).

$H_1: Cov(\varepsilon_{it}, \varepsilon_{i,t-1}) \neq 0$ (ada autokorelasi).

Menurut [8], autokorelasi dapat diketahui dengan menggunakan statistik uji Durbin Watson (DW) sebagai berikut

$$DW = \frac{\sum_{i=1}^N \sum_{t=2}^T (\varepsilon_{it} - \varepsilon_{i,t-1})^2}{\sum_{i=1}^N \sum_{t=1}^T \varepsilon_{it}^2} \quad (10)$$

Adapun kriteria penolakan dalam uji autokorelasi yaitu :

a) Tidak ada autokorelasi positif dan negatif jika $d_U < DW < 4 - d_U$.

b) Ada autokorelasi positif jika $0 < DW < d_L$.

c) Ada autokorelasi negatif bila $4 - d_L < DW < 4$.

d) Tidak ada keputusan bila $d_L \leq DW < d_U$ dan $4 - d_U \leq DW \leq 4 - d_L$.

Notasi d_U merupakan nilai batas atas, sedangkan d_L menyatakan nilai batas bawah yang nilainya dapat ditentukan melalui tabel DW. Pelanggaran asumsi klasik autokorelasi dapat diatasi dengan melakukan pembobotan *Seemingly Unrelated Regression* (SUR).

7. Setelah data dalam model terpilih regresi data panel terbebas dari pelanggaran asumsi klasik, maka selanjutnya dilakukan uji signifikansi parameter yang terdiri dari :

a. Uji Simultan.

Untuk mengetahui apakah ada minimal satu variabel prediktor di dalam model regresi data panel yang mempengaruhi variabel respon, maka dilakukan uji simultan. Hipotesis penelitian yang digunakan dalam uji simultan yaitu

$H_0: \beta_1 = \beta_2 = \dots = \beta_K = 0$ (variabel-variabel prediktor tidak berpengaruh signifikan).

H_1 : minimal terdapat satu $\beta_k \neq 0$ (variabel-variabel prediktor berpengaruh signifikan).

Statistik uji dalam uji simultan yaitu [10]

$$F_{hitung} = \frac{\sum_{i=1}^N \sum_{t=1}^T (\hat{Y}_{it} - \bar{Y}_{it})^2}{(N + K - 1)} \div \frac{\sum_{i=1}^N \sum_{t=1}^T (Y_{it} - \hat{Y}_{it})^2}{NT - N - K} \quad (11)$$

Jika nilai $F_{hitung} \geq F_{tabel} = F_{\alpha(N+K-1, NT-N-K)}$ atau $p\text{-value} \leq 0.05$, maka tolak hipotesis H_0 artinya minimal ada satu variabel prediktor yang berpengaruh signifikan terhadap variabel respon di dalam model regresi data panel. Sebaliknya, jika $F_{hitung} < F_{tabel} = F_{\alpha(N+K-1, NT-N-K)}$ atau $p\text{-value} > 0.05$, maka terima hipotesis H_0 . Pada komponen waktu nilai $F_{tabel} = F_{\alpha(T+K-1, NT-T-K)}$.

b. Uji Parsial.

Uji parsial dilakukan untuk mengetahui hubungan atau pengaruh dari masing-masing variabel prediktor terhadap variabel respon secara terpisah. Hipotesis penelitiannya adalah

$H_0: \beta_k = 0$ (variabel prediktor ke $-k$ tidak berpengaruh signifikan).

$H_1: \beta_k \neq 0$ (variabel prediktor ke $-k$ berpengaruh signifikan).

Statistik uji parsial adalah sebagai berikut [4]

$$t_{hitung} = \frac{\beta_k}{se(\beta_k)} \quad (12)$$

Hipotesis H_0 ditolak jika $t_{hitung} \geq t_{tabel} = t_{(\frac{\alpha}{2}, NT-N-K)}$ atau nilai $p\text{-value} \leq 0,05$, yang artinya $\hat{\beta}_k$ berpengaruh signifikan terhadap variabel respon. Sebaliknya, terima hipotesis H_0 jika $t_{hitung} < t_{tabel} = t_{(\frac{\alpha}{2}, NT-N-K)}$ atau nilai $p\text{-value} > 0.05$.

8. Seleksi variabel prediktor menggunakan prosedur eliminasi mundur untuk menentukan variabel-variabel prediktor yang berpengaruh signifikan di dalam model terbaik regresi data panel.

9. Interpretasi model terbaik regresi data panel dengan efek komponen satu arah yang terpilih.

3. Hasil dan Pembahasan

3.1. Deskriptif Statistik

Langkah awal sebelum dilakukan proses pengolahan data adalah melakukan deskriptif statistik. Karakteristik dari masing-masing variabel prediktor dapat diinformasikan melalui deskriptif statistik seperti halnya nilai IPM dan semua variabel prediktor yang diduga berpengaruh signifikan terhadap perubahan nilai IPM di kabupaten dan kota pada Provinsi Sumatera Selatan dari tahun 2007 hingga 2014. Deskriptif statistik yang digunakan berupa nilai maksimum, minimum, dan rata-rata dari setiap variabel seperti yang ditampilkan dalam Tabel 3.1.

Tabel 3.1 Deskriptif Statistik Variabel-Variabel Penelitian

Variabel	Notasi	Nilai Maksimum	Nilai Minimum	Nilai Rata-Rata	Satuan
Indeks Pembangunan Manusia (IPM)	Y	76.02	58.88	66.892	-
Tingkat keluhan kesehatan masyarakat	X_1	50.02	11.3	28.62	Persen
Jumlah rumah tangga dengan akses air bersih	X_2	98.37	12.08	49.014	Persen
Angka partisipasi SMA	X_3	82.65	31.57	57.648	Persen
Angka melek huruf	X_4	99.91	93.79	97.271	Persen
Tingkat partisipasi angkatan kerja	X_5	80.89	54.86	70.053	Persen
Nilai produk domestik regional bruto	X_6	78173	538	10476.95	Miliar rupiah

Pada Tabel 3.1, rentang nilai antara nilai maksimum dan minimum dapat dikatakan sangat jauh kecuali pada variabel X_4 (angka melek huruf). Kota Palembang selalu mendominasi perolehan tertinggi untuk seluruh variabel kecuali variabel X_1 (tingkat keluhan kesehatan masyarakat) dan X_5 (tingkat partisipasi angkatan kerja). Sedangkan kabupaten atau kota yang memiliki nilai terendah untuk masing-masing variabel selalu mengalami perubahan. Berdasarkan nilai rata-rata, angka melek huruf pada kabupaten dan kota di Provinsi Sumatera Selatan dari tahun 2007 hingga 2014 tergolong sangat tinggi karena melebihi 90 %.

3.2. Uji Multikolinearitas

Uji multikolinearitas bertujuan memeriksa ada atau tidak korelasi antara sesama variabel prediktor di dalam model regresi data panel yang seharusnya variabel-variabel prediktor tersebut hanya berkorelasi dengan variabel responnya saja. Multikolinearitas dapat mengakibatkan hasil uji parsial lebih sering menerima H_0 , sehingga variabel prediktornya banyak yang tidak berpengaruh signifikan. Multikolinearitas dapat dideteksi melalui metode koefisien korelasi sampel (r). Nilai r dalam penelitian ini diperoleh dengan bantuan *software* Eviews 9 yang ditampilkan dalam Tabel 3.2.

Tabel 3.2 Koefisien Korelasi Sampel

	X_1	X_2	X_3	X_4	X_5	X_6
X_1	-	-0.057	0.188	0.091	0.023	-0.075
X_2	-0.057	-	0.365	0.16	-0.578	0.547
X_3	0.188	0.365	-	0.467	-0.368	0.106
X_4	0.091	0.16	0.467	-	-0.241	0.181
X_5	0.023	-0.578	-0.368	-0.241	-	-0.298
X_6	-0.075	0.547	0.106	0.181	-0.298	-

Menurut Gujarati dan Dawn [8], terjadinya multikolinearitas antar dua variabel prediktor yang berbeda ditandai dengan nilai $|r| > 0.8$, sehingga dari Tabel 3.2, nilai semua r pada setiap kolom yaitu $|r| < 0.8$. Jadi, semua variabel prediktor dalam penelitian ini terbebas dari masalah multikolinearitas.

3.3. Pemilihan Model Regresi Data Panel Komponen Wilayah

Uji Chow dilakukan untuk mengetahui ada atau tidaknya perbedaan karakteristik dari model regresi data panel antar kabupaten dan kota yang dapat diketahui dari perbedaan konstanta model regresinya. Uji Chow melibatkan MEU dan MET. Dikarenakan nilai $F_{hitung} = 3.992 > F_{0,05(14;99)} = 1.79$ dan $p\text{-value} = 0.000 < 0.05$, maka keputusannya adalah tolak hipotesis H_0 yang artinya MET terpilih dengan taraf kepercayaan 95 %, sehingga ada perbedaan karakteristik dari model regresi data panel IPM antar

wilayah yang disebabkan oleh perbedaan nilai konstanta dari masing-masing model persamaan regresinya.

Berdasarkan uji Hausman yang melibatkan MEA dan MET untuk mengetahui ada atau tidaknya korelasi antara efek spesifikasi wilayah dengan variabel prediktor. Diperoleh bahwa nilai $W_{hitung} = 36.453 > \chi^2_{tabel} = 12.59$ dan $p-value = 0.000 < 0.05$ yang berarti tolak hipotesis H_0 , sehingga ada korelasi antara efek spesifikasi wilayah dengan variabel prediktor, sehingga model regresi yang paling sesuai adalah MET.

3.4. Pemilihan Model Regresi Data Panel Komponen Waktu

Sama halnya dengan komponen wilayah, pada komponen waktu, uji Chow melibatkan MEU dan MET untuk mengetahui ada atau tidak perbedaan karakteristik di periode waktu pengamatan. Diperoleh bahwa nilai $F_{hitung} = 27.439 > F_{tabel} = 2.1$ dan $p-value = 0.000 < 0.05$, maka keputusannya adalah tolak hipotesis H_0 . Hal ini berarti MET terpilih, sehingga pada taraf kepercayaan 95 % ada perbedaan karakteristik dari model regresi data panel komponen waktu pada periode waktu pengamatan yang disebabkan oleh perbedaan nilai konstanta modelnya.

Uji Hausman komponen waktu dilakukan untuk memilih antara MEA dan MET. Berdasarkan uji Hausman, diperoleh nilai $W_{hitung} = 113.931 > \chi^2_{tabel} = 12.592$ dan $p-value = 0.000 < 0.05$ yang artinya tolak hipotesis H_0 , sehingga ada korelasi antara efek spesifikasi waktu dengan variabel prediktor, sehingga model regresi yang paling sesuai adalah MET.

3.5. Pemilihan Efek Komponen Satu Arah

Setelah model regresi data panel terpilih untuk komponen wilayah dan komponen waktu, selanjutnya akan dilakukan pemilihan efek komponen satu arah antara komponen wilayah dan komponen waktu yang bertujuan untuk mengetahui efek komponen satu arah yang paling baik dalam menjelaskan tingkat keragaman dari nilai IPM. Berdasarkan nilai R^2 (koefisien determinasi), maka efek komponen satu arah yang paling baik dalam menjelaskan tingkat keragaman nilai IPM pada seluruh kabupaten dan kota di Provinsi Sumatera Selatan dari tahun 2007 hingga 2014 adalah model regresi data panel MET komponen waktu dikarenakan $R^2_{komponen waktu} = 0,807 > R^2_{komponen wilayah} = 0,643$.

3.6. Uji Asumsi Klasik

3.6.1. Uji Normalitas

Agar dapat dilakukan uji dengan menggunakan statistik parametrik pada data yang berskala interval atau rasio, dibutuhkan terpenuhinya syarat asumsi normalitas pada data tersebut. Uji asumsi klasik normalitas bertujuan untuk mengetahui data pengamatan yang diambil berasal dari sampel yang berdistribusi normal atau tidak. Berdasarkan uji normalitas dengan menggunakan metode Jarque Bera, dan bantuan *software* Eviews 9, diperoleh nilai $JB = 1.509 < \chi^2_{(0,05;2)} = 5.991$ dan $p-value = 0.470 > 0.05$, maka keputusannya adalah terima hipotesis H_0 yang artinya dengan taraf kepercayaan sebesar 95 %, data dalam model regresi data panel terpilih MET komponen waktu berasal dari sampel yang berdistribusi normal.

3.6.2. Uji Heteroskedastisitas

Setelah diketahui bahwa model regresi data panel terpilih MET komponen waktu memiliki data yang berdistribusi normal, maka langkah berikutnya adalah mendeteksi kehomogenan varians residualnya dengan menggunakan uji heteroskedastisitas. Varians residual yang tidak sama (heteroskedastisitas) akan menyebabkan estimator MET komponen waktu tidak mempunyai varians yang minimum lagi sehingga uji signifikansi yang dilakukan secara parsial akan menghasilkan kesimpulan yang tidak benar. Heteroskedastisitas dalam MET komponen waktu dapat dideteksi dengan menggunakan metode Lagrange *Multiplier* (LM). Dikarenakan nilai $LM_{hitung} = 7.961 < \chi^2_{(0,05;6)} = 12.592$ dan $p-value = 0.1665 > 0.05$, sehingga keputusannya adalah terima hipotesis H_0 yang berarti dengan taraf kepercayaan sebesar 95 %, variansi residual dari model regresi data panel terpilih MET komponen waktu adalah konstan (homoskedastisitas).

3.6.3. Uji Autokorelasi

Uji autokorelasi bertujuan untuk mengetahui ada atau tidaknya korelasi antar residual pada waktu pengamatan saat ini (t) dengan waktu pengamatan sebelumnya ($t - 1$) yang saling mempengaruhi. Autokorelasi dapat diketahui dengan menggunakan metode Durbin Watson (DW). Diperoleh nilai statistik

$DW = 0.569 < d_L = 1.599$, sehingga keputusannya adalah tolak hipotesis H_0 . Hal ini berarti bahwa dengan taraf kepercayaan sebesar 95 %, MET komponen waktu memiliki masalah autokorelasi positif. Autokorelasi akan berpengaruh pada nilai koefisien penduga parameter yang dihasilkan menjadi tidak tepat lagi (bias) dikarenakan nilai standar residual yang dihasilkan menjadi lebih besar, sehingga untuk uji parsial jika terjadi autokorelasi diragukan kebenarannya. Untuk mengatasi masalah autokorelasi pada model regresi data panel terpilih MET komponen waktu, dapat diatasi dengan menggunakan pembobotan *Seemingly Unrelated Regression* (SUR) guna mengoreksi nilai standar residual yang dihasilkan. Perubahan nilai DW ditampilkan dalam Tabel 3.3.

Tabel 3.3 Nilai DW Sebelum dan Setelah Pembobotan

Sebelum Pembobotan	Setelah Pembobotan
0.569	1.82

Pada Tabel 3.3, nilai DW setelah pembobotan SUR waktu adalah 1.82 yang mengindikasikan bahwa model regresi data panel terpilih MET komponen waktu memiliki residual yang saling bebas atau terbebas dari masalah autokorelasi positif dikarenakan nilai DW setelah pembobotan teletak pada daerah penerimaan hipotesis H_0 yaitu $d_U = 1.808 < DW = 1.82 < 4 - d_U = 2.192$.

3.7. Uji Signifikansi Parameter

3.7.1. Uji Simultan

Uji simultan dilakukan untuk mengetahui ada atau tidaknya minimal satu variabel prediktor yang berpengaruh signifikan terhadap variabel respon. Berdasarkan statistik uji F , diperoleh nilai $F_{hitung} = 79.043 > F_{0,05;(13;106)} = 1.81$ dan $p\text{-value} = 0.000 < 0.05$, maka dapat disimpulkan untuk tolak hipotesis H_0 yang berarti minimal ada satu variabel prediktor yang berpengaruh signifikan terhadap variabel respon. Jadi, dengan taraf kepercayaan 95 %, secara simultan variabel-variabel prediktor berpengaruh signifikan terhadap variabel respon.

3.7.2. Uji Parsial

Setelah hasil uji simultan ternyata minimal ada satu variabel prediktor memiliki pengaruh yang signifikan terhadap variabel respon, maka dapat dilanjutkan pada uji parsial untuk menentukan variabel-variabel prediktor manakah yang benar-benar memiliki pengaruh yang signifikan terhadap perubahan nilai IPM secara parsial. Hasil uji parsial ditampilkan pada Tabel 3.4.

Tabel 3.4 Hasil Uji Parsial

Variabel	t_{hitung}	$P\text{-value}$
X_1	-1.73	0.087
X_2	10.202	0.000 *
X_3	9.999	0.000 *
X_4	4.06	0.000 *
X_5	4.168	0.000 *
X_6	1.118	0.266

Pada Tabel 3.4, variabel prediktor yang berpengaruh signifikan terhadap variabel respon pada penelitian ini dengan taraf signifikansi 5 % adalah X_2 , X_3 , X_4 , dan X_5 yang ditandai dengan *, sedangkan variabel X_1 dan X_6 tidak berpengaruh signifikan.

3.8. Prosedur Eliminasi Mundur

Dikarenakan pada hasil uji parsial sebelumnya ada dua variabel prediktor yang tidak berpengaruh signifikan terhadap variabel respon, maka perlu dilakukan proses penyeleksian variabel-variabel prediktor guna memastikan bahwa variabel-variabel prediktor yang akan dimasukkan di dalam model terbaik regresi data panel semuanya memiliki pengaruh yang signifikan terhadap variabel respon (IPM). Prosedur eliminasi mundur dapat digunakan untuk melakukan hal tersebut dengan meregresikan semua variabel prediktor kemudian mengeluarkan satu per satu variabel prediktor yang tidak berpengaruh signifikan pada taraf signifikansi 5 %.

Pada Tabel 3.4, variabel X_6 dikeluarkan terlebih dahulu dikarenakan memiliki nilai $p\text{-value}$ tertinggi. Setelah variabel X_6 dikeluarkan, maka perlu dilakukan uji Chow dan Hausman kembali untuk

memastikan bahwa MET komponen waktu terboboti SUR adalah model yang masih terpilih. Hasil uji Chow tanpa variabel X_6 yaitu $F_{hitung} = 17.187 > F_{0,05(7;107)} = 2.1$ dan $p\text{-value} = 0.000 < 0.05$ yang berarti MET terpilih. Setelah itu, hasil uji Hausman yaitu nilai $W_{hitung} = 69.342 > \chi_{0,05(5)}^2 = 11.07$ dan $p\text{-value} = 0.000 < 0.05$, sehingga yang terpilih adalah MET komponen waktu terboboti SUR. Selanjutnya, hasil uji simultan yaitu $F_{hitung} = 67.952 > F_{0,05(12;107)} = 1.84$ dan $p\text{-value} = 0.000 < 0.05$ yang mengindikasikan bahwa minimal ada satu variabel prediktor yang berpengaruh signifikan terhadap perubahan nilai IPM, sehingga dapat dilanjutkan dengan uji parsial (Tabel 3.5).

Tabel 3.5 Hasil Uji Parsial Tanpa Variabel X_6

Variabel	t_{hitung}	$p\text{-value}$
X_1	-0.707	0.481
X_2	10.199	0.000 *
X_3	9.072	0.000 *
X_4	5.178	0.000 *
X_5	5.11	0.000 *

Variabel prediktor yang tidak berpengaruh signifikan terhadap variabel respon (tanpa tanda “**”) dalam taraf signifikansi 5% adalah variabel X_1 , sehingga untuk proses selanjutnya variabel X_1 tidak diikuti dalam MET komponen waktu terboboti SUR. Setelah variabel X_1 (tingkat keluhan kesehatan masyarakat) dan X_6 (nilai produk domestik regional bruto) dikeluarkan, hasil uji Chow yaitu nilai $F_{hitung} = 17.244 > F_{0,05(7;108)} = 2.1$ dan $p\text{-value} = 0,000 < 0.05$, sehingga tolak H_0 yang berarti MET terpilih. Selanjutnya, berdasarkan uji Hausman diperoleh nilai $W_{hitung} = 18.441 > \chi_{0,05(4)}^2 = 9.488$ dan $p\text{-value} = 0.001 < 0.05$ yang berarti tolak hipotesis H_0 sehingga MET komponen waktu terboboti SUR masih menjadi model regresi data panel terpilih. Adapun hasil uji simultan yakni nilai $F_{hitung} = 73.577 > F_{0,05(11;108)} = 1.88$ dan $p\text{-value} = 0.000 < 0.05$ yang berarti bahwa minimal ada satu variabel prediktor berpengaruh signifikan terhadap nilai IPM, sehingga dapat dilanjutkan untuk uji parsial pada masing-masing variabel prediktor (Tabel 3.6).

Tabel 3.6 Hasil Uji Parsial Tanpa Variabel X_1 dan X_6

Variabel	Koefisien	t_{hitung}	$p\text{-value}$
X_2	0.086	9.55	0.000 *
X_3	0.051	8.367	0.000 *
X_4	0.362	5.886	0.000 *
X_5	0.074	5.107	0.000 *

Hasil uji parsial di dalam Tabel 3.6 memperlihatkan bahwa semua variabel prediktor berpengaruh signifikan terhadap perubahan nilai IPM pada taraf signifikansi 5 %, sehingga tanpa memasukkan variabel X_1 dan X_6 dapat diperoleh model terbaik regresi data panel yaitu MET komponen waktu terboboti SUR dengan melibatkan hanya variabel $X_2, X_3, X_4,$ dan X_5 yang berpengaruh signifikan terhadap nilai IPM.

3.9. Interpretasi Model Terbaik Regresi Data Panel

Dari Tabel 3.6, bentuk umum model terbaik regresi data panel yaitu MET komponen waktu terboboti SUR untuk nilai IPM dengan konstanta tahun pembeda $\tau_0 = 29.689$ adalah

$$\hat{Y}_{it} = \hat{\beta}_{0t} + 0.086X_{2it} + 0.051X_{3it} + 0.362X_{4it} + 0.074X_{5it} \quad (13)$$

Indeks $i = 1, 2, \dots, 15$ menyatakan banyaknya wilayah dan $t = 1, 2, \dots, 8$ menyatakan banyaknya tahun. Adapun nilai konstanta MET komponen waktu terboboti SUR ialah $\hat{\beta}_{0t} = 29,689 + \theta_t$ dengan θ_t merupakan efek spesifikasi waktu yang diinterpretasikan sebagai nilai koefisien variabel *dummy* dari tahun yang bukan sebagai tahun pembeda. Pemilihan tahun pembeda dilakukan secara bebas [6], sehingga yang dipilih ialah tahun 2007, maka $\tau_0 = 29.689$ menjadi konstanta MET komponen waktu terboboti SUR tahun 2007. Maka, bentuk model terbaik regresi data panel yaitu MET komponen waktu terboboti SUR untuk nilai IPM di sepanjang tahun 2007 pada seluruh kabupaten dan kota di Provinsi Sumatera Selatan adalah

$$\hat{Y}_{i1} = 29.689 + 0.086X_{2i1} + 0.051X_{3i1} + 0.362X_{4i1} + 0.074X_{5i1}. \quad (14)$$

Persamaan (14) memiliki nilai $R^2 = 0,882$ yang berarti kemampuan seluruh variabel prediktor untuk menjelaskan tingkat keragaman nilai IPM adalah sebesar 88.2 % sedangkan sisanya sebesar 11.8 % dijelaskan oleh variabel-variabel prediktor lain yang tidak dimasukkan ke dalam Persamaan (14). Nilai konstanta sebesar 29.689 yang berarti jika semua variabel prediktor bernilai nol, maka nilai IPM pada seluruh kabupaten dan kota di Provinsi Sumatera Selatan sepanjang tahun 2007 ialah 29.689. Variabel X_2 memiliki arah korelasi positif seperti pada penelitian [9] dan berpengaruh signifikan terhadap perubahan

nilai IPM, sehingga setiap kenaikan 1 satuan banyaknya rumah tangga yang menggunakan akses air bersih, maka akan menaikkan nilai IPM sebesar 0.086 dengan asumsi variabel prediktor lainnya bernilai konstan. Sejalan dengan hal tersebut, variabel X_3 memiliki arah korelasi positif dan hal ini sesuai dengan penelitian dari [3]. Dari hasil uji parsial, variabel X_3 memiliki hubungan atau pengaruh yang signifikan terhadap perubahan nilai IPM, sehingga nilai IPM akan bertambah sebesar 0.051 jika terjadi kenaikan setiap 1 satuan variabel angka partisipasi Sekolah Menengah Atas (SMA) dengan asumsi ketiga variabel prediktor lainnya bernilai konstan. Selain itu, arah korelasi positif pada variabel X_4 juga sesuai dengan penelitian dari [10] dan berdasarkan uji parsial variabel X_4 berpengaruh signifikan terhadap perubahan nilai IPM, sehingga dengan mengasumsikan ketiga variabel prediktor lain bernilai konstan, maka setiap pertambahan 1 satuan variabel angka melek huruf, maka akan mengakibatkan nilai IPM naik sebesar 0.362. Selanjutnya, variabel prediktor X_5 yang juga berpengaruh signifikan terhadap perubahan nilai IPM dan memiliki arah korelasi yang sama dengan penelitian [9] yaitu positif, sehingga jika setiap kenaikan 1 satuan variabel tingkat partisipasi angkatan kerja maka akan berdampak pada naiknya nilai IPM sebesar 0.074 dengan mengasumsikan ketiga variabel prediktor lainnya bernilai konstan. Adapun, pada tahun yang bukan merupakan tahun pembeda (selain tahun 2007) dibutuhkan efek spesifikasi waktu guna menghasilkan nilai konstanta dari MET komponen waktu terboboti SUR, sehingga dari hal tersebut perbedaan karakteristik dari masing-masing periode waktu terlihat dari sebesar besar perbedaan konstanta di masing-masing wilayah berbeda dengan nilai konstanta tahun pembeda. Nilai efek spesifikasi waktu (θ_t) dan nilai konstantanya ($\beta_{0t} = 29,689 + \theta_t$) dapat dilihat pada Tabel 3.7.

Tabel 3.7 Nilai Efek Spesifikasi Waktu dan Konstanta

t	Tahun	θ_t	β_{0t}
2	2008	4.227	33.916
3	2009	4.122	33.811
4	2010	-3.084	26.605
5	2011	-2.598	27.091
6	2012	-2.246	27.443
7	2013	-2.183	27.506
8	2014	-2.114	27.575

Misal akan dilihat MET komponen waktu terboboti SUR pada tahun 2013 ($t = 7$), sehingga bentuk model regresi data panelnya yaitu

$$\hat{Y}_{i7} = 29.689 + \theta_2 D_2 + \theta_3 D_3 + \dots + \theta_8 D_8 + 0.086 X_{2i7} + 0.051 X_{3i7} + 0.362 X_{4i7} + 0.074 X_{5i7} \quad (15)$$

Dikarenakan waktu pengamatan yang ingin diketahui adalah tahun 2013, maka selain dari tahun 2013 nilai variabel *dummy* adalah nol (t lainnya), sedangkan nilai variabel *dummy* tahun 2013 yaitu $D_7 = 1$, sehingga dengan mensubstitusikan $D_7 = 1$ dan $\theta_7 = -2.183$ ke Persamaan (15), maka diperoleh bentuk model terbaik regresi data panel yaitu MET komponen waktu terboboti SUR untuk nilai IPM di sepanjang tahun 2013 pada seluruh kabupaten dan kota di Provinsi Sumatera Selatan adalah

$$\hat{Y}_{i7} = 27.506 + 0.086 X_{2i7} + 0.051 X_{3i7} + 0.362 X_{4i7} + 0.074 X_{5i7} \quad (16)$$

Dari Persamaan (16), nilai konstanta menandakan bahwa apabila semua variabel prediktor bernilai nol, maka nilai IPM untuk seluruh kabupaten dan kota pada Provinsi Sumatera Selatan di sepanjang tahun 2013 sebesar 27.506. Interpretasi untuk variabel banyaknya rumah tangga yang menggunakan air bersih (X_2), angka partisipasi SMA (X_3), angka melek huruf (X_4), dan tingkat partisipasi angkatan kerja (X_5) sama halnya pada interpretasi di Persamaan (14) dikarenakan nilai koefisien dan arah korelasi semua variabel prediktor untuk masing-masing tahun sama.

4. Simpulan

Dalam penelitian ini, dibahas tiga model regresi data panel dalam memodelkan nilai IPM di Provinsi Sumatera Selatan, yaitu MEU, MET, dan MEA. Berdasarkan beberapa uji statistik yang dilakukan, diperoleh bahwa MET merupakan model regresi data panel yang paling sesuai dengan data nilai IPM di Provinsi Sumatera Selatan. Lebih spesifik lagi, MET komponen waktu terboboti SUR merupakan model regresi data panel terbaik yang dapat menjelaskan keberagaman nilai IPM di Provinsi Sumatera Selatan.

Selanjutnya, berdasarkan model terbaik tersebut diperoleh bahwa persentase banyaknya rumah tangga yang menggunakan akses air bersih (X_2), persentase angka partisipasi SMA (X_3), persentase angka melek huruf (X_4), dan persentase tingkat partisipasi angkatan kerja (X_5) adalah faktor-faktor yang signifikan berpengaruh terhadap nilai IPM di Provinsi Sumatera Selatan.

Daftar Pustaka

1. Badan Pusat Statistik Provinsi Sumatera Selatan (BPSPSS). 2015. *Sumatera Selatan dalam Angka 2015*. Palembang : Badan Pusat Statistik Provinsi Sumatera Selatan.
2. Badan Pusat Statistik Provinsi Sumatera Selatan (BPSPSS). 2016. *Buku Saku Data dan Indikator Sosial Sumatera Selatan 2006-2015*. Palembang: Badan Pusat Statistik Provinsi Sumatera Selatan.
3. Baltagi, B. H. 2010. *Econometric Analysis of Panel Data 4th edition*. West Sussex : John Wiley & Sons, Ltd.
4. Destiluna F.G. dan Ismaini Z. 2015. Pengaruh dan Pemetaan Pendidikan, Kesehatan, serta UMKM terhadap Indeks Pembangunan Manusia di Jawa Timur Menggunakan Regresi Panel dan Biplot. *Jurnal Sains dan Seni ITS 4(2) : D-292 – D-298*.
5. Fajriyah, N dan Santi P.R. 2016. Pemodelan Faktor-Faktor yang Mempengaruhi Kemiskinan Kabupaten/Kota di Jawa Timur Menggunakan Regresi Data Panel. *Jurnal Sains dan Seni ITS 5(1) : D-45 – D-50*
6. Greene, W. H. 2003. *Econometric Analysis (5th edition)*. New Jersey : Prentice Hall International.
7. Gujarati, D. N. 2004. *Basic Econometrics 4th edition* . New York: The McGraw-Hill Companies.
8. Gujarati, D. N dan Dawn C. P. 2010. *Dasar-dasar Ekonometrika Edisi 5 (Terjemahan)*. New York: The McGraw-Hill Companies.
9. Melliana, A. dan Ismaini Z. 2013. Analisis Statistika Faktor yang Mempengaruhi Indeks Pembangunan Manusia di Kabupaten/Kota Provinsi Jawa Timur dengan Menggunakan Regresi Panel. *Jurnal Sains dan Seni POMITS 2(2) : D-237 – D-242*.
10. Pangestika, S. 2015. *Analisis Estimasi Model Regresi Data Panel dengan Pendekatan Common Effect Model (CEM), Fixed Effect Model (FEM), dan Random Effect Model (REM)*. Semarang : Jurusan Matematika FMIPA Universitas Negeri Semarang.
11. Setiawan, I. dan Krismanti T. W. 2015. Pendekatan Random Effect Regresi Data Panel Kajian Konsumsi Rokok Di Indonesia Tahun 2007 – 2012. *Seminar Nasional Matematika dan Pendidikan Matematika UNY 2015*. Yogyakarta.
12. Sugiharso dan Ester. 2007. *Determinan Investasi Portfolio Internasional Negara-Negara ASEAN, Amerika Serikat dan Jepang*. Depok : Fakultas Ekonomi Universitas Indonesia