

BAB 3

METODE PENELITIAN

3.1 Desain Penelitian

Penelitian ini menggunakan pendekatan penelitian kuantitatif dengan menggunakan data panel (*pool data*).

3.2 Metode Analisis Data

3.2.1 Analisis *Weighted Coefficient Variation (CVw)* atau *Index Williamson (Iw)*

Objek dari penelitian ini adalah disparitas dengan indikator PDRB per kapita atas dasar harga konstan tahun dasar 2000 menurut kabupaten/kota di Jawa Barat selama 6 (enam) tahun dari tahun 2003 sampai dengan tahun 2008. Periode pengambilan data tersebut dilakukan bersamaan dengan penerapan sistem desentralisasi di Indonesia khususnya Provinsi Jawa Barat, dimana sejak tahun 2001 terdapat pembentukan tiga kota baru yaitu Kota Tasikmalaya dan Kota Cimahi pada tahun 2001 dan Kota Banjar pada tahun 2002. Karena Kota Banjar baru terbentuk pada tahun tersebut maka untuk kemudahan memperoleh data dilakukan pembatasan pengambilan data dari tahun 2003 sampai dengan tahun 2008. Teknik analisis yang digunakan adalah analisis *Weighted Coefficient Variation (CVw)* atau *Indeks Williamson (Iw)* untuk melihat tingkat disparitas antar kecamatan di Kota Tasikmalaya pada periode Tahun 2002-2008.

Menurut Tambunan (2003), menghitung Indeks atau Koefisien Variasi Williamson adalah :

$$CV_w = \sqrt{\frac{\sum(Y_i - \bar{Y})^2 \left[\frac{1}{n} \right]}{\bar{Y}^2}}$$

Keterangan :

CV_w = Koefisien Variasi Williamson

Y_i = PDRB per kapita atas dasar harga konstan masing-masing kabupaten/ kota di provinsi yang bersangkutan

Y = PDRB per kapita atas dasar harga konstan provinsi yang bersangkutan

f_i = Penduduk pertengahan t:31 pada masing-masing kabupaten/kota di provinsi yang bersangkutan

n = Penduduk pertengahan tahun masing-masing kabupaten/kota di provinsi yang bersangkutan

Koefisien Variasi Williamson atau tingkat disparitas tersebut berada di antara nilai 0 dan 1, semakin mendekati nilai 0 berarti tingkat disparitas pendapatan antar kota/kabupaten di provinsi tersebut semakin rendah atau terjadi pemerataan pendapatan antar kabupaten/kota di provinsi tersebut, tetapi jika koefisien variasi Williamson mendekati nilai 1 berarti tingkat disparitas antar kota/kabupaten di provinsi tersebut semakin tinggi atau diindikasikan terjadi ketidak merataan pendapatan kabupaten/kota di provinsi tersebut.

3.2.2 Model Regresi

Berdasarkan uraian pada bab sebelumnya bahwa disparitas pendapatan dipengaruhi oleh beberapa faktor seperti pendidikan, kesehatan, infrastruktur jalan, tingkat partisipasi angkatan kerja dan investasi maka rancangan model yang dibangun adalah:

$$\text{Disparitas} = f(\text{pendidikan, kesehatan, jalan, tenaga kerja, investasi})$$

disparitas dicerminkan oleh indikator PDRB per kapita atas dasar harga konstan 2000, pendidikan dicerminkan oleh indikator ketersediaan prasarana pendidikan dasar dan rasio guru dan murid tingkat sekolah dasar, kesehatan dicerminkan oleh indikator ketersediaan prasarana kesehatan dasar yaitu puskesmas dan jumlah tenaga kesehatan/paramedis, jalan dicerminkan oleh indikator panjang jalan dalam kondisi baik, tenaga kerja dicerminkan oleh indikator Tingkat Partisipasi Angkatan Kerja (TPAK) dan investasi dicerminkan oleh indikator jumlah alokasi investasi di masing-masing kabupaten/kota.

Model penelitian ini menggunakan PDRB per kapita sebagai indikator dalam menganalisis disparitas pendapatan antar kabupaten/kota di Provinsi Jawa Barat berdasarkan apa yang disampaikan Tambunan (2001) bahwa PDRB per kapita dapat dijadikan sebagai indikator dalam menganalisis *development gap* selain konsumsi rumah tangga per kapita, *human development index*, kontribusi sektoral terhadap PDRB, tingkat kemiskinan dan struktur fiskal. Secara hipotesis dapat dikatakan bahwa semakin baik pembangunan di suatu daerah maka semakin tinggi PDRB per kapita daerah tersebut.

Model penelitian ini merupakan modifikasi dari model penelitian terdahulu yang dilakukan oleh Sadikin pada tahun 2005 yang menganalisis faktor-faktor yang mempengaruhi kesenjangan antar kecamatan di Kabupaten Kudus. Penelitian tersebut menggunakan variabel PDRB per kapita sebagai variabel dependent dan menggunakan variabel independent sebanyak sembilan variabel yaitu: (1) rasio luas lahan pertanian terhadap luas wilayah, (2) produktivitas tenaga kerja sektor primer, (3) produktivitas tenaga kerja sektor sekunder, (4) produktivitas tenaga kerja sektor tersier, (5) tingkat partisipasi angkatan kerja, (6) rasio jumlah fasilitas pendidikan per 1.000 jumlah penduduk, (7) rasio jumlah fasilitas kesehatan per 1.000 jumlah penduduk, (8) rasio jumlah jalan terhadap luas wilayah serta (9) rasio jumlah kendaraan bermotor (mobil) terhadap jumlah rumah tangga. Model awal tersebut dapat dituliskan sebagai berikut:

$$PDRB_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \beta_3 X_{3it} + \beta_4 X_{4it} + \beta_5 X_{5it} + \beta_6 X_{6it} + \beta_7 X_{7it} + \beta_8 X_{8it} + \beta_9 X_{9it}$$

dimana:

X_1 = rasio luas lahan pertanian terhadap luas wilayah

X_2 = produktivitas tenaga kerja sektor primer

X_3 = produktivitas tenaga kerja sektor sekunder

X_4 = produktivitas tenaga kerja sektor tersier

X_5 = tingkat partisipasi angkatan kerja

X_6 = rasio jumlah fasilitas pendidikan per 1.000 jumlah penduduk

X_7 = rasio jumlah fasilitas kesehatan per 1.000 jumlah penduduk

X_8 = rasio jumlah jalan terhadap luas wilayah

X_9 = rasio jumlah kendaraan bermotor (mobil) terhadap jumlah rumah tangga

Model awal tersebut kemudian dimodifikasi dengan mempertahankan variabel ketersediaan sarana prasarana pendidikan dan kesehatan, infrastruktur jalan dan TPAK serta menambahkan variabel baru yaitu variabel investasi ke dalam model sehingga model yang dimodifikasi menjadi berikut:

$$PDRB_{it} = \beta_0 + \beta_1 SAPDIKSD_{it} + \beta_2 RASIOGURSD_{it} + \beta_3 JALAN_{it} + \beta_4 SAPKES_{it} + \beta_5 DOKTER_{it} + \beta_6 TPAK_{it} + \beta_7 INVESTASI_{it} + e_{it}$$

Keterangan :

$PDRB_{it}$ = PDRB per kapita ADHK th.2000 kab/kota ke-i dan periode ke-t

$SAPDIKSD_{it}$ = sarana prasarana pendidikan dasar kab/kota ke-i dan periode ke-t

$RASIOGURSD_{it}$ = rasio guru SD per murid kab/kota ke-i dan periode ke-t

$JALAN_{it}$ = sarana prasarana transportasi kab/kota ke-i dan periode ke-t

$SAPKES_{it}$ = sarana prasarana kesehatan kab/kota ke-i dan periode t

$DOKTER_{it}$ = jumlah dokter kab/kota ke-i dan periode ke-t

- TPAK_{it}** = Tingkat Partisipasi Angkatan Kerja kab/kota ke-i dan periode ke-t
- INVESTASI_{it}** = jumlah investasi kab/kota ke-i dan periode ke-t
- e_{it}** = variabel gangguan pada kab/kota ke-i dan periode ke-t

Berdasarkan rancangan model ekonometrika di atas dapat dibuat definisi operasional variabel sebagai berikut:

- Disparitas Pendapatan dalam hal ini diwakili oleh PDRB per kapita Atas Dasar Harga Konstan tahun 2000 masing-masing kabupaten/kota;
- Sarana dan Prasarana Pendidikan Dasar dalam hal ini diwakili oleh jumlah Sekolah Dasar negeri dan swasta yang berada di masing-masing kabupaten/kota;
- Rasio Guru terhadap Murid Sekolah Dasar adalah perbandingan jumlah guru terhadap jumlah murid Sekolah Dasar negeri dan swasta masing-masing kabupaten/kota;
- Sarana dan Prasarana Transportasi dalam hal ini diwakili oleh panjang jalan berbagai fungsi dalam kondisi baik yang berada di wilayah masing-masing kabupaten/kota;
- Sarana dan Prasarana Kesehatan dalam hal ini diwakili oleh jumlah puskesmas, puskesmas pembantu dan puskesmas keliling yang berada di wilayah masing-masing kabupaten/kota;
- Jumlah dokter adalah jumlah dokter dan paramedis yang bertugas di puskesmas yang berada di masing-masing kabupaten/kota;
- Tingkat Partisipasi Angkatan Kerja (TPAK) adalah ukuran yang menggambarkan jumlah angkatan kerja untuk setiap 100 tenaga kerja masing-masing kabupaten/kota atau dengan kata lain TPAK adalah perbandingan angkatan kerja terhadap penduduk usia kerja;
- Investasi adalah jumlah uang yang diinvestasikan di masing-masing kabupaten/kota.

3.2.3 Regresi Data Panel

Dalam penelitian ini digunakan regresi data panel karena tujuan dari penelitian ini adalah untuk menganalisis faktor-faktor apa saja yang mempengaruhi disparitas pendapatan di Provinsi Jawa Barat antar wilayah (*cross section*) dan antar waktu (*time series*). Ada beberapa keuntungan yang diperoleh dengan menggunakan data panel. Pertama, data panel yang merupakan gabungan dua data *time series* dan *cross section* mampu menyediakan data yang lebih banyak sehingga akan menghasilkan *degree of freedom* yang lebih besar. Kedua, menggabungkan informasi dari data *time series* dan *cross section* dapat mengatasi masalah yang timbul ketika ada masalah penghilangan variabel (*omitted-variabel*).

Terdapat beberapa metode yang biasa digunakan untuk mengestimasi model regresi dengan data panel diantaranya yaitu:

1. Pendekatan *common effect*

Pendekatan ini hanya mengkombinasikan data *time series* dan *cross section* tanpa melihat perbedaan antar waktu dan individu. Dalam pendekatan ini tidak memperhatikan dimensi individu maupun waktu yang bisa menggunakan metode OLS (*Ordinary Least Square*) untuk mengestimasi model data panel.

2. Pendekatan *fixed effect*

Pendekatan ini mengestimasi data panel menggunakan variabel *dummy* untuk menangkap adanya perbedaan intersep. *fixed effect* ini didasarkan adanya perbedaan intersep antar individu namun intersepanya sama antar waktu (*time invariant*). Disamping itu, model ini juga mengasumsikan bahwa koefisien regresi (*slope*) tetap antar individu dan antar waktu.

Untuk menjelaskan perbedaan intersep digunakan metode teknik variabel *dummy*, teknik ini seringkali disebut teknik *Least Squares Dummy Variabel (LSDV)*.

3. Pendekatan *random effect*

Dimasukkannya variabel *dummy* di dalam model *fixed effect* bertujuan untuk mewakili ketidaktahuan tentang model yang sebenarnya. Namun, ini juga membawa konsekuensi berkurangnya derajat kebebasan (*degree of freedom*) yang pada akhirnya mengurangi efisiensi parameter. Masalah

ini bisa diatasi dengan menggunakan variabel gangguan (*error terms*) yang dikenal sebagai metode *random effect*. Di dalam model ini kita akan mengestimasi data panel dimana variabel gangguan mungkin akan saling berhubungan antar waktu dan antar individu.

Nama metode *random effect* berasal dari pengertian bahwa variabel gangguan v_{it} terdiri dari dua komponen yaitu variabel gangguan secara

menyeluruh μ_{it} yaitu kombinasi *time series* dan *cross section* dan

variabel gangguan secara individu μ_i . Dalam hal ini variabel gangguan μ_{it}

adalah berbeda-beda antar individu tetapi tetap antar waktu.

Dari ketiga teknik estimasi tersebut maka dipilih salah satu teknik estimasi yang paling tepat untuk mengestimasi regresi data panel. Pemilihan tersebut didasarkan atas uji-uji sebagai berikut:

1. Uji statistik F untuk memilih antara metode OLS tanpa variabel dummy atau *fixed effect* dengan metode OLS (*common effect*) atau sering disebut sebagai Uji Chow. Hipotesis nulnya adalah bahwa intersep adalah sama sedangkan nilai statistik hitung F adalah sebagai berikut:

$$F_{stat} = \frac{(RSS_1 - RSS_2)/m}{(RSS_2)/(n - k)}$$

dimana

RSS_1 = *Residual Sum of Squares* dari Model *Common Effect*

RSS_2 = *Residual Sum of Squares* dari Model *Fixed Effect*

m = Jumlah restriksi dari Model *Common Effect*

n = Jumlah observasi

k = Jumlah Parameter dalam Model *Fixed Effect*

Statisitk hitung tersebut berdistribusi F dengan derajat kebebasan (df) sebanyak m untuk numerator dan sebanyak $n-k$ untuk denominator.

Apabila $F_{stat} \geq F_{(\alpha)(m)(n-k)}$ maka H_0 ditolak yang berarti asumsi bahwa

koefisien intersep dan slope adalah sama tidak berlaku atau dapat dikatakan model yang cocok adalah model *fixed effect* sedangkan jika

$F_{stat} \leq F_{(alpha, k-1)}$ maka H_0 diterima yang berarti asumsi bahwa

koefisien intersep dan slope adalah sama berlaku atau dapat dikatakan model yang cocok adalah model *common effect*.

2. Uji statistik Hausman untuk memilih antara metode *fixed effect* dengan metode *random effect*. Terdapat dua hal yang mendasari pengujian statistik ini yaitu: (1) tentang ada tidaknya korelasi antara error terms e_{it}

dan variabel independen X . Jika diasumsikan terjadi korelasi antara error

terms e_{it} dan variabel independen X maka model *random effect* lebih

tepat. Sebaliknya jika tidak ada korelasi antara error terms e_{it} dan

variabel independen X maka model *fixed effect* lebih tepat; (2) berkaitan

dengan jumlah sampel di dalam penelitian. Jika sampel yang kita ambil adalah hanya bagian kecil dari populasi maka kita akan mendapatkan error terms e_{it} yang bersifat random sehingga model *random effect* lebih

tepat.

Uji secara formal ini dikembangkan oleh Hausman. Hausman telah mengembangkan suatu uji statistik untuk memilih apakah menggunakan model *fixed effect* atau *random effect*. Uji ini didasarkan pada ide bahwa LSDV di dalam metode *fixed effect* adalah efisien sedangkan metode GLS tidak efisien, di lain pihak alternatifnya metode GLS efisien dan LSDV tidak efisien. Karena itu hipotesis nulnya adalah hasil estimasi keduanya tidak berbeda sehingga uji Hausman bisa dilakukan berdasarkan perbedaan estimasi tersebut. Unsur penting untuk uji ini adalah kovarian matrik dari perbedaan vektor $[\beta - \beta_{GLS}]$. statistik uji

Hausman ini adalah

$$m = X'_{stat} = q' Var(q)^{-1} q$$

dimana

$$\hat{q} = [\hat{\beta} - \hat{\beta}_{GLS}] \text{ dan } \text{Var}(\hat{q}) = \text{Var}(\hat{\beta}) - \text{Var}(\hat{\beta}_{GLS})$$

Statistik uji Hausman ini mengikuti distribusi *chi squares* dengan *degree of freedom* sebanyak k dimana k adalah jumlah variabel independen. Jika nilai statistik Hausman lebih besar dari nilai kritisnya maka model yang tepat adalah model *fixed effect* sedangkan sebaliknya jika nilai statistik Hausman lebih kecil dari nilai kritisnya maka model yang tepat adalah model *random effect*.

Setelah melewati pengujian pemilihan model maka analisis dilanjutkan dengan melakukan pengujian-pengujian statistik diantaranya yaitu:

1. Pengujian koefisien regresi secara parsial (uji t)

Uji distribusi t secara garis besar adalah pengujian untuk menguji apakah koefisien regresi parsial berbeda secara signifikan dari nol atau apakah suatu variabel bebas secara individu berhubungan dengan variabel terikat.

Hipotesis

H_0 : $\beta_i = 0$, menyatakan koefisien regresi tidak berbeda nyata dari nol
(tidak signifikan)

H_1 : $\beta_i \neq 0$, menyatakan koefisien regresi berbeda nyata dari nol
(signifikan)

Pengujian dilakukan dengan membandingkan nilai t-hitung terhadap nilai t-tabel dengan derajat bebas $n-2$ pada tingkat kepercayaan α tertentu.

Nilai t-hitung dinyatakan dengan formula berikut:

$$t_{hitung} = \frac{\hat{\beta}_i - \beta_i}{se(\hat{\beta}_i)}$$

dimana:

$\hat{\beta}_i$ = nilai dugaan koefisien regresi

$se(\hat{\beta}_i)$ = standar error pendugaan koefisien regresi

Kriteria pengujian:

- $t_{hitung} \leq t_{(\alpha/2, n-2)}$ berarti terima H_0 sedangkan
- $t_{hitung} > t_{(\alpha/2, n-2)}$ berarti tolak H_0

Selain membandingkan t_{hitung} dengan t_{tabel} pengujian t juga dapat dilakukan dengan melihat probabilitas yang terdapat dalam output komputer, jika probabilitas pada output komputer dibawah α yang

ditentukan maka koefisien regresi signifikan.

2. Pengujian model secara keseluruhan (uji F)

Hipotesis

$$H_0 : \beta_i = 0, i = 0, 1, 2, \dots, k$$

$$H_1 : \text{minimal ada satu } \beta_i \neq 0$$

Alat uji yang digunakan adalah Uji F yang dinyatakan dengan formula berikut:

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

dimana:

$$SSR = \sum (\hat{Y}_i - \bar{Y})^2 = \text{jumlah kuadrat regresi}$$

$$SSE = \sum (Y_i - \hat{Y}_i)^2 = \text{jumlah kuadrat error peramalan}$$

$$k = \text{banyaknya variabel bebas}$$

$$n = \text{banyaknya observasi}$$

Kriteria pengujian:

- $F_{hitung} \leq F(\alpha, n_1, n_2)$ berarti terima H_0 sedangkan
- $F_{hitung} > F(\alpha, n_1, n_2)$ berarti tolak H_0

dimana:

$$n_1 = \text{derajat bebas pembilang}$$

n_2 = derajat bebas penyebut

Jika melihat pada output komputer apabila probabilitas dibawah α yang ditentukan maka model signifikan menjelaskan variabel terikat Y .

3. Pengujian koefisien determinasi (R^2)

Koefisien determinasi (R^2) mempunyai nilai antara 0 dan 1. Semakin besar nilai R^2 berarti semakin besar variasi variabel terikat yang dapat dijelaskan oleh variasi variabel bebas. Sebaliknya semakin kecil nilai R^2 berarti semakin kecil variasi variabel terikat yang dapat dijelaskan oleh variasi-variabel bebas.

3.3 Jenis dan Sumber Data

Jenis data yang digunakan adalah data sekunder yang diambil dari dokumen resmi yang dikeluarkan oleh Badan Pusat Statistik Provinsi Jawa Barat. Data tersebut diambil dari buku Produk Domestik Regional Bruto Kabupaten/Kota se Jawa Barat, buku Jawa Barat Dalam Angka, buku Penyusunan Basis Data Untuk Perhitungan IPM Provinsi Jawa Barat dan berbagai dokumen yang berkaitan dengan penelitian ini.

3.4 Teknik Pengumpulan dan Pengolahan Data

Data sekunder yang diperoleh dari data statistik, laporan, atau referensi yang tersedia, diolah dan dianalisis dengan menggunakan alat dan metode statistik yang sesuai. Dalam penelitian ini pengolahan data dilakukan dengan menggunakan bantuan dari software statistika yaitu *E-Views 4.0*. agar proses pengolahan data lebih cepat dan akurat.