

## BAB III

### METODE PENELITIAN

#### 3.1 Data dan Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder. Data sekunder untuk variabel dependen Ketimpangan Pendapatan serta variabel-variabel independen berupa Laju Pertumbuhan Produk Domestik Regional Bruto (yang dilihat dari Laju Pertumbuhan PDRB ADHK 2010) dan Inflasi Daerah diambil dari Badan Pusat Statistik (BPS). *Dummy* Dana Desa memakai kode 0 untuk tahun 2011 – 2014 (sebelum implementasi Dana Desa) dan kode 1 untuk tahun 2015 – 2018 (sesudah implementasi Dana Desa). Data panel pada penelitian ini menggabungkan data *time series* dari tahun 2011 – 2018 dan data *cross section* sebanyak 34 provinsi. Jumlah observasi secara keseluruhan adalah 272.

#### 3.2 Alat Analisis

Penelitian ini mengimplementasikan data panel adalah gabungan antara data runtut waktu (*time series*) dan data silang (*cross section*). Menurut Widarjono (2009) penggunaan data panel dalam sebuah observasi mempunyai beberapa kelebihan. Pertama, data panel yang merupakan gabungan dua data *time series* dan *cross section* mampu menyediakan data yang lebih banyak sehingga akan lebih 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-variable*). Hsiao (1986), mencatat bahwa penggunaan

panel data dalam penelitian ekonomi memiliki beberapa keuntungan utama dibandingkan data jenis *cross section* maupun *time series*. Pertama, dapat memberikan peneliti jumlah pengamatan yang besar, meningkatkan *degree of freedom* (derajat kebebasan), data memiliki variabilitas yang besar dan mengurangi kolinieritas antara variabel penjelas, dimana dapat menghasilkan estimasi ekonometri yang efisien. Kedua, panel data dapat memberikan informasi lebih banyak yang tidak dapat diberikan hanya oleh data *cross section* atau *time series* saja. Ketiga, panel data dapat memberikan penyelesaian yang lebih baik dalam inferensi perubahan dinamis dibandingkan data *cross section*.

Menurut Wibisono (2005) keunggulan regresi data panel antara lain: pertama, panel data mampu memperhitungkan heterogenitas individu secara ekspilisit dengan mengizinkan variabel spesifik individu. Kedua, kemampuan mengontrol heterogenitas ini selanjutnya menjadikan data panel dapat digunakan untuk menguji dan membangun model perilaku lebih kompleks. Ketiga, data panel mendasarkan diri pada observasi *cross section* yang berulang-ulang (*time series*), sehingga metode data panel cocok digunakan sebagai *study of dynamic adjustment*. Keempat, tingginya jumlah observasi memiliki implikasi pada data yang lebih informatif, lebih variatif, dan multikolinearitas antara data semakin berkurang, dan derajat kebebasan (*degree of freedom/df*) lebih tinggi sehingga dapat diperoleh hasil estimasi yang lebih efisien. Kelima, data panel dapat digunakan untuk mempelajari model-model perilaku yang kompleks. Terakhir, data panel dapat digunakan untuk meminimalkan bias yang mungkin ditimbulkan oleh agregasi data individu.

### 3.3 Tahapan Analisis Regresi Data Panel

#### 3.3.1 Estimasi Model Regresi Data Panel

Model persamaan data panel yang merupakan gabungan dari data *cross section* dan data *time series* adalah sebagai berikut:

$$Y_{it} = \alpha + \beta_1 X_{1it} + \beta_2 X_{2it} + \dots + \beta_n X_{nit} + e_{it} \dots \dots \dots (1.1)$$

keterangan:

$Y_{it}$  = variabel terikat (dependen)

$\alpha$  = konstanta

$\beta$  = koefisien variabel bebas

$X_{it}$  = variabel bebas (independen)

$i$  = entitas ke- $i$

$t$  = periode ke- $t$ .

Persamaan di atas merupakan model regresi linier berganda dari beberapa variabel bebas dan satu variabel terikat. Estimasi model regresi linier berganda bertujuan untuk memprediksi parameter model regresi yaitu nilai konstanta ( $\alpha$ ) dan koefisien regresi ( $\beta_i$ ). Konstanta biasa disebut dengan *intercept* dan koefisien regresi biasa disebut dengan *slope*. Regresi data panel memiliki tujuan yang sama dengan regresi linier berganda, yaitu memprediksi nilai *intercept* dan *slope*. Penggunaan data panel dalam regresi akan menghasilkan *intercept* dan *slope* yang berbeda pada setiap entitas/perusahaan dan setiap periode waktu. Model regresi data panel yang akan diestimasi membutuhkan asumsi terhadap *intercept*, *slope* dan variabel gangguannya.

Menurut Widarjono (2007) ada beberapa kemungkinan yang akan muncul atas adanya asumsi terhadap *intercept*, *slope* dan variabel gangguannya.

- a. Diasumsikan *intercept* dan *slope* adalah tetap sepanjang periode waktu dan seluruh entitas/perusahaan. Perbedaan *intercept* dan *slope* dijelaskan oleh variabel gangguan (residual).
- b. Diasumsikan *slope* adalah tetap tetapi *intercept* berbeda antar entitas/perusahaan.
- c. Diasumsikan *slope* tetap tetapi *intercept* berbeda baik antar waktu maupun antar individu.
- d. Diasumsikan *intercept* dan *slope* berbeda antar individu.
- e. Diasumsikan *intercept* dan *slope* berbeda antar waktu dan antar individu.

Dari berbagai kemungkinan yang disebutkan di atas, muncullah berbagai kemungkinan model/teknik yang dapat dilakukan oleh regresi data panel. Dalam banyak literatur hanya asumsi pertama sampai ketiga saja yang sering menjadi acuan dalam pembentukan model regresi data panel.

Menurut Widarjono (2007: 251), untuk mengestimasi parameter model dengan data panel, terdapat tiga teknik (model) yang sering ditawarkan, yaitu:

(1) Model *Common Effect*

Teknik ini merupakan teknik yang paling sederhana untuk mengestimasi parameter model data panel, yaitu dengan mengkombinasikan data *cross section* dan *time series* sebagai satu kesatuan tanpa melihat adanya perbedaan waktu dan entitas

(individu). Dimana pendekatan yang sering dipakai adalah metode *Ordinary Least Square* (OLS). Model *Common Effect* mengabaikan adanya perbedaan dimensi individu maupun waktu atau dengan kata lain perilaku data antar individu sama dalam berbagai kurun waktu. Persamaan regresinya dapat ditulis sebagai berikut:

$$Y_{it} = \alpha + \beta'X_{it} + \varepsilon_{it} ; i = 1, 2, \dots, n; t = 1, 2, \dots, t \dots \dots \dots (1.2)$$

keterangan:

n = jumlah unit/individu *cross section*

t = jumlah periode waktu.

### (2) Model Efek Tetap (*Fixed Effect*)

Pendekatan model *Fixed Effect* mengasumsikan bahwa *intercept* dari setiap individu adalah berbeda sedangkan *slope* antar individu adalah tetap (sama). Teknik ini menggunakan variabel *dummy* untuk menangkap adanya perbedaan *intercept* antar individu. Persamaan regresinya dapat ditulis sebagai berikut:

$$Y_{it} = \alpha_i + \beta'X_{it} + \varepsilon_{it} ; i = 1, 2, \dots, n; t = 1, 2, \dots, t \dots \dots \dots (1.3)$$

keterangan:

n = jumlah unit/individu *cross section*

t = jumlah periode waktu.

### (3) Model Efek Random (*Random Effect*)

Pendekatan yang dipakai dalam *Random Effect* mengasumsikan setiap perusahaan mempunyai perbedaan *intercept*, yang mana *intercept* tersebut adalah variabel *random* atau stokastik. Model ini sangat berguna jika individu (entitas) yang

diambil sebagai sampel adalah dipilih secara random dan merupakan wakil populasi. Teknik ini juga memperhitungkan bahwa *error* mungkin berkorelasi sepanjang *cross section* dan *time series*. Persamaan regresinya dapat ditulis sebagai berikut:

$$Y_{it} = \alpha + \beta' X_{it} + u_i + \varepsilon_{it}; i = 1, 2, \dots, n; t = 1, 2, \dots, t \dots \dots \dots (1.4)$$

keterangan:

n = jumlah unit/individu *cross section*

t = jumlah periode waktu.

### 3.3.2 Pemilihan Model (Teknik Estimasi) Regresi Data Panel

Pada dasarnya ketiga teknik (model) estimasi data panel dapat dipilih sesuai dengan keadaan penelitian, dilihat dari jumlah individu bank dan variabel penelitiannya. Namun demikian, ada beberapa cara yang dapat digunakan untuk menentukan teknik mana yang paling tepat dalam mengestimasi parameter data panel. Menurut Widarjono (2007: 258), ada tiga uji untuk memilih teknik estimasi data panel. Pertama, uji statistik F digunakan untuk memilih antara metode *Common Effect* atau metode *Fixed Effect*. Kedua, uji Hausman yang digunakan untuk memilih antara metode *Fixed Effect* atau metode *Random Effect*. Ketiga, uji Lagrange Multiplier (LM) digunakan untuk memilih antara metode *Common Effect* atau metode *Random Effect*.

#### (1) Uji Statistik F (Uji Chow)

Untuk mengetahui model mana yang lebih baik dalam pengujian data panel, bisa dilakukan dengan penambahan variabel *dummy* sehingga dapat diketahui bahwa

*intercept*-nya berbeda. Hal ini dapat diuji dengan uji statistik F. Uji ini digunakan untuk mengetahui apakah teknik regresi data panel dengan metode *Fixed Effect* lebih baik dari regresi model data panel tanpa variabel *dummy* atau metode *Common Effect*.

Ho: *Common Effect*

Ha: *Fixed Effect*.

(2) Uji Hausman

Hausman telah mengembangkan suatu uji untuk memilih apakah metode *Fixed Effect* dan metode *Random Effect* lebih baik dari metode *Common Effect*. Uji Hausman ini didasarkan pada ide bahwa *Least Squares Dummy Variables* (LSDV) dalam metode *Fixed Effect* dan *Generalized Least Squares* (GLS) dalam metode *Random Effect* adalah efisien sedangkan *Ordinary Least Squares* (OLS) dalam metode *Common Effect* tidak efisien. Dilain pihak, alternatifnya adalah metode OLS efisien dan GLS tidak efisien. Karena itu, uji hipotesis nulnya adalah hasil estimasi keduanya tidak berbeda sehingga uji Hausman bisa dilakukan berdasarkan perbedaan estimasi tersebut.

Ho: *Random Effect*

Ha: *Fixed Effect*.

Apabila nilai statistik Hausman lebih besar dari nilai kritis Chi-Squares maka Ho ditolak yang artinya model yang tepat untuk regresi data panel adalah model *Fixed Effect*. Dan sebaliknya, apabila nilai statistik Hausman lebih kecil dari nilai kritis Chi-Squares maka Ho tidak ditolak yang artinya model yang tepat untuk regresi data panel adalah model *Random Effect*.

### (3) Uji Lagrange Multiplier

Menurut Widarjono (2007: 260), untuk mengetahui apakah model *Random Effect* lebih baik dari model *Common Effect* digunakan Lagrange Multiplier (LM). Uji Signifikansi *Random Effect* ini dikembangkan oleh Breusch-Pagan. Pengujian didasarkan pada nilai residual dari metode *Common Effect*.

Ho: *Common Effect*

Ha: *Random Effect*.

Apabila nilai LM hitung lebih besar dari nilai kritis Chi-Squares maka Ho ditolak yang artinya model yang tepat untuk regresi data panel adalah model *Random Effect*. Sebaliknya, apabila nilai LM hitung lebih kecil dari nilai kritis Chi-Squares maka Ho tidak ditolak yang artinya model yang tepat untuk regresi data panel adalah model *Common Effect*.

### 3.3.3 Pengujian Asumsi Klasik (Multikolinearitas dan Heteroskedastisitas)

Regresi data panel memberikan alternatif model, *Common Effect*, *Fixed Effect* dan *Random Effect*. Model *Common Effect* dan *Fixed Effect* menggunakan pendekatan *Ordinary Least Squared* (OLS) dalam teknik estimasinya, sedangkan *Random Effect* menggunakan *Generalized Least Squares* (GLS) sebagai teknik estimasinya. Uji asumsi klasik yang digunakan dalam regresi linier dengan pendekatan *Ordinary Least Squared* (OLS) meliputi uji Linearitas, Autokorelasi, Heteroskedastisitas, Multikolinearitas dan Normalitas. Walaupun demikian, tidak semua uji asumsi klasik harus dilakukan pada setiap model regresi linier dengan pendekatan OLS.



Uji linearitas hampir tidak dilakukan pada setiap model regresi linier karena sudah diasumsikan bahwa model bersifat linier. Jika harus dilakukan semata-mata untuk melihat sejauh mana tingkat linieritasnya. Autokorelasi hanya terjadi pada data *time series*. Pengujian autokorelasi pada data yang tidak bersifat *time series* (*cross section* atau panel) akan sia-sia semata atau tidaklah berarti. Multikolinearitas perlu dilakukan pada saat regresi linier menggunakan lebih dari satu variabel bebas. Jika variabel bebas hanya satu, maka tidak mungkin terjadi multikolinearitas. Heteroskedastisitas biasanya terjadi pada data *cross section*, dimana data panel lebih dekat ke ciri data *cross section* dibandingkan *time series*. Uji normalitas pada dasarnya tidak merupakan syarat BLUE (*Best Linier Unbias Estimator*) dan beberapa pendapat tidak mengharuskan syarat ini sebagai sesuatu yang wajib dipenuhi.

Dari penjelasan di atas dapat ditarik kesimpulan bahwa pada regresi data panel, tidak semua uji asumsi klasik yang ada pada metode OLS dipakai, hanya multikolinearitas dan heteroskedastisitas saja yang diperlukan.

#### (1) Uji Multikolinearitas

Uji multikolinearitas digunakan untuk menguji apakah di dalam model regresi ditemukan adanya korelasi antar variabel bebas atau variabel independen (Ghozali, 2013: 105). Menurut Chatterjee dan Price dalam Nachrowi (2002), adanya korelasi antara variabel independen menjadikan interpretasi koefisien-koefisien regresi menjadi tidak benar lagi. Meskipun demikian, bukan berarti korelasi yang terjadi antara variabel independen tidak diperbolehkan. Hanya kolinieritas yang sempurna (*perfect collinierity*) saja yang tidak diperbolehkan atau terjadinya korelasi linier antara

sesama variabel independennya. Sifat kolinier yang hampir sempurna (hubungannya tidak bersifat linier atau korelasi mendekati nol) masih diperbolehkan alias tidak termasuk dalam pelanggaran asumsi.

Sebagai aturan main yang kasar (*rule of thumb*), jika koefisien korelasi cukup tinggi di atas 0,85 maka diduga terdapat multikolinearitas di dalam model. Sebaliknya, jika koefisien korelasi relatif rendah, maka diduga model tidak mengandung unsur multikolinearitas (Farrar dan Glauber, 1967) dalam Widarjono (2013). Menurut Gujarati (2003: 68) apabila koefisien korelasi antar variabel bebas lebih dari atau sama dengan 0,8 maka dalam data teridentifikasi multikolinearitas yang serius.

## (2) Uji Heteroskedastisitas

Uji heteroskedastisitas bertujuan untuk menguji apakah dalam model regresi terjadi ketidaksamaan *variance* dari satu pengamatan ke pengamatan yang lain. Jika *variance* dari residual satu pengamatan ke pengamatan lain tetap, maka disebut homoskedastisitas dan jika berbeda disebut heteroskedastisitas. Model regresi yang baik adalah homoskedastisitas atau bisa dikatakan tidak terjadi heteroskedastisitas (Ghozali, 2013: 111). Heteroskedastisitas dapat diuji dengan uji Glejser. Uji Glesjer menyarankan untuk melakukan regresi nilai absolut residual dengan variabel independennya (Widarjono, 2013: 118). Menurut Gujarati (2006: 93) kriteria yang dipakai dalam pengambilan kesimpulan uji Glesjer, yaitu:

- a. Jika nilai t-statistik tidak signifikan pada derajat 5%, maka tidak terdapat masalah heteroskedastisitas.
- b. Jika nilai t-statistik signifikan pada derajat 5%, maka terdapat masalah heteroskedastisitas.

### 3.3.4 Koefisien Determinasi ( $R^2$ )

Koefisien determinasi ( $R^2$ ) mengukur tingkat kemampuan model dalam menjelaskan variasi perubahan variabel dependen. Koefisien determinasi juga dapat didefinisikan sebagai proporsi atau persentase dari total variasi variabel tak bebas Y yang dijelaskan oleh sebuah variabel penjelas X (Gujarati, 2006: 187). Nilai koefisien determinasi terletak di antara 0 – 1. Semakin besar nilai koefisien determinasi maka semakin baik kemampuan variabel independen dalam menerangkan variabel dependen. Jika nilai koefisien determinasi semakin kecil, maka kemampuan variabel independen dalam menjelaskan variabel dependen dikatakan terbatas. Pada regresi dengan variabel independen lebih dari satu, maka nilai koefisien determinasi diambil dari nilai *Adjusted R<sup>2</sup>*.

### 3.3.5 Uji Hipotesis

Menurut Nachrowi (2006) uji hipotesis digunakan untuk menguji signifikansi koefisien regresi yang diperoleh. Penelitian ini menggunakan pengujian hipotesis secara serentak (uji F) dan secara individu (uji t). Berikut penjelasan dari masing-masing pengujian.

(1) Uji F

Uji F merupakan pengujian hubungan regresi secara simultan atau serentak yang dilakukan untuk mengetahui apakah secara bersama-sama variabel independen memiliki pengaruh yang signifikan terhadap variabel dependen. Kriteria pengambilan keputusan dalam uji F yakni sebagai berikut:

- a. Jika nilai F-statistik  $<$  F-kritis atau nilai probabilitas F-statistik  $>$   $\alpha$ , maka tidak signifikan. Artinya secara serentak variabel-variabel independen tidak berpengaruh terhadap variabel dependen.
- b. Jika nilai F-statistik  $>$  F-kritis atau nilai probabilitas F-statistik  $<$   $\alpha$ , maka signifikan. Artinya secara serentak variabel-variabel independen berpengaruh terhadap variabel dependen.

(2) Uji t

Uji t digunakan untuk menguji hubungan regresi secara parsial atau individu yang menunjukkan pengaruh masing-masing variabel independen terhadap variabel dependen. Kriteria yang dipakai untuk menentukan keputusan adalah:

- a. Jika nilai t-statistik  $<$  t-kritis atau nilai probabilitas t-statistik  $>$   $\alpha$ , maka tidak signifikan. Artinya secara individu variabel independen tidak berpengaruh terhadap variabel dependen.
- b. Jika nilai t-statistik  $>$  t-kritis atau nilai probabilitas t-statistik  $<$   $\alpha$ , maka signifikan. Artinya secara individu variabel independen berpengaruh terhadap variabel dependen.

### 3.4 Model

Data panel merupakan gabungan dari data *cross section* dan data *time series*, dengan model yang dapat dituliskan sebagai berikut.

$$\hat{Y}_{it} = \alpha + \beta X_{it} + \varepsilon_{it}; i = 1, 2, \dots, n; t = 1, 2, \dots, t \dots \dots \dots (1.5)$$

keterangan:

n = jumlah unit/individu *cross section*

t = jumlah periode waktu

n . t = banyaknya data panel.

Berdasarkan teori di atas, model estimasi yang diterapkan dalam penelitian ini adalah sebagai berikut.

$$\widehat{GINI} = \hat{\beta}_0 + \hat{\beta}_1 PDRB_{it} + \hat{\beta}_2 INF_{it} + \hat{\beta}_3 DDANA_{it} \dots \dots \dots (1.6)$$

di mana:

GINI = Indeks Gini (Ketimpangan Distribusi Pendapatan)

$\beta_0$  = konstanta/*intercept*

$\beta_1$  = koefisien variabel PDRB

PDRB = Laju PDRB ADHK 2010 (dalam persen)

$\beta_2$  = koefisien variabel INF

INF = Inflasi Daerah (dalam persen)

$\beta_3$  = koefisien variabel DDAN

DDANA = *Dummy* Dana Desa (sebelum implementasi: 0, sesudah implementasi: 1)

$\varepsilon$  = *error term*.

### **3.5 Batasan Operasional**

#### **3.5.1 Indeks Gini (Ketimpangan Distribusi Pendapatan)**

Indeks Gini merupakan indikator yang menunjukkan tingkat ketimpangan pendapatan secara menyeluruh. Penelitian ini menggunakan nilai Koefisien Gini yang berkisar antara 0 hingga 1. Jika 0 berarti menunjukkan pemerataan pendapatan sempurna. Sedangkan angka 1 mengisyaratkan distribusi pendapatan sangat timpang. Koefisien Gini biasanya diperlihatkan oleh kurva yang disebut Kurva Lorenz. Kurva ini memperlihatkan hubungan kuantitatif antara persentase jumlah penduduk dan persentase pendapatan yang diperoleh selama kurun waktu tertentu, biasanya setahun.

#### **3.5.2 Laju Pertumbuhan Produk Domestik Regional Bruto**

PDRB dihitung menggunakan 2 cara, yaitu berdasarkan harga berlaku dan berdasarkan harga konstan yang sudah memperhitungkan tingkat kenaikan harga/inflasi. Data PDRB berdasarkan harga konstan lebih menggambarkan perkembangan produksi riil barang dan jasa yang dihasilkan oleh kegiatan ekonomi daerah tersebut. Pada penelitian ini, penulis akan menggunakan Laju Pertumbuhan Produk Domestik Regional Bruto Atas Dasar Harga Konstan 2010 dalam satuan persen.

#### **3.5.3 Inflasi Daerah**

Peneliti akan menggunakan Indeks Harga Konsumen (IHK) sebagai indikator yang paling sering digunakan untuk mengukur tingkat inflasi. Perubahan IHK dari

waktu ke waktu menunjukkan pergerakan harga dari paket barang dan jasa yang dikonsumsi masyarakat.

#### **3.5.4 Dummy Dana Desa**

Peraturan mengenai Dana Desa termuat dalam UU Nomor 6 Tahun 2014. Kebijakan ini sekaligus mengintegrasikan dan mengoptimalkan seluruh skema pengalokasian anggaran dari pemerintah kepada desa yang selama ini sudah ada. Peneliti menggunakan kode 0 untuk mewakili *dummy* sebelum implementasi dana desa (tahun 2011 – 2014) dan kode 1 sesudah implementasi dana desa yakni tahun 2015 – 2018.

