

## **BAB III**

### **METODE PENELITIAN**

#### **A. Objek dan Subjek Penelitian**

##### **1. Objek Penelitian**

Objek penelitian dalam penelitian ini adalah seluruh provinsi di Indonesia, kecuali 3 provinsi yaitu provinsi Bangka Belitung, provinsi Kalimantan Utara, dan provinsi Maluku.

##### **2. Subjek Penelitian**

Variabel dependen yang digunakan pada penelitian ini adalah Indeks Gini, sedangkan variabel independennya adalah Pertumbuhan Ekonomi, Tingkat Pengangguran Terbuka, Indeks Pembangunan Manusia (IPM), Produk Domestik Regional Bruto (PDRB) per kapita, Penanaman Modal Asing (PMA), Penanaman Modal Dalam Negeri (PMDN).

#### **B. Jenis Data**

Jenis data yang digunakan dalam penelitian ini adalah data sekunder yang berupa data panel, yakni gabungan dari data *time series* dari tahun 2010-2016 (7 tahun) dan data *cross section* dari 31 provinsi di Indonesia. Data yang digunakan dalam penelitian ini bersifat data kuantitatif, yaitu data yang diperoleh dari Badan Pusat Statistik dalam bentuk angka dan membutuhkan analisis lebih lanjut. Data yang digunakan dalam penelitian ini antara lain:

1. Indeks Gini dari 31 provinsi di Indonesia periode 2010-2016

## 2. Pertumbuhan Ekonomi 31 provinsi di Indonesia periode 2010-2016

3. Tingkat Pengangguran Terbuka 31 provinsi di Indonesia periode 2010-2016
4. Indeks Pemabangunan Manusia (IPM) 31 provinsi di Indonesia periode 2010-2016
5. PDRB per kapita atas harga konstan 2010 dari 31 provinsi di Indonesia periode 2010-2016
6. Penanaman Modal Asing (PMA) 31 provinsi di Indonesia periode 2010-2016
7. Penanaman Modal Dalam Negeri (PMDN) 31 provinsi di Indonesia periode 2010-2016

### C. Sumber Data

Sumber data merupakan segala sesuatu yang dapat memberikan informasi mengenai hal yang ingin diteliti. Berdasarkan sumbernya data dibedakan menjadi dua yaitu data primer dan data sekunder. Dalam penelitian ini peneliti menggunakan data sekunder. Sumber data yang digunakan dalam penelitian ini adalah literatur, artike, jurnal, publikasi Badan Pusat Statistik (BPS), Badan Koordinasi Penanaman Modal (BKPM), serta situs internet yang berkaitan dengan penelitian ini (Sugiyono, 2009). Adapun sumber data yang digunakan dalam penelitian ini sebagai berikut:

Data	Sumber
Ketimpangan distribusi Pendapatan (Indeks Gini)	Badan Pusat Statistik (BPS)
Pertumbuhan Ekonomi	Badan Pusat Statistik (BPS), Indonesia Dalam angka 2010-2016
Tingkat Pengangguran Terbuka	Badan Pusat Statistik (BPS), Indonesia Dalam angka 2010-2016
Indeks Pembangunan Manusia (IPM)	Badan Pusat Statistik (BPS), Indonesia Dalam angka 2010-2016

Pendapatan Domestik Regional Bruto (PDRB) per kapita	Badan Pusat Statistik (BPS), Indonesia Dalam angka 2010-2016
Penanaman Modal Asing (PMA)	Badan Pusat Statistik (BPS), Badan Koordinasi Penanaman Modal (BKPM)
Penanaman Modal Dalam Negeri	Badan Pusat Statistik (BPS), Badan Koordinasi Penanaman Modal (BKPM)

#### D. Definisi Operasional Variabel

Variabel penelitian adalah suatu atribut atau sifat atau nilai dari orang, objek, atau kegiatan yang mempunyai variasi tertentu yang ditentukan peneliti untuk dipelajari dan kemudian ditarik kesimpulan (Sugiyono, 2013). Dalam penelitian ini menggunakan dua macam variabel, yaitu:

##### 1. Variabel Dependen

Variabel dependen dalam penelitian ini adalah ketimpangan distribusi pendapatan. ketimpangan distribusi pendapatan menggambarkan distribusi pendapatan penduduk di suatu wilayah tertentu pada periode waktu tertentu yang dapat diukur dengan menggunakan Indeks Gini. Indeks Gini dapat digunakan untuk mengukur besaran ketimpangan distribusi pendapatan di masyarakat.

##### 2. Variabel Independen

Dalam penelitian ini variabel independen merupakan beberapa indikasi faktor yang mempengaruhi ketimpangan distribusi pendapatan di Indonesia. beberapa faktor tersebut antara lain:

a. Pertumbuhan ekonomi

Pertumbuhan ekonomi menggambarkan perubahan tingkat kegiatan ekonomi yang berlaku dari tahun ke tahun. Sehingga untuk mengetahuinya harus diadakan perbandingan pendapatan nasional dari tahun ke tahun atau yang lebih dikenal dengan laju pertumbuhan ekonomi (Sukirno, 1996). Data pertumbuhan ekonomi Indonesia dalam penelitian ini dinyatakan dalam persentase dan dari tahun 2010 hingga tahun 2016.

b. Tingkat Pengangguran Terbuka

Tingkat Pengangguran Terbuka (TPT) adalah persentase jumlah pengangguran terhadap jumlah angkatan kerja. Nilai TPT menunjukkan persentase angkatan kerja yang termasuk dalam pengangguran. Nilai TPT didapat dari jumlah pengangguran dibagi jumlah angkatan kerja dikalikan dengan seratus persen. Dalam penelitian ini menggunakan data TPT dari tahun 2010 hingga tahun 2016.

c. Indeks Pembangunan Manusia (IPM)

Indeks Pembangunan Manusia merupakan pengukuran perbandingan dari harapan hidup, melek huruf, pendidikan, dan standar hidup untuk semua negara di seluruh dunia. IPM digunakan untuk mengklasifikasikan apakah negara tersebut merupakan negara maju atau negara berkembang (UNDP, 1995). Data IPM yang digunakan dalam penelitian ini dinyatakan dalam persentase dan terdiri dari tahun 2010 hingga tahun 2016.

d. Pendapatan Domestik Regional Bruto (PDRB) per kapita

Produk Domestik Regional Bruto (PDRB) per kapita menggambarkan tingkat kemakmuran ekonomi suatu daerah. PDRB per kapita diperoleh dari PDRB dibagi dengan jumlah penduduk dalam suatu wilayah dan periode tertentu. PDRB per kapita dalam penelitian ini merupakan nilai PDRB per kapita berdasarkan harga konstan tahun 2010. Nilai PDRB per kapita ini dinyatakan dalam Ribu Rupiah dan terdiri dari data tahun 2010 hingga tahun 2016

e. Penanaman Modal Dalam Negeri (PMDN)

Menurut Undang-Undang No 25 tahun 1997 menyatakan bahwa penanaman modal adalah kegiatan menanam modal untuk melakukan usaha di wilayah negara Indonesia yang dilakukan oleh warga negara Indonesia, badan usaha Indonesia, negara republik Indonesia atau daerah daerah yang melakukan penanaman modal di wilayah negara Indonesia. Data PMDN dalam penelitian ini dinyatakan dalam Miliar Rupiah dan terdiri dari data tahun 2010 hingga 2016.

f. Penanaman Modal Asing (PMA)

Penanaman modal asing dapat diartikan sebagai penanaman modal yang dilakukan oleh pihak swasta di negara asal pemilik modal, atau penanaman modal suatu negara ke negara lain atas nama negara pemilik modal (Jhingan, 1994). Data PMA dalam penelitian ini

dinyatakan dalam Juta Dollar dan terdiri dari tahun 2010 hingga tahun 2016.

### 3. Alat Ukur data

Dalam mengolah data sekunder yang telah terkumpul, penulis menggunakan beberapa alat statistik, yakni: program *Microsoft Excel* 2010 dan program *Eviews 9.0*. *Microsoft Excel 2010* digunakan untuk pengolahan data berupa pembuatan data, sementara *E-views 9.0* digunakan untuk pengolahan regresi data.

### **E. Teknik Pengumpulan Data**

Teknik pengumpulan data yang digunakan dalam penelitian ini adalah teknik studi dokumenter. Teknik dokumentasi digunakan untuk memperoleh data dengan cara menghimpun dan menganalisis dokumen-dokumen, baik dokumen tertulis, gambar, maupun elektronik (Syaodih, 2009)

### **F. Teknik Analisis Data**

Pada penelitian ini menggunakan teknik analisis regresi data panel. 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 keuntungan. 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*). Persamaan dasar regresi data panel adalah sebagai berikut:

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

Dimana:

- $Y_{it}$  = variabel terikat (dependen)  
 $X_{it}$  = variabel terikat (Independen)  
 $i$  = *cross section*  
 $t$  = *time series*

Adapun tiga model pendekatan dalam regresi data panel meliputi:

1. *Common Effect (Pooled Least Square)*

Metode *Pooled Least Square* (PLS) merupakan metode yang digunakan untuk mengestimasi data panel dengan menggabungkan seluruh observasi pada masing-masing variabel. Sehingga intersep dari semua objek *cross-section* sama, dengan kata lain metode ini mengasumsikan tidak ada perbedaan setiap individu dalam kurun waktu (*time-invariant*) (Gujarati, 2012)

2. Model Efek Tetap (*Fixed Effect model*)

Menurut Gujarati (2012) *Fixed Effect Model* mengasumsikan bahwa terdapat perbedaan intersep antar individu. Akan tetapi koefisien (*slope*) dari variabel independen tetap sama antar individu atau antar waktu. Pada pendekatan *Fixed Effect Model* terdapat intersep yang berbeda-beda pada setiap individu *cross-section* sehingga menunjukkan perbedaan setiap individu, setiap intersep tidak berubah seiring dengan berjalannya waktu, dan hal tersebut

dinamakan *time-variant*. Sedangkan, koefisien (*slope*) dari masing-masing variabel independen sama untuk setiap individu atau antar waktu. Persamaan matematis model panel data yang menggunakan model *Fixed Effect* adalah sebagai berikut:

$$Y_i = \beta_{1it} + \beta_2 X_{2it} + \beta_3 X_{3it} + \mu_{it}$$

$$i = 1, 2 \dots N, t = 1, 2 \dots T$$

Dimana:

Y = Variabel Dependen

$\beta$  = Koefisien Regresi

i = *cross section*

t = waktu/*time series*

it = Data Penelitian

$\mu_{it}$  = nilai residual (faktor pengganggu) yang berada di luar model

menurut Gujarati (2012) *Fixed Effect Model* dapat dibedakan menjadi dua yaitu :

a. Model *fixed effect* satu arah

Setiap individu mempunyai intersep yang berbeda-beda, tetapi tidak berubah dengan seiring berjalannya waktu.

b. Model *fixed effect* dua arah

Dengan menambahkan efek waktu (*time effect*) pada setiap intersep dan efek waktu tersebut dapat dihitung dengan menggunakan dummy waktu.

3. Model Efek Random (*Random Effect model*)

Dalam pendekatan ini mengasumsikan terdapat perbedaan intersep, yang mana perbedaan intersep tersebut adalah variabel random atau stokastik. Model ini sangat berguna jika setiap individu (entitas) yang diambil sebagai

sampel adalah dipilih secara random dan merupakan wakil populasi. Pendekatan ini juga memperhitungkan bahwa *error* mungkin berkorelasi sepanjang *cross section* dan *time series* (Widarjono, 2007). Persamaan model random effect sebagai berikut:

$$Y_{it} = \alpha + \beta X_{it} + \varepsilon_{it}; \varepsilon_{it} = u_i + V_t + W_{it}$$

Dimana:

$u_i$	= error <i>cross section</i>
$v_t$	= error <i>time series</i>
$w_{it}$	= error gabungan

Dalam Penelitian ini, menggunakan model persamaan dalam bentuk semi log. Bentuk semi log dipilih karena adanya transformasi dua variabel dalam bentuk logaritma natural yaitu PDRB per kapita, Penanaman Modal Dalam Negeri (PMDN), dan Penanaman Modal Asing (PMA). Hal tersebut bertujuan untuk memperhalus data dimana data PDRB per kapita, PMDN, PMA dalam bentuk ribuan sementara Indeks Gini, TPT, PE, dan IPM dalam bentuk satuan dan puluhan atau bisa dikatakan memiliki perbedaan angka yang jauh. Berikut adalah model yang dipakai dalam penelitian ini:

$$IG_{it} = \beta_{1it} + \beta_2 PE_{it} + \beta_3 TPT_{it} + \beta_4 IPM_{it} + \beta_5 LOGPDRBPK_{it} + \beta_6 LOGPMDN_{it} + \beta_7 LOGPMA_{it} + e_{it}$$

Dimana :

IG	= Ketimpangan distribusi pendapatan (diukur dengan Indeks Gini) 31 provinsi di Indonesia periode 2010-2016
PE	= Pertumbuhan Ekonomi 31 Provinsi di Indonesia Periode 2010-2016
TPT	= Tingkat Pengangguran Terbuka 31 Provinsi di Indonesia Periode

	2010-2016
IPM	= Indeks Pembangunan Manusia 31 Provinsi di Indonesia Periode 2010-2016
PDRBPK	= Pendapatan Domestik Regional Bruto per kapita 31 Provinsi di Indonesia Periode 2010-2016
PMDN	= Penanaman Modal Dalam Negeri 31 Provinsi di Indonesia Periode 2010-2016
$i$	= <i>cross section</i>
$t$	= <i>time series</i>
$\beta_1$	= <i>Intercept</i>
$\beta_2 - \beta_7$	= koefisien regresi
$e_{it}$	= <i>error</i>

## 1. Uji Pemilihan Model

Untuk menentukan model terbaik yang digunakan, peneliti harus melakukan uji pemilihan teknik estimasi regresi. Terdapat tiga cara dalam melakukan pemilihan teknik estimasi untuk menentukan teknik yang paling tepat digunakan dalam mengestimasi parameter data panel, yakni pertama, uji Chow (*Likelihood Ratio*) digunakan untuk memilih antara metode *Common Effect* atau *Fixed Effect*. Kedua, uji Hausman yang digunakan untuk memilih antara model *Fixed Effect* atau *Random Effect*. Ketiga, uji Lagrange Multiplier yang digunakan untuk memilih antara model *Random Effect* atau *Common Effect* (Gujarati,2013)

### a. Uji *Likelihood Ratio* (Uji Chow)

Uji *Likelihood Ratio* digunakan untuk memilih model terbaik antara *common effect model* atau *fixed effect model*. Hipotesis nol menyatakan bahwa model yang tepat untuk regresi panel data adalah *common effect model* dan

hipotesis alternatifnya menyatakan bahwa model yang tepat untuk regresi panel data adalah model *fixed effect model* (Gujarati, 2013)

$H_0 = \text{common effect model}$

$H_1 = \text{fixed effect model}$

Apabila dalam hasil uji Chow nilai F-statistik  $> F$  tabel atau probabilitas  $<$  taraf signifikan, maka  $H_0$  ditolak dan *Fixed Effect Model* yang baik digunakan. Tetapi, jika hasil uji Chow menunjukkan nilai F-statistik  $< F$  tabel atau probabilitasnya  $>$  taraf signifikansi, maka  $H_0$  diterima dan *Common Effect Model* yang digunakan (Gujarati, 2013)

**b. Uji Hausman**

Uji *hausman* digunakan untuk memilih model terbaik antara *fixed effect model* atau *random effect model*. Hipotesis nol menyatakan bahwa model yang tepat untuk regresi data panel adalah model *random effect model* dan hipotesis alternatifnya menyatakan bahwa model yang tepat untuk regresi data panel adalah *fixed effect model* (Widarjono, 2013)

$H_0 = \text{random effect model}$

$H_1 = \text{fixed effect model}$

Apabila nilai statistik Hausman  $> \text{Chi-Square}$  dan nilai probabilitas  $< \alpha$  (nilai kritis) maka  $H_0$  ditolak dan pendekatan Fixed Effect Model yang dipilih. Sedangkan, jika nilai statistik Hausman  $< \text{Chi-square}$  dan nilai probabilitas  $> \alpha$  (nilai kritis) maka  $H_0$  diterima dan pendekatan Random Effect Model yang dipilih (Widarjono, 2013)

**c. Uji LM (*Lagrange Multiplier*)**

Uji LM (*Lagrange Multiplier*) digunakan untuk menentukan model terbaik antara *Common Effect model* atau *Random Effect model*. Uji *Lagrange Multiplier* ini dikembangkan oleh Breusch Pagan. Metode Breusch Pagan digunakan untuk menguji signifikansi *Random Effect* yang didasarkan pada nilai residual dari metode OLS (Widarjono, 2009). Hipotesis yang digunakan sebagai berikut:

$H_0 = \text{common effect model}$

$H_1 = \text{random effect model}$

Uji LM ini didasari oleh distribusi *Chi-squares* dengan *degree of freedom* sebesar jumlah variabel independen. Jika nilai LM statistik lebih besar dari nilai kritis statistik *chi-squares* maka kita menolak hipotesis nul, yang artinya estimasi yang tepat untuk model regresi data panel adalah metode *Random Effect* dari pada metode *Common Effect*. Sebaliknya jika nilai LM statistik lebih kecil dari nilai statistik *chi-squares* sebagai nilai kritis, maka kita menerima hipotesis nul, yang artinya estimasi yang digunakan dalam regresi data panel adalah metode *Common Effect* bukan metode *Random Effect* (Widarjono, 2009).

**2. Uji Asumsi Klasik**

**a. Uji Multikolinearitas**

Uji Multikolinearitas bertujuan untuk menguji apakah model regresi ditemukan adanya hubungan antar variabel bebas atau independen. Model

regresi yang baik seharusnya tidak terjadi korelasi antar variabel independen. Jika terjadi multikolinearitas dalam model, estimator masih bersifat *Best Linear Unbiased Estimator* (BLUE) namun estimator mempunyai varian dan kovarian yang besar sehingga sulit didapatkan estimasi yang tepat (Widarjono, 2013)

Menurut Gujarati (2012), masalah multikolinearitas dapat di deteksi dari beberapa gejala sebagai berikut:

1. Bila nilai  $R^2$  yang dihasilkan terlalu tinggi, tetapi secara individual variabel-variabel independen banyak yang tidak signifikan terhadap variabel dependen
2. Melakukan regresi parsial dengan cara
  - A) Melakukan estimasi model awal dalam persamaan sehingga didapat nilai  $R^2$
  - B) Melakukakan *auxiliary regression* pada masing masing variabel penjelasan yang akan diteliti
  - C) Membandingkan nilai  $R^2$  dalam model persamaan awal dengan  $R^2$  pada model persamaan regresi parsial, apabila nilai regresi lebih tinggi maka di dalamnya terdapat multikolinearitas
  - D) Melakukan korelasi antara variabel variabel independen yang akan diteliti dan bilai nilai korelasi independen lebih dari 0,10 maka terdapat multikolinearitas.

#### **b. Uji Heteroskedastisitas**

Uji Heteroskedastisitas digunakan untuk menguji apakah dalam model regresi mempunyai variabel pengganggu yang tidak konstan. Model regresi yang baik adalah model yang homoskedastisitas atau tidak terjadi heteroskedastisitas

(Widarjono, 2013). Adanya masalah heteroskedastisitas dalam data dapat di deteksi menggunakan *Glejser heteroskedasticity* yang ada pada program eviews. Uji ini pada hasil regresi dengan menggunakan prosedur *equations* dan metode OLS. Hasil yang perlu diperhatikan dari uji ini adalah apabila nilai F dan *Obs\*Rsquared*, dibandingkan dengan *Obs\*Rsquared* dengan  $X^2$  (*chi-squared*) tabel. Dan apabila nilai *Obs\*Rsquared* lebih kecil dari pada nilai  $X^2$  Tabel maka tidak terdapat Heteroskedastisitas pada model yang diteliti.

### 3. Uji Statistik

#### a. Uji Signifikansi Simultan (Uji F)

Uji Statistik F digunakan untuk mengetahui pengaruh semua variabel bebas secara serentak terhadap variabel terikatnya. Berikut ini adalah tahap tahap untuk Uji F (Gujarati, 2012):

1. menentukan hipotesisnya terlebih dahulu

$H_0$  :  $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$ , diduga tidak ada pengaruh sama sekali antara variabel bebas terhadap variabel terikat.

$H_1$  :  $\beta_1 \neq \beta_2 \neq \beta_3 \neq \beta_4 \neq \beta_5 \neq 0$ , diduga tidak ada pengaruh sama sekali antara variabel bebas terhadap variabel terikat.

2. Hitunglah nilai F Hitung dan carilah nilai F tabel dengan menggunakan rumus sebagai berikut:

$$F = \frac{R^2 (K-2)}{(1-R^2)(n-k+1)}$$

Dimana:

$R^2$  = Koefisien determinasi  
 N = Jumlah Observasi  
 K = Jumlah Variabel

3. Menurut Gujarati (2012) menyatakan bahwa dengan membandingkan nilai F hitung dengan nilai F tabel untuk mengambil sebuah keputusan yang akan menolak atau menerima  $H_0$

- a) Jika  $F_{obs} > F_{table (a;k-1,n-k)}$  atau signifikansi F kurang dari  $\alpha = 0,05$ , maka  $H_0$  ditolak dan  $H_1$  diterima. Hal ini berarti bahwa variabel bebas secara bersama-sama berpengaruh terhadap variabel terikat yaitu angka kejahatan umum
- b) Jika  $F_{obs} < F_{table (a;k-1,n-k)}$  atau signifikansi F kurang dari  $\alpha = 0,05$ , maka  $H_0$  diterima dan  $H_1$  ditolak. Hal ini berarti bahwa variabel bebas secara bersama-sama tidak berpengaruh terhadap variabel terikat yaitu angka kejahatan umum.

#### **b. Uji Signifikansi Parsial (Uji t)**

Menurut Gujarati (2012) uji signifikansi parsial (uji t) merupakan sebuah prosedur yang digunakan untuk untuk menguji kebenaran atau kesalahan dari hasil hipotesis nol dari sampel. Ide dasar pengujian ini dilatar belakangi oleh uji statistic (estimator) dari distribusi sampel dari suatu statistik di bawah hipotesis nol. Keputusan untuk  $H_0$  diterima atau ditolak dibuat berdasarkan nilai uji statistik yang diperoleh dari data yang telah ada. Dalam uji signifikansi parsial, sebuah statistic dikatakan signifikan secara statistik apabila nilai dari uji statistiknya berada di daerah tolak. Sebaliknya, sebuah pengujian dikatakan

tidak signifikan secara statistik, jika nilai dari uji statistiknya berada di daerah penerimaan.

**c. Uji *Goodness of Fit* (Koefisien Determinasi/Adjusted R<sup>2</sup>)**

Koefisien Determinan (R<sup>2</sup>) merupakan ukuran ringkasan yang dapat menginformasikan kepada kita seberapa baik sebuah garis regresi sampel sesuai dengan data (Gujarati, 2012) Adjusted R<sup>2</sup> dapat dirumuskan dengan:

$$R_2 = \frac{\sum(\hat{Y}_1 - \bar{Y})^2}{\sum(Y_1 - \bar{Y})^2}$$

Adapun menurut (Gujarati, 2012) terdapat dua sifat yang dimiliki oleh R<sub>2</sub> yaitu:

- a) Besaran yang dihasilkan tidak pernah negative
- b) Batasannya yaitu  $0 \leq R_2 \leq 1$ . Jika R<sub>2</sub> bernilai 1, maka kesesuaian garis yang dimilikinya tepat. Tetapi jika R<sub>2</sub> memiliki nilai nol

