

BAB III

METODOLOGI PENELITIAN

A. Tujuan Penelitian

Berdasarkan masalah-masalah yang telah peneliti rumuskan, maka tujuan penelitian ini adalah untuk mengetahui besarnya pengaruh jumlah penduduk, investasi dan pengeluaran pemerintah terhadap Ketimpangan antar Provinsi di Jawa dan Bali.

B. Objek dan Ruang Lingkup Penelitian

Penelitian mengambil Provinsi di Jawa Bali sebagai objek studi dan sekaligus sebagai lokasi penelitian. Lokasi ini diambil dengan pertimbangan provinsi ini mempunyai perekonomian yang baik dan kondisi sumber daya yang banyak di Indonesia, namun kesejahteraan antar regional tidak merata.

Objek penelitian ini adalah Provinsi Banten, DKI Jakarta, Jawa Barat, Jawa Tengah, Yogyakarta, Jawa Timur dan Bali untuk melengkapi informasi data yang diperlukan maka peneliti menggunakan data dari Badan Pusat Statistik (BPS) Pusat, Badan Pusat Statistik (BPS) daerah, dan Badan Koordinasi Penanaman Modal.

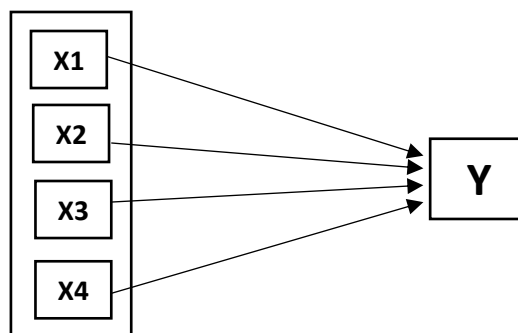
Ruang lingkup penelitian ini adalah mengkaji hubungan antara jumlah penduduk, investasi dan Pengeluaran Pemerintah terhadap ketimpangan di Provinsi Jawa dan Bali. Penelitian ini menggunakan data dari kabupaten/kota Provinsi di Jawa dan Bali diantaranya PRDB, investasi PMA dan PMDN, Jumlah penduduk, dan Pengeluaran Pemerintah.

Penelitian ini dilakukan pada bulan Januari-April 2018 karena merupakan waktu yang efektif bagi peneliti untuk melaksanakan penelitian sehingga peneliti dapat fokus pada saat penelitian dan keterbatasan peneliti dalam waktu, tenaga, dan materi.

C. Metode Penelitian

Metode Penelitian yang digunakan dalam penelitian ini adalah metode *ekpos fakto* dan jenis data yang digunakan adalah data sekunder. Metode *Ex Post Facto* adalah metode penelitian yang dilakukan untuk mengetahui peristiwa yang telah terjadi dan kemudian meruntut kebelakang untuk mengetahui faktor-faktor yang dapat menimbulkan kejadian tersebut (Sugiyono, 2004, p. 7).

Penelitian ini melihat ketimpangan seluruh Provinsi di Jawa dan Bali menggunakan indeks Williamson. Variabel yang digunakan dalam penelitian ini adalah jumlah penduduk, investasi asing dan dalam negeri serta Pengeluaran Pemerintah. Dari penjelasan tersebut maka variabel yang ditetapkan terhadap ketimpangan regional di Jawa dan Bali dapat dilihat pada gambar berikut ini:



Gambar III.1
Konstelasi Penelitian

Keterangan :

- Y : Ketimpangan Regional
- X 1 : Jumlah Penduduk
- X 2 : Investasi PMA (Penanam Modal Asing)
- X 3 : Investasi PMDN (Penanam Modal Dalam Negeri)
- X 4 : Pengeluaran Pemerintah

D. Jenis dan Sumber Data

Jenis data yang digunakan dalam penelitian ini adalah data sekunder yang bersifat kuantitatif, yaitu data yang telah tersedia dalam bentuk angka. Sedangkan data yang digunakan dalam penelitian ini adalah data runtut waktu (*time series*) dan data deret lintang (*cross section*). Menurut Nachrowi (2006, p. 309) Data *time series* adalah data yang dikumpulkan dari waktu ke waktu terhadap suatu individu, sedangkan data *cross section* adalah data yang dikumpulkan dalam satu waktu terhadap banyak individu. Data *time series* sebanyak tujuh tahun dari tahun 2010 sampai 2016 dan data *cross section* dari seluruh Provinsi di Jawa dan Bali. Data sekunder tersebut diperoleh dari berbagai publikasi Badan Pusat Statistik, dan Badan Pusat Statistik daerah.

E. Operasional Variabel Penelitian

a. Ketimpangan

1. Definisi Konseptual

Ketimpangan Regional adalah tingkat kesenjangan yang terjadi antar wilayah dari segi pemerataan pembangunan yang menghasilkan wilayah maju (*Development Region*) dan wilayah terbelakang (*Underdevelopment Region*).

2. Definisi Operasional

Ketimpangan Regional diukur melalui indeks ketimpangan Williamson tahun 2010 – 2016 di setiap Provinsi Pulau Jawa dan Bali. Hasil data yang diperoleh berdasarkan perhitungan peneliti.

b. Jumlah Penduduk

1. Definisi Konseptual

Jumlah Penduduk adalah total masyarakat yang bersama – sama mendiami suatu wilayah geografis tertentu.

2. Definisi Operasional

Jumlah penduduk yang digunakan dalam penelitian ini merupakan jumlah penduduk di setiap Provinsi Pulau Jawa dan Bali tahun 2010 – 2016 yang datanya diperoleh dari BPS.

c. Investasi

1. Definisi Konseptual

Investasi dari sisi Pemerintah merupakan upaya penempatan sejumlah dana untuk meningkatkan pertumbuhan ekonomi diberbagai sektor dalam rangka memajukan kesejahteraan umum.

2. Definisi Operasional

Dalam penelitian ini investasi dipisahkan menjadi PMA (Penanaman Modal Asing) dan PMDN (Penanaman Modal Dalam negeri) untuk mengetahui pengaruh dari masing – masing investasi tersebut. Data investasi yang diperoleh merupakan realisasi menurut lokasi mulai dari tahun 2010 – 2016 dari tiap provinsi di Pulau Jawa dan Bali. Satuan

yang digunakan adalah satuan dollar untuk PMA dan rupiah untuk PMDN.

d. Pengeluaran Pemerintah

1. Definisi Konseptual

Pengeluaran pemerintah merupakan seluruh belanja yang dikeluarkan pemerintah yang disusun dalam APBD dalam rangka pembangunan daerah dan masyarakat.

2. Definisi Operasional

Pengeluaran Pemerintah adalah realisasi pengeluaran Pemerintah Daerah Kota dan Kabupaten di Provinsi Jawa dan Bali tahun 2010 – 2016 yang datanya diperoleh dari BPS.

F. Teknik Analisis Data

1. Analisis Data Panel

Data yang digunakan dalam analisis ekonometrika dapat berupa data *time series*, data *cross section*, atau data panel. Data panel (*panel pooled data*) merupakan gabungan data *time series* dan data *cross section*. Dengan kata lain, data panel merupakan unit-unit individu yang sama yang diamati dalam kurun waktu tertentu. Jika kita memiliki T periode waktu ($t= 1,2,\dots,T$) dan N jumlah individu ($i= 1,2,\dots,N$), maka dengan data panel kita akan memiliki total unit observasi sebanyak NT . Jika jumlah unit waktu sama untuk setiap individu, maka data disebut *balanced panel*. Jika sebaliknya, yakni jumlah unit waktu berbeda untuk setiap individu, maka disebut *unbalanced panel*. Penggunaan data panel pada dasarnya merupakan solusi akan

ketidaktersediaan data *time series* yang cukup panjang untuk kepentingan analisis ekonometrika.

Sesuai dengan tujuan penelitian yaitu untuk mengetahui pengaruh variabel bebas yaitu Jumlah Penduduk (X1), Investasi Modal Asing (X2), Investasi Modal Dalam Negeri (X3), dan Pengeluaran Pemerintah (X4) terhadap variabel terikat yaitu Ketimpangan Regional (Y), maka dapat dinyatakan dalam bentuk persamaan regresi linier berganda seperti berikut :

Persamaan Regresi :

$$\log_{IW}_{it} = a + \beta_1 \cdot \log_{Pop}_{it} + \beta_2 \cdot \log_{PMA}_{it} + \beta_3 \cdot \log_{PMDN}_{it} + \beta_3 \cdot \log_{Spend}_{it} + e \quad (2)$$

Dimana :

\log_{IW} = Ketimpangan Regional (Y)

\log_{Pop}_{it} = Jumlah Penduduk (X₁)

\log_{PMA}_{it} = Investasi PMA (X₂)

\log_{PMDN}_{it} = Investasi PMDN (X₃)

\log_{Spend}_{it} = Pengeluaran pemerintah (X₄)

a = Konstanta

$\beta_1, \beta_2, \beta_3, \beta_4$ = Koefisien variabel X₁, X₂, X₃ dan X₄

e = Variabel gangguan (*error term*)

a) Estimasi Model Regresi Panel

Estimasi model regresi data panel terdapat tiga spesifikasi model yang mungkin digunakan, yakni model *common effect*, *fixed effect*, dan *random effect*.

1. Model Common Effect

Model *common effect* atau *pooled regression* merupakan model regresi data panel yang paling sederhana. Model ini pada dasarnya mengabaikan struktur panel dari data, sehingga diasumsikan bahwa perilaku antar individu sama dalam berbagai kurun waktu atau dengan kata lain pengaruh spesifik dari masing-masing individu diabaikan atau dianggap tidak ada. Dengan demikian, akan dihasilkan sebuah persamaan regresi yang sama untuk setiap unit *cross section*. Persamaan regresi untuk model *common effect* dapat dituliskan sebagai berikut

$$Y_{it} = \alpha + \beta X_{it} + u_{it} \quad i = 1, 2, \dots, N \quad t = 1, 2, \dots, T \quad (3)$$

Keterangan:

Y : variabel dependen

α : koefisien regresi

X : variabel independen

β : estimasi parameter (koefisien)

u : *error term*

N : jumlah (individu)

T : jumlah periode waktu

Berdasarkan asumsi struktur matriks varians-kovarians residual, maka pada model *common effect* metode yang dapat digunakan, yaitu *Ordinary Least Square* (OLS), jika struktur matrik varians-kovarians

residualnya diasumsikan bersifat homoskedastik dan tidak ada *cross sectional correlation*.

2. Model *Fixed Effect*

Jika model *common effect* cenderung mengabaikan struktur panel dari data dan pengaruh spesifik masing-masing individu, maka model *fixed effect* adalah sebaliknya. Pada model ini, terdapat efek spesifik individu α_i dan diasumsikan berkorelasi dengan variabel penjelas yang teramati X_{it} . Pada model *fixed effect* metode estimasi yang dapat digunakan, yaitu *Ordinary Least Square* (OLS/LSDV), jika struktur matriks varians-kovarians residualnya diasumsikan bersifat homoskedastik dan tidak ada *cross sectional correlation*.

3. Model *Random Effect*

Pada model *random effect*, efek spesifik dari masing-masing individu α_i diperlakukan sebagai bagian dari komponen *error* yang bersifat acak dan tidak berkorelasi dengan variabel penjelas yang teramati X_{it} . Dengan demikian, persamaan model *random effect* dapat dituliskan sebagai berikut:

$$Y_{it} = \alpha + \beta X_{it} + w_{it}; i = 1, \dots, N; t = 1, \dots, T \quad (4)$$

Dimana

$$W_u = \alpha + u_{it}; E(w_u) = 0; E(w_u^2) = \sigma^2 + \sigma_u^2; E(w_u w_{jt-1}) = 0; i \neq j$$

Meskipun komponen *error* w_{it} bersifat homoskedastik, nyatanya terdapat korelasi antara w_{it} dan w_{it-s} (*equicorrelation*). Karena itu, metode

OLS tidak bisa digunakan untuk mendapatkan estimator yang efisien bagi model *random effect*. Metode yang tepat untuk mengestimasi model *random effect* adalah *Generalized Least Square* (GLS) dengan asumsi homoskedastik dan tidak ada *cross sectional correlation*.

Analisis regresi dalam penelitian dilakukan empat kali, yaitu pengaruh masing masing variabel x yaitu jumlah penduduk, PMA, PMDN, dan pengeluaran pemerintah terhadap ketimpangan regional, model regresi yang diuji sebagai berikut:

$$Y = \alpha + \beta X_1 + u \quad (5)$$

$$Y = \alpha + \beta X_2 + u \quad (6)$$

$$Y = \alpha + \beta X_3 + u \quad (7)$$

$$Y = \alpha + \beta X_4 + u \quad (8)$$

Keterangan :

Y : Ketimpangan regional

X₁ : Jumlah penduduk

X₂ : Penanaman modal asing

X₃ : Penanaman modal dalam negeri

X₄ : Pengeluaran pemerintah

α : konstanta

β : parameter (koefisien)

u : *error term*

2. Penyeleksian Model Estimasi Data Panel

Menurut Gujarati (2004, p. 650-651) untuk memilih model mana yang paling tepat digunakan untuk pengolahan data panel, maka terdapat beberapa pengujian yang dapat dilakukan, antara lain *chow test*, *hausman test* dan *LM test*.

a. Chow Test

Chow Test adalah pengujian untuk memilih apakah model yang digunakan model *common effects* atau *fixed effects*. Dalam pengujian ini dilakukan dengan hipotesis sebagai berikut:

H₀: Model Common Effects

H_a: Model Fixed Effects

Statistik uji yang digunakan merupakan uji F, yaitu dengan membandingkan antara nilai F-statistik dan nilai F-tabel. Nilai F-tabel diperoleh dari $\{(\alpha, df_1(n-1), df_2(nt-n-k))\}$. Dimana α adalah taraf signifikansi (0.05), n adalah jumlah provinsi, t adalah jumlah waktu, nt adalah jumlah provinsi dikali waktu (jumlah observasi), dan k adalah jumlah variabel independen. Dasar penolakan terhadap hipotesis di atas adalah dengan membandingkan perhitungan F-statistik dan F-tabel. Apa bila hasil F-statistik lebih besar dari F-tabel, maka H₀ ditolak, yang berarti model *fixed effects* yang paling baik untuk digunakan dalam mengmodel estimasi regresi data panel. Sebaliknya, apabila F-statistik lebih kecil dari F-tabel, maka H₀ diterima, yang berarti model *common effects* yang paling baik untuk digunakan dalam mengmodel estimasi regresi data panel.

Selain dengan membandingkan F-tabel dan F-statistik, dapat juga dilakukan dengan membandingkan antar nilai probabilitas dari F-statistik dan *alpha* (0,05). Apabila nilai probabilitas dari F-

statistik $> 0,05$, maka H_0 diterima yang artinya model *common effects* yang paling baik untuk digunakan. Jika sebaliknya, maka H_0 ditolak yang artinya model *fixed effects* yang paling baik digunakan.

b. Hausman Test

Hausman Test adalah pengujian statistik sebagai dasar pertimbangan kita dalam memilih apakah menggunakan *model fixed effects* atau *random effects*. Uji ini bekerja dengan menguji apakah terdapat hubungan antara *error component* dengan satu atau lebih variabel independen dalam suatu model. Hipotesis awalnya adalah tidak terdapat hubungan antara *error component* dengan variabel independen. Hipotesis dari uji Hausman adalah sebagai berikut:

H_0 : Model Random Effects

H_a : Model Fixed Effects

Statistik uji yang digunakan adalah uji *chi square*. Jika nilai *chi square*-statistik $>$ *chi square*-tabel ($\alpha, k-1$) atau nilai *p-value* kurang dari taraf signifikansi yang ditentukan, maka hipotesis awal (H_0) ditolak sehingga model yang terpilih adalah model *fixed effects*. Hal ini dilakukan untuk mengetahui apakah terdapat efek random di dalam data panel.

Dalam perhitungan uji Hausman diperlukan asumsi bahwa banyaknya kategori *cross-section* lebih besar dibandingkan jumlah variabel independen (termasuk konstanta) dalam model. Lebih

lanjut, dalam estimasi uji Hausman diperlukan estimasi variansi *cross-section* yang positif, yang tidak selalu dapat dipenuhi oleh model. Apabila kondisi-kondisi ini tidak dapat dipenuhi, maka hanya dapat digunakan model *fixed effects*.

c. Lagrange Multiplier Test

Lagrange Multiplier Test digunakan untuk menguji model apakah yang terbaik untuk digunakan dalam penelitian, yaitu untuk menguji model *common effects* dan model *random effects*. Hipotesis yang digunakan dalam uji Lagrange Multiplier adalah sebagai berikut:

$H_0 = \text{Model Random Effects}$

$H_a = \text{Model Common Effects}$

Untuk dapat menentukan jawaban dari hipotesis di atas, maka diperlukanlah perhitungan LM-statistik nya. Perhitungan LM-statistik dapat dituliskan dengan rumus sebagai berikut:

$$\text{LM - statistik} = \frac{nT}{2(T-1)} \left[\frac{T^2 \sum \bar{e}^2}{\sum e^2} - 1 \right]^2 \quad (9)$$

Keterangan:

n = jumlah *cross-section*

T = jumlah *time-series*

$\sum \bar{e}^2$ = jumlah rata-rata kuadrat residual

$\sum e^2$ = jumlah residual kuadrat

Nilai LM-statistik akan dibandingkan dengan nilai *Chi Square*-tabel dengan derajat kebebasan (*degree of freedom*) sebanyak jumlah variabel independen (bebas) dan alpha atau tingkat signifikansi sebesar 5%. Apabila nilai LM-statistik > *Chi Square*-tabel, maka H_0 di terima, yang artinya model yang dipilih adalah model *random effects*, jika sebaliknya maka H_0 ditolak, yang artinya model yang dipilih adalah model *common effects*.

3. Pengujian Asumsi Klasik

Untuk membangun persamaan regresi panel yang terbaik dari kriteria ekonometrika, perlu dilakukan penyelidikan dan penanganan adanya masalah-masalah yang berkaitan dengan pelanggaran asumsi dasar. Berikut ini adalah asumsi-asumsi yang diperlukan dalam analisis regresi:

a) Normalitas

Pengujian normalitas dilakukan untuk mengetahui apakah residual berdistribusi normal atau tidak. Hal tersebut didasarkan pada asumsi bahwa faktor kesalahan (residual) didistribusikan secara normal. Salah satu metode yang dapat digunakan untuk menguji normalitas adalah *Jarque-Bera test*. Uji statistik ini dapat dihitung dengan rumus berikut:

$$JB = n \left[\frac{\mu_3^2}{6\mu_2^3} + \frac{(\mu_4 - 3)^2}{24} \right] \quad (10)$$

Keterangan:

n : jumlah sampel

μ_2 : varians

μ_3 : skewness

μ_4 : kurtosis

Jarque-Bera test mempunyai distribusi *chi square* dengan derajat bebas dua. Jika hasil *Jarque-Bera test* lebih besar dari nilai *chi square* pada $\alpha = 5\%$, maka tolak hipotesis nol yang berarti tidak berdistribusi normal. Jika hasil *Jarque-Bera test* lebih kecil dari nilai *chi square* pada $\alpha = 5\%$, maka terima hipotesis nol yang berarti *error term* berdistribusi normal.

b) Heteroskedastisitas

Uji untuk melihat apakah terdapat ketidaksamaan varian dari residual satu pengamatan ke pengamatan yang lain. Model regresi yang memenuhi persyaratan adalah dimana terdapat kesamaan varians dari residual satu pengamatan ke pengamatan yang lain tetap atau disebut homoskedastisitas.

Metode yang bisa digunakan untuk uji heteroskedastisitas adalah Uji White, Glejser, Breusch-Pagan-Godfrey, Harvey, dan ARCH. Model memenuhi persyaratan apabila nilai probabilitas chi-square nyanya melebihi nilai alpha 0,5. (Winarno, 2011:5.14)

c) Uji Multikolineritas

Multikolinearitas adalah keadaan dimana kedua variabel independen atau lebih pada model regresi terjadi hubungan linear yang

sempurna atau mendekati sempurna. Model regresi yang baik mensyaratkan tidak adanya masalah multikolinearitas. Apabila koefisien korelasi lebih besar dari rule of thumb 0,9 maka tidak ada masalah multikolinearitas antar variabel independen.

4. Uji Hipotesis

Untuk mengetahui keberartian model regresi yang dihasilkan, hal pertama yang harus diperhatikan adalah melihat kesesuaian tanda dan nilai koefisien estimasi. Jika kriteria terpenuhi, maka selanjutnya keberartian model regresi yang diperoleh dapat dinilai dengan memperhatikan hal-hal berikut:

a). Uji Keberartian Koefisien Regresi (Uji t)

Uji t digunakan untuk mengetahui apakah variabel bebas secara parsial berpengaruh signifikan terhadap variabel tak bebasnya. Hipotesis pengujian: $H_0: \beta_i = 0, H_1: \beta_i \neq 0$.

Statistik uji yang digunakan adalah statistik uji t -student. Adapun rumusnya adalah sebagai berikut:

$$t_{hitung} = \frac{\hat{\beta}_i}{se(\hat{\beta}_i)} \quad (11)$$

Hipotesis nol ditolak jika $t_{hitung} > t_{\alpha/2; (nT-n-k-1)}$. Keputusan ini dapat juga didasarkan pada perbandingan nilai p -value dengan tingkat signifikansinya (α). Hipotesis nol ditolak jika nilai p -value lebih kecil dari (α). Hal ini berarti secara parsial variabel bebasnya dengan tingkat kepercayaan sebesar $(1 - \alpha) \times 100$ persen.

b). Uji keberartian Regresi (Uji F)

Uji F digunakan untuk menguji apakah variabel independen secara bersama-sama berpengaruh secara signifikan terhadap variabel dependen. Metode yang digunakan dalam uji ini adalah dengan cara membandingkan antara F_{hitung} dengan F_{tabel} atau $F_{(\alpha; n+k-1; nT-n-k)}$ pada tingkat kesalahan 5% dengan hipotesis:

$$H_0: \beta_1 + \beta_2 = 0$$

$$H_1: \beta_1 + \beta_2 \neq 0$$

Hipotesis nol ditolak jika $F_{hitung} > F_{tabel}$, maka seluruh variabel independen berpengaruh signifikan terhadap variabel dependen secara simultan dan sebaliknya. Untuk menguji kedua hipotesis ini digunakan nilai statistik F yang dihitung dengan rumus sebagai berikut:

$$F = \frac{R^2 / (k - 1)}{(1 - R^2) / (n - k)} \quad (12)$$

Keterangan:

R^2 : koefisien determinasi

k : jumlah variabel bebas

n : jumlah data

c). Perhitungan Koefisien Determinasi (R^2)

R^2 digunakan untuk mengukur kebaikan atau kesesuaian suatu model persamaan regresi. Besaran R^2 dihitung dengan rumus:

$$R^2 = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2} = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS} \quad (13)$$

Sedangkan $R^2_{adjusted}$ dihitung dengan rumus:

$$\bar{R} = 1 - (1 - R^2) \frac{nT-1}{nT-n-k} \quad (14)$$

Keterangan:

ESS : jumlah kuadrat yang dijelaskan

RSS : jumlah kuadrat residual

TSS : jumlah kuadrat total

n : jumlah observasi

T : jumlah periode waktu

k : banyaknya variabel bebas tanpa intersep

Adjusted R² digunakan karena sudah menghilangkan pengaruh penambahan variabel bebas dalam model, karena R^2 akan terus naik seiring dengan penambahan variabel bebas. Penggunaan *adjusted R²* sudah memperhitungkan jumlah derajat bebas.