

## **BAB III**

### **METODE PENELITIAN**

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

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

Objek dari penelitian ini adalah seluruh provinsi di Indonesia, kecuali 1 provinsi yaitu Provinsi Kalimantan Utara. Kalimantan Utara tidak diikutsertakan dalam penelitian ini karena provinsi tersebut merupakan daerah pemekaran baru yang diresmikan tahun 2012, sehingga terdapat hambatan berupa ketidaklengkapan data.

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

Variabel dependen yang digunakan dalam penelitian ini yaitu ketimpangan pembangunan wilayah yang diukur dengan Indeks Williamson, sedangkan untuk variabel independen yang digunakan adalah Penanaman Modal Asing (PMA), Penanaman Modal Dalam Negeri (PMDN), Indeks Pembangunan Manusia (IPM), aglomerasi, dan Tingkat Pengangguran Terbuka (TPT).

#### **B. Jenis dan Sumber Data**

Jenis data yang digunakan di dalam penelitian ini yaitu data panel dan tergolong dalam jenis data kuantitatif. Data panel yang dimaksudkan yaitu berupa gabungan dari data runtut waktu (*time series*) dan data silang

(*cross section*). Menurut Widarjono dalam Basuki dan Yuliadi (2015), penggunaan data panel dalam sebuah observasi memiliki beberapa keuntungan seperti mampu menyediakan data yang lebih banyak sehingga akan lebih menghasilkan *degree of freedom* yang lebih besar dan dapat mengatasi masalah yang timbul ketika ada masalah penghilangan variabel (*omitted-variable*). Data panel yang digunakan pada penelitian ini berbentuk data tahunan dari tahun 2010 sampai dengan tahun 2016 yang dipublikasikan oleh Badan Pusat Statistik (BPS) Indonesia.

#### **A. Teknik Pengumpulan Data**

Data yang diolah penulis dalam melakukan penelitian ini diperoleh dengan menggunakan metode *library research* atau studi pustaka, yaitu teknik pengumpulan data dengan cara menggunakan bahan kepustakaan seperti buku atau laporan tahunan yang dicetak, serta dari *website* yang dipublikasikan oleh Badan Pusat Statistik (BPS) Indonesia.

#### **B. Definisi Operasional Variabel Penelitian**

Di dalam penelitian ini penulis menggunakan metode analisis regresi data panel untuk mengetahui seberapa besar pengaruh variabel bebas (*independent*) dalam mempengaruhi variabel terikat (*dependent*). Penulis menggunakan Indeks Williamson sebagai ukuran ketimpangan daerah yang kemudian dijadikan sebagai variabel dependen, sedangkan untuk variabel independennya penulis menggunakan Penanaman Modal Asing (PMA), Penanaman Modal Dalam Negeri (PMDN), Indeks Pembangunan Manusia (IPM), aglomerasi, dan Tingkat Pengangguran

Terbuka (TPT). Berikut ini merupakan penjelasan mengenai masing-masing variabel yang digunakan:

## 1. Variabel Terikat (*Dependent Variable*)

### a. Ketimpangan Wilayah

Ketimpangan wilayah merupakan fenomena dimana tingkat kemakmuran di suatu wilayah sangat berbeda dengan kemakmuran yang berada di wilayah lain. Hal ini diakibatkan oleh perbedaan potensi daerah yang berbeda-beda, perbedaan kondisi demografis, serta perbedaan kondisi sosial budaya. Data yang digunakan dalam penelitian ini merupakan perhitungan dari Indeks Williamson yang dihitung dari PDRB beserta jumlah penduduk tiap provinsi dari tahun 2010 sampai tahun 2016.

Indeks Williamson sendiri merupakan indeks yang digunakan sebagai ukuran ketimpangan antar daerah. Poin yang ditampilkan dari Indeks Williamson adalah antara angka 0 sampai dengan angka 1, dengan poin maksimal 1 yang menunjukkan ketimpangan antar daerah yang tinggi. Untuk mengukur ketimpangan antar daerah sendiri, formula yang digunakan adalah sebagai berikut:

$$IW = \frac{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2 \left(\frac{f_i}{n}\right)}}{\bar{y}} \dots\dots\dots(3.1)$$

Dimana:

$y_i$  = PDRB per kapita daerah i

$y$  = PDRB per kapita rata-rata seluruh daerah

$f_i$  = jumlah penduduk daerah i

$n$  = jumlah penduduk seluruh daerah

## 2. Variabel Bebas (*Independent Variable*)

### a. Investasi

Investasi merupakan pengeluaran penanam modal atau perusahaan untuk membeli barang-barang modal dan perlengkapan produksi untuk menambah kemampuan memproduksi barang dan jasa yang tersedia dalam perekonomian. Investasi dibedakan menjadi dua, yaitu:

- 1) Penanaman Modal Asing (PMA), adalah kegiatan penanaman modal untuk melakukan usaha di Indonesia yang dilakukan penanam modal asing.
- 2) Penanaman Modal Dalam Negeri (PMDN), adalah kegiatan penanaman modal untuk melakukan usaha di Indonesia yang dilakukan penanam modal dalam negeri.

Data yang digunakan di dalam penelitian ini merupakan data PMA dan PMDN di setiap provinsi di Indonesia dari tahun 2010 sampai 2016.

**b. Indeks Pembangunan Manusia (IPM)**

Indeks Pembangunan Manusia (IPM) merupakan pengukuran perbandingan dari harapan hidup, melek huruf, pendidikan dan standar hidup untuk semua negara di seluruh dunia. Data yang digunakan di dalam penelitian ini merupakan data IPM di setiap provinsi dari tahun 2010 sampai tahun 2016.

**c. Aglomerasi**

Aglomerasi merupakan pemusatan kegiatan ekonomi di kawasan tertentu sehingga terjadinya daerah strategis atau perkotaan. Data yang digunakan dalam penelitian ini adalah data aglomerasi dari perhitungan PDRB Atas Dasar Harga Konstan 2010 tiap provinsi dibagi dengan PDB Atas Dasar Harga Konstan 2010 Indonesia dari tahun 2010 sampai tahun 2016.

Untuk menghitung aglomerasi di dalam penelitian ini, penulis menggunakan berupa PDRB wilayah atau provinsi dibandingkan dengan PDB Indonesia tahun 2010 sampai tahun 2016. Satuan yang dihasilkan adalah dalam bentuk poin, dan formula yang digunakan adalah sebagai berikut:

$$Aglomerasi = \frac{PDRB_i}{PDRB_0} \dots\dots\dots(3.2)$$

Dimana:

$PDRB_i$  = PDRB ADHK 2010 Provinsi di Indonesia

$PDRB_o$  = PDB ADHK 2010 Indonesia

#### **d. Tingkat Pengangguran Terbuka (TPT)**

Pengangguran terbuka merupakan angkatan kerja yang sama sekali tidak memiliki pekerjaan, yang disebabkan oleh penambahan lowongan pekerjaan lebih rendah dibanding penambahan tenaga kerja. Data yang digunakan dalam penelitian ini merupakan data persentase dari pengangguran terbuka yaitu Tingkat Pengangguran Terbuka (TPT) tiap provinsi dari tahun 2010 sampai tahun 2016.

### **C. Alat Ukur Data**

Untuk mengolah data sekunder yang telah dikumpulkan, penulis menggunakan beberapa alat statistik, yaitu *Microsoft Excel 2010* dan *Stata 13*. *Microsoft Excel 2010* digunakan sebagai pengolah data yang menyangkut pembuatan tabel, sedangkan *Stata 13* digunakan sebagai pengolahan regresi.

### **D. Uji Kualitas Instrumen Data**

#### **1. Uji Asumsi Klasik**

Menurut Basuki dan Yuliadi (2015) bahwa pada regresi data panel tidak semua uji asumsi klasik yang ada pada metode

OLS digunakan, maka dari itu hanya multikolinieritas dan heteroskedastisitas saja yang diperlukan.

**a. Uji Multikolinieritas**

Menurut Basuki dan Yuliadi (2015) uji multikolinieritas merupakan uji yang dilakukan apakah terdapat kolinearitas atau interkorelasi antar variabel bebas (*independent variable*) di dalam suatu analisis regresi, dan salah satu asumsi regresi linier klasik adalah tidak adanya multikolinieritas sempurna (*no perfect multicollinearity*) yaitu tidak adanya hubungan linier antara variabel penjelas dalam suatu model regresi.

Salah satu asumsi *classical linear regression model* (CLRM) adalah tidak ada multikolinieritas di antara variabel penjelas, X. Singkatnya, multikolinieritas mengacu pada situasi di mana terdapat hubungan linear yang hampir sempurna di antara variabel X. Konsekuensi dari multikolinieritas adalah apabila terdapat kolinearitas sempurna di antara variabel X, koefisien regresi tidak dapat ditentukan dan standard error-nya tidak dapat didefinisikan (Gujarati dan Porter, 2012).

Apabila sampai terjadi masalah multikolinieritas, maka akan terjadi invalidnya signifikansi variabel maupun

besaran koefisien variabel dan konstanta. Multikolinearitas diduga terjadi apabila estimasi menghasilkan nilai  $R^2$  yang tinggi ( $> 0,8$ ), nilai F tinggi, dan nilai t-statistik semua atau hampir semua variabel penjelas tidak signifikan (Basuki dan Yuliadi, 2015).

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

Menurut Basuki dan Yuliadi (2015) uji heteroskedastisitas merupakan uji yang menilai apakah ada ketidaksamaan varian dari residual untuk semua pengamatan pada model regresi linear. Heteroskedastisitas merupakan situasi dimana varians tidak konstan. Akibat dari masalah heteroskedastisitas ini adalah varians menjadi bias sehingga terjadi invalidnya uji signifikansi. Uji ini merupakan salah satu dari uji asumsi klasik yang harus dilakukan, karena apabila asumsi heteroskedastisitas tidak terpenuhi maka model regresi yang digunakan dinyatakan tidak valid. Variabel yang digunakan di dalam penelitian dikatakan terdapat gejala heteroskedastisitas apabila nilai signifikansinya  $< 0,05$ .

Menurut Gujarati dan Porter (2012), asumsi kritis pada *classical linear regression model* (CLRM) adalah faktor gangguan memiliki varians yang sama. Jika asumsi ini tidak terpenuhi, maka terdapat heteroskedastisitas.

Heteroskedastisitas tidak menghilangkan sifat ketidakbiasan dan konsistensi karakteristik dari estimator-estimator *ordinary least square* (OLS).

## **E. Uji Hipotesis dan Analisis Data**

### **1. Uji Hipotesis**

Menurut Basuki dan Yuliadi (2015), dalam metode estimasi model regresi dengan menggunakan data panel dapat dilakukan melalui 3 pendekatan, yaitu:

#### ***a. Common Effect Model***

*Common Effect Model* merupakan pendekatan model data panel sederhana karena hanya mengkombinasikan data *time series* dan *cross section*. Pendekatan ini mengasumsikan perilaku data perusahaan sama dalam berbagai kurun waktu, sehingga tidak diperhatikan dimensi waktu maupun individunya. Metode ini menggunakan teknik kuadrat terkecil dalam mengestimasi model data panel atau *Ordinary Least Square* (OLS).

#### ***b. Fixed Effect Model***

*Fixed Effect Model* mengasumsikan bahwa perbedaan intersep dapat mengakomodasi perbedaan antar individunya. Estimasi data panel model ini menggunakan

teknik *variable dummy*, yang digunakan sebagai penangkap perbedaan intersep antar perusahaan. Model estimasi yang kedua ini juga disebut dengan teknik *Least Squares Dummy Variable* (LSDV).

**c. *Random Effect Model***

*Random Effect Model* mengestimasi data panel dimana variabel gangguan mungkin saling berhubungan antar individu dan waktu. *Error terms* masing-masing perusahaan akan mengakomodasi perbedaan intersepanya. *Random Effect Model* memiliki keuntungan untuk menghilangkan masalah heteroskedastisitas. Model ini juga dinamakan *Error Component Model* (ECM) atau teknik *Generalized Least Square* (GLS).

**2. Pemilihan Model Estimasi Analisis Data**

Menurut Basuki dan Yuliadi (2015), pemilihan model yang tepat untuk digunakan dalam mengolah data panel dapat dilakukan dengan beberapa pengujian, yaitu:

**a. *Chow Test***

*Chow test* merupakan uji yang dilakukan dalam menentukan model yang tepat, antara *Common Effect* atau *Fixed Effect*.

$H_0$  : *Common Effect Model*

$H_1$  : *Fixed Effect Model*

Apabila nilai probabilitas dalam *Chow Test*  $< 0,05$ , maka  $H_0$  ditolak dan model yang tepat untuk digunakan yaitu *Fixed Effect Model*. Namun apabila hasil dalam *Chow Test*  $> 0,05$ , maka model yang tepat digunakan yaitu *Common Effect Model*. Menurut Baltagi (dalam Mulyadi dan Basuki, 2015), perhitungan F statistik dari *chow test* didapat dengan rumus:

$$F = \frac{\frac{(SSE_1 - SSE_2)}{(n-1)}}{\frac{SSE_2}{(nt-n-k)}} \dots \dots \dots (3.3)$$

Dimana:

$SSE_1$  = *Sum Square Error* dari *Common Effect Model*

$SSE_2$  = *Sum Square Error* dari *Fixed Effect Model*

n = jumlah *cross section*

t = jumlah *time series*

k = jumlah variabel independen

Sedangkan F tabel dapat dihitung dengan rumus:

$$F \text{ tabel} = \{\alpha: df(n-1, nt-n-k)\} \dots \dots \dots (3.4)$$

Dimana:

$\alpha$  = tingkat signifikansi yang dipakai

n = jumlah *cross section*

$nt$  = jumlah *cross section* x *time series*

$k$  = jumlah variabel independen

**b. Hausman Test**

*Hausman Test* merupakan uji yang dilakukan untuk menentukan model yang tepat, antara *Fixed Effect* atau *Random Effect*.

$H_0$  : *Fixed Effect Model*

$H_1$  : *Random Effect Model*

*Hausman test* menggunakan nilai chi-square sehingga pemilihan metode data panel dapat ditentukan secara statistik. Dengan asumsi bahwa eror secara individual tidak saling berkorelasi, begitu pula dengan eror kombinasinya. Apabila nilai dari hasil *Hausman Test*  $< 0,05$  maka  $H_0$  diterima, artinya model yang tepat untuk digunakan yaitu *Fixed Effect Model*. Namun apabila hasil dari *Hausman Test*  $> 0,05$  maka model yang tepat untuk digunakan yaitu *Random Effect Model*.

**c. Lagrange Multiplier Test**

*Lagrange Multiplier Test* merupakan uji yang dilakukan untuk menentukan model yang tepat digunakan, antara *Common Effect* atau *Random Effect*.

$H_0$  : *Common Effect Model*

$H_1$  : *Random Effect Model*

Apabila dari hasil *Lagrange Multiplier Test*  $< 0,05$  maka  $H_0$  ditolak, sehingga model yang tepat digunakan adalah *Random Effect*. Namun jika hasilnya  $> 0,05$  maka model yang tepat digunakan yaitu *Common Effect*.

### 3. Model Regresi Data Panel

Penelitian ini menggunakan analisis regresi data panel untuk mengetahui pengaruh variabel Penanaman Modal Asing (PMA), Penanaman Modal Dalam Negeri (PMDN), Indeks Pembangunan Manusia (IPM), aglomerasi (Aglo), dan Tingkat Pengangguran Terbuka (TPT) terhadap ketimpangan daerah (IW) di provinsi-provinsi di Indonesia. Dari beberapa variabel tersebut, maka model penelitian yang digunakan adalah sebagai berikut:

$$IW = f(PMA, PMDN, IPM, Aglo, TPT) \dots \dots \dots (3.5)$$

$$IW_{it} = \beta_0 + \beta_1 PMA_{it} + \beta_2 PMDN_{it} + \beta_3 IPM_{it} + \beta_4 Aglo_{it} + \beta_5 TPT_{it} + e_{it} \dots \dots \dots (3.6)$$

Dikarenakan terdapat perbedaan satuan dan besaran variabel independen dalam persamaan, menyebabkan persamaan regresi di atas harus dibuat dengan model logaritma-linier (log). Maka model persamaan regresi yang baru menjadi sebagai berikut:

$$IW_{it} = \beta_0 + Log \beta_1 PMA_{it} + Log \beta_2 PMDN_{it} + Log \beta_3 IPM_{it} + \beta_4 Aglo_{it} + \beta_5 TPT_{it} + e_{it} \dots \dots \dots (3.7)$$

Dimana :

IW	= nilai ketimpangan antar daerah
LogPMA	= Penanaman Modal Asing (PMA)
LogPMDN	= Penanaman Modal Dalam Negeri (PMDN)
IPM	= Indeks Pembangunan Manusia (IPM)
Aglo	= aglomerasi
TPT	= Tingkat Pengangguran Terbuka (TPT)
$i$	= <i>cross section</i>
$t$	= <i>time series</i>
$\beta$	= koefisien
$e$	= <i>error term</i>

#### 4. Uji Statistik Analisis Regresi

##### a. Uji Koefisien Determinasi $R^2$ (*Goodness of Fit*)

Gujarati dan Porter (2012) menyatakan bahwa koefisien determinasi  $R^2$  untuk mengukur seberapa jauh kemampuan model dalam menerangkan variasi variabel independen untuk mengukur kebaikan suatu model (*Goodness of Fit*). Nilai koefisien determinasi adalah antara 0 dan 1, dan nilai yang mendekati 1 memiliki arti bahwa variabel memiliki kesesuaian yang baik terhadap modelnya.

Menilai *goodness of fit* merupakan tujuan utama dalam persamaan struktural yaitu ingin mengetahui sampai

seberapa jauh model yang dihipotesakan “fit” atau cocok dengan sampel data (Ghozali, 2014). Formula yang digunakan dalam menghitung  $R^2$  adalah sebagai berikut:

$$R^2 = \frac{(\bar{Y}_1 - \bar{Y})^2}{(Y_1 - \bar{Y})^2} \dots \dots \dots (3.8)$$

**b. Uji F-Statistik**

Uji F digunakan untuk mengetahui seberapa besar pengaruh variabel independen secara keseluruhan dalam mempengaruhi variabel dependen (Widarjono, 2013). Dalam melakukan uji ini terdapat langkah yang harus ditempuh yaitu merumuskan hipotesis dan pengambilan keputusan. Dalam merumuskan hipotesis dapat ditulis sebagai berikut:

$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = 0$ , diduga variabel independen secara keseluruhan tidak memiliki pengaruh terhadap variabel dependen.

$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 \neq 0$ , diduga variabel independen secara keseluruhan memiliki pengaruh terhadap variabel dependen.

Formula yang digunakan dalam menentukan F hitung adalah sebagai berikut:

$$F = \frac{R^2 (k-2)}{1-R^2(n-k+1)} \dots\dots\dots(3.9)$$

Dimana:

$R^2$  = koefisien determinasi

n = jumlah observasi

k = jumlah variabel

Dari formula tersebut dapat diambil kesimpulan sebagai berikut:

- 1) Jika  $F_{hitung} > F_{tabel}$  pada tingkat signifikansi 5%, maka  $H_0$  ditolak dan  $H_1$  diterima. Hal ini berarti bahwa variabel independen secara keseluruhan berpengaruh secara signifikan terhadap variabel dependen.
- 2) Jika  $F_{hitung} < F_{tabel}$  pada tingkat signifikansi 5%, maka  $H_0$  diterima dan  $H_1$  ditolak. Hal ini berarti bahwa variabel independen secara keseluruhan tidak berpengaruh secara signifikan terhadap variabel dependen.

**c. Uji t-Statistik**

Uji t dilakukan untuk mengetahui seberapa besar pengaruh variabel independen secara individual dalam menerangkan pengaruhnya terhadap variabel dependen. Uji

t merupakan sebuah pendekatan alternatif, namun juga sebagai pelengkap untuk metode interval kepercayaan dari pengujian hipotesis statistik. Atau secara umum uji signifikansi merupakan sebuah prosedur dimana hasil sampel digunakan untuk membuktikan kebenaran atau kesalahan dari hipotesis nol (Gujarati dan Porter, 2012).

Kemudian Gujarati dan Porter (2012) juga menambahkan bahwa untuk menguji pengaruh variabel independen secara individu dapat dibuat hipotesis terlebih dahulu sebagai berikut:

1) Menentukan hipotesis

Hipotesis 1

$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = 0$ , diduga tidak ada pengaruh variabel independen terhadap variabel dependen.

$H_1 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 \neq 0$ , diduga terdapat pengaruh variabel independen terhadap variabel dependen.

2) Menghitung nilai t hitung untuk setiap koefisien regresi kemudian cari nilai t tabel. Rumus untuk menghitung t hitung adalah sebagai berikut:

$$t = \frac{\beta_i - \beta}{S\beta_i} \dots\dots\dots(3.10)$$

Dimana:

$\beta_i$  = koefisien variabel independen ke-i

$\beta$  = nilai hipotesis nol

$S\beta_i$  = simpangan baku variabel independen ke-i

Basuki dan Yuliadi (2015) menjelaskan bahwa dalam mengambil keputusan menolak atau menerima  $H_0$  adalah apabila nilai t hitung > nilai t kritis, maka  $H_0$  ditolak. Namun sebaliknya, apabila nilai t hitung < nilai t kritis, maka  $H_0$  diterima.