

BAB III

METODELOGI PENELITIAN

A. Tujuan Penelitian

Berdasarkan masalah yang telah dirumuskan oleh peneliti, maka tujuan penelitian ini adalah untuk mendapatkan pengetahuan berdasarkan data dan fakta yang tepat (benar, valid) dan dapat dipercaya (dapat diandalkan/reliable) tentang faktor-faktor yang mempengaruhi ketimpangan distribusi pendapatan antar daerah kota/kabupaten Administratif di DKI Jakarta

B. Objek dan Ruang Lingkup Penelitian

Penelitian ini dilaksanakan dengan mengambil data PDRB per kapita dan jumlah penduduk untuk menghitung ketimpangan distribusi pendapatan dengan menggunakan indeks Williamson dan indeks entropy theil, pertumbuhan ekonomi, tingkat pendidikan dan tingkat pengangguran antar kota/kabupaten administratif di DKI Jakarta periode 2006-2013.

C. Metode Penelitian

Metode penelitian yang digunakan dalam penelitian ini adalah metode *expost facto* dengan jenis data yang digunakan adalah data sekunder. Data sekunder adalah jenis data yang diperoleh dan digali melalui hasil pengolahan pihak kedua dari hasil penelitian lapangannya, baik berupa data kualitatif maupun data kuantitatif¹⁷. Menurut Keurlinger, metode *Ex*

¹⁷ Muhammad Teguh, *Metodologi Penelitian Ekonomi* (Jakarta: PT. Raja Grafindo Persada, 2005), p.121

Post Facto adalah merupakan empirik yang sistematis di mana peneliti tidak dapat mengendalikan variabel bebasnya karena peristiwa itu telah terjadi atau sifatnya tidak dapat dimanipulasi.¹⁸ Metode ini dipilih karena sesuai dengan tujuan yang ingin dicapai yakni untuk memperoleh data berdasarkan runtun waktu.

Sedangkan pendekatan yang digunakan adalah pendekatan korelasional bertujuan untuk mengetahui seberapa besar pengaruh antara variabel-variabel yang diteliti yaitu ketimpangan distribusi pendapatan beserta faktor-faktor yang mempengaruhinya.

D. Jenis dan Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder berupa data tahunan per kota/kabupaten administratif PDRB per kapita dan jumlah penduduk untuk menghitung ketimpangan distribusi pendapatan dengan menggunakan indeks Williamson dan indeks entropy theil, pertumbuhan ekonomi, tingkat pendidikan, dan tingkat pengangguran yang tersedia di Badan Pusat Statistik Provinsi DKI Jakarta.

Teknik pengambilan data dalam penelitian ini adalah dengan menggunakan data panel. Data panel adalah gabungan antara data *time series* (antar waktu) dan *cross section* (antar individu/ruang). Data yang digunakan dengan menggunakan *cross section* dari 5 kota dan 1 kabupaten administrative di DKI Jakarta dan *time series* selama 8 tahun dari tahun 2006-2013. Sehingga jumlah data secara keseluruhan dengan

¹⁸Husein Umar, *Metode Penelitian untuk Skripsi dan Tesis Bisnis* Edisi 2 (Jakarta: PT Raja Grafindo Persada, 2009), p.28

menggunakan *cross section* dan *time series* dalam bentuk data panel menjadi sebanyak 48 data analisis.

E. Operasionalisasi Variabel Penelitian

1. Ketimpangan Pendapatan

a. Definisi Konseptual

Ketimpangan pendapatan adalah perbedaan pembagian pendapatan antara sekelompok masyarakat dengan sekelompok masyarakat lain. Sedangkan ketimpangan pendapatan antar regional merupakan perbedaan faktor *endowment* atau faktor yang terdapat dalam masing-masing daerah.

b. Definisi Operasional

Ketimpangan pendapatan regional merupakan perbedaan faktor *endowment* yang dimiliki oleh setiap daerah. Ketimpangan pendapatan antar regional dihitung berdasarkan indeks Williamson, dengan rumus sebagai berikut yaitu:

$$IW = \frac{\sqrt{\sum(Y_i - Y)^2 \cdot f_i/n}}{Y}$$

Dimana:

IW = Indeks Williamson

Y_i = PDRB per kapita (dalam penelitian ini adalah kecamatan)

Y = PDRB per kapita rata-rata (kabupaten/kota)

f_i = Jumlah penduduk (dalam penelitian ini adalah kecamatan)

n = Jumlah penduduk (kabupaten/kota)

2. Pertumbuhan Ekonomi

a. Definsi Konseptual

Pertumbuhan ekonomi adalah suatu proses perkembangan perekonomian suatu negara yang ditandai dengan peningkatan produksi barang dan jasa rill yang diukur dengan PDB pada perekonomian nasional dan PDRB pada perekonomian daerah.

b. Definisi Operasional

Data Pertumbuhan ekonomi dalam penelitian ini diperoleh dari data Laporan Statistik Indonesia, indikator yang dilihat yaitu pertumbuhan Produk Domestik Regional Bruto per Kapita (PDRBK) atas dasar harga konstan 2000. Data yang dipergunakan dalam penelitian ini adalah data tahun 2006 dengan tahun 2013.

3. Tingkat Pendidikan

a. Definisi Konseptual

Tingkat pendidikan adalah tingkat pendidikan tertinggi yang dicapai penduduk suatu Negara. Menurut BPS (2000), tingkat pendidikan tertinggi yang ditamatkan adalah mereka yang meninggalkan sekolah setelah mengikuti pelajaran pada kelas tertinggi suatu tingkatan sekolah sampai akhir dengan mendapatkan ijazah baik dari sekolah negeri ataupun swasta.

b. Definisi Operasional

Tingkat pendidikan adalah jenjang pendidikan tertinggi yang ditamatkan oleh penduduk. Jenjang itu terbagi menjadi jenjang pendidikan dasar, menengah pertama, menengah atas, dan perguruan tinggi. Data tingkat pendidikan diambil merupakan nilai rata-rata lama sekolah 6 kota/kabupaten di DKI Jakarta yang diambil di Badan Pusat Statistik.

4. Tingkat Pengangguran**a. Definisi Konseptual**

Pengangguran adalah seseorang yang sudah digolongkan dalam angkatan kerja (usia 15-64 tahun) yang secara aktif sedang mencari pekerjaan pada suatu tingkat upah tertentu, tetapi tidak memperoleh pekerjaan yang diinginkannya. Sedangkan menurut BPS (Badan Pusat Statistik) meliputi penduduk yang sedang mencari pekerjaan, penduduk yang sedang mempersiapkan suatu usaha, penduduk yang merasa tidak mungkin mendapatkan pekerjaan, penduduk yang sudah mempunyai pekerjaan tetapi belum mulai bekerja

b. Definisi Operasional

Data yang digunakan untuk melihat pengangguran adalah pengangguran terbuka di Provinsi DKI Jakarta tahun 2006-2013 (dalam satuan persen)

F. Teknik Analisis Data

1. Metode Analisis

a. Analisis Data Panel

Analisis dengan menggunakan panel data adalah kombinasi antar deret waktu (*time series*) dan deret lintang (*cross section*). Gujarati menyatakan bahwa untuk menggambarkan data panel secara singkat, misalnya pada data *cross section*, nilai dari satu variabel atau lebih dikumpulkan untuk beberapa unit sampel pada suatu waktu. Dalam data panel, unit *cross section* yang sama disurvei dalam beberapa waktu. Dalam model panel data, persamaan model dengan menggunakan data *cross section* dapat ditulis sebagai berikut :

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} ; i = 1, 2, \dots, N \dots\dots\dots (3.1)$$

dimana N adalah banyaknya data cross section

Sedangkan persamaan model dengan time series adalah

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 X_t + \beta_3 X_t ; t = 1, 2, \dots, T \dots\dots\dots (3.2)$$

dimana T adalah banyaknya data time-series

Mengingat data panel merupakan gabungan dari time series dan cross section, maka model dapat ditulis dengan :

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \beta_2 X_{it} + \beta_3 X_{it} + \mu_{it} \dots\dots\dots (3.3)$$

$i = 1, 2, 3, \dots, N ; t = 1, 2, 3, \dots, T$

Keterangan :

Y = variabel Ketimpangan Pendapatan
 X1 = Pertumbuhan Ekonomi
 X2 = Tingkat Pendidikan
 X3 = Tingkat Pengangguran
 i = cross section

- t = time series
 β_0 = konstanta
 $\beta_1, \beta_2, \beta_3$ = koefisien yang dicari untuk mengukur pengaruh variabel X_1 , X_2 , dan X_3
 μ = kesalahan pengganggu

Model tersebut dapat ditransformasikan kedalam persamaan logaritma :

$$\text{Ln}Y = \beta_0 + \beta_1\text{Ln}X_1 + \beta_2\text{Ln}X_2 + \beta_3\text{Ln}X_3 + \mu$$

Keterangan:

- Y = Ketimpangan Pendapatan
 β_0 = Konstanta
 X_1 = Pertumbuhan Ekonomi
 X_2 = Tingkat Pendidikan
 X_3 = Tingkat Pengangguran
 β_1, β_2 = Koefisien yang dicari untuk mengukur pengaruh variabel X_1 , X_2 , dan X_3
 μ = Kesalahan pengganggu
 Ln = logaritma natural

Pemilihan model ini didasarkan pada penggunaan model logaritma natural (Ln). Damodar Gujarati menyebutkan bahwa salah satu keuntungan dari penggunaan logaritma natural adalah memperkecil bagi variabel-variabel yang diukur karena penggunaan logaritma dapat memperkecil salah satu penyimpangan dalam asumsi klasik atau OLS (Ordinary Least Square) yaitu heterokedastisitas.¹⁹

Penggunaan data panel pada dasarnya merupakan solusi akan ketidakterediaan data time series yang cukup panjang untuk kepentingan analisis ekonometrika. Menurut Hsiao dalam Greene keunggulan

¹⁵ Damodar Gujarati, *Ekonometrika Dasar* (Jakarta: Erlangga, 1997) p.68

penggunaan data panel dibandingkan deret waktu dan kerta lintang adalah:

- 1) Dapat memberikan peneliti jumlah pengamatan yang besar, meningkatkan degrees of freedom (derajat kebebasan), data memiliki variabilitas yang besar dan mengurangi kolinearitas antara variabel penjelas, dimana dapat menghasilkan ekonometri yang efisien.
- 2) Data panel data, data lebih informatif, lebih bervariasi, yang tidak dapat diberikan hanya oleh data cross section dan time series saja.
- 3) Panel data dapat memberikan penyelesaian yang lebih baik dalam inferensi perubahan dinamis dibandingkan data cross section.

b. Estimasi Model

Dalam data panel, terdapat tiga spesifikasi model yang mungkin digunakan, yakni model *common effect*, *fixed effect*, dan *random effect*. Ketiga model tersebut, yaitu :

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:

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

Keterangan :	Y	= Variabel dependen
	α	= Koefisien regresi
	X	= Variabel independen
	i	= <i>cross section</i>
	t	= time series
	β	= estimasi parameter
	N	= jumlah (individu)
	μ	= kesalahan pengganggu (<i>error term</i>)
	T	= jumlah periode waktu

Berdasarkan asumsi struktur matriks varians-covarians residual, maka pada model *common effect*, terdapat empat metode estimasi yang dapat digunakan, yaitu:

1. *Ordinary Least Square (OLS)*, jika struktur matriks varians-kovarians residualnya diasumsikan bersifat homoskedastik dan tidak ada *cross sectional correlation*,
2. *Generalized Least Square (GLS)/ Weighted Least Square (WLS): Cross Sectional Weight*, jika struktur matriks varians-kovarians residualnya diasumsikan bersifat heteroskedastik dan tidak ada *cross sectional correlation*,

3. *Feasible Generalized Least Square (FGLS) / Seemingly Uncorrelated Regression (SUR)* atau *Maximum Likelihood Estimator (MLE)*, jika struktur matriks varians-kovarians residualnya diasumsikan bersifat heteroskedastik dan ada *cross sectional correlation*,
4. *Feasible Generalized Least Square (FGLS)* dengan proses *auto regressive (AR)* pada *error term*-nya, jika struktur matriks varians-kovarians residualnya diasumsikan bersifat heteroskedastik dan ada korelasi antar waktu pada residualnya.

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 variable penjelas yang teramati X_{it} .

Berdasarkan asumsi struktur matriks varians-kovarians residual, maka pada model *fixed effect*, terdapat 3 metode estimasi yang dapat digunakan, yaitu:

1. *Ordinary Least Square (OLS/LSDV)*, jika struktur matriks varians-kovarians residualnya diasumsikan bersifat homoskedastik dan tidak ada *cross sectional correlation*,

2. *Weighted Least Square (WLS)*, jika struktur matriks varians-kovarians residualnya diasumsikan bersifat heteroskedastik dan tidak ada *cross sectional correlation*,
3. *Seemingly Uncorrelated Regression (SUR)*, jika struktur matriks varians-kovarians residualnya diasumsikan bersifat heteroskedastik dan ada *cross sectional correlation*.

3) *Model Random Effect*

Pendekatan ini mengasumsikan *unobservable individual effect* (u_{it}) tidak berkorelasi dengan *regressor* (X) atau dengan kata lain u_{it} diasumsikan bersifat random. Sebelum model diestimasi dengan model yang tepat, terlebih dahulu dilakukan uji spesifikasi apakah *fixed effect* atau *random effect* atau keduanya memberikan hasil yang sama.

2. Uji Metode Estimasi Data Panel

Sebelum menentukan metode estimasi data panel yang akan digunakan dalam penelitian ini, maka harus dilakukan beberapa pengujian. Untuk menentukan apakah model panel data dapat diregresi dengan metode *common effect*, metode *Fixed Effect* (FE) atau metode *random effect* (RE), maka dilakukan uji-uji sebagai berikut:

1) Signifikansi fixed effect (common vs fixed)

H_0 : Common effect lebih baik dari fixed effect

H_1 : Fixed effect lebih baik dari common effect

Alpha : 5 %

Ketentuan : Tolak H_0 jika nilai p-VALUE < ALPHA atau dengan membandingkan besar F hitung dibandingkan dengan F tabel. Jika $F_{hitung} > F_{tabel}$, maka H_0 ditolak, maka fixed effect lebih baik dibandingkan dengan common effect

2) Uji Hausman

Uji Hausman digunakan untuk memilih antara metode pendekatan *Fixed effect* (FE) atau *random effect* (RE). prosedur Uji Hausman adalah sebagai berikut:

- a. Buat hipotesis dari Uji Hausman: $=random\ effect$ dan $=fixed\ effect$.
- b. Menentukan kriteria uji: apabila *Chi-square* statistic > *Chi-square* tabel dan *p-value* signifikan, maka hipotesis ditolak, sehingga metode FE lebih tepat untuk digunakan. Dan apabila *Chi-square* statistic < *Chi-square* tabel dan *p-value* signifikan, maka hipotesis diterima, sehingga metode RE lebih tepat untuk digunakan.

3. Indeks Williamson

Indeks Williamson digunakan untuk menentukan besarnya kesenjangan pendapatan. Williamson meneliti hubungan antara disparitas regional dengan tingkat pembangunan ekonomi, dengan menggunakan data ekonomi yang sudah maju dan ekonomi yang sedang berkembang, ditemukan bahwa selama tahap awal pembangunan, disparitas regional menjadi lebih besar dan pembangunan terkonsentrasi di daerah – daerah tertentu (Mudrajad Kuncoro, 1997). Metode ini diperoleh dari perhitungan pendapatan regional per kapita dan jumlah penduduk masing – masing daerah.

Rumus :

$$IW = \frac{\sqrt{\sum(Y_i - Y)^2 \cdot f_i/n}}{Y}$$

Keterangan :

IW = Indeks Williamson

Y_i = PDRB per kapita (dalam penelitian ini adalah kecamatan)

Y = PDRB per kapita (kabupaten/kota)

f_i = jumlah penduduk (dalam penelitian ini adalah kecamatan)

n = jumlah penduduk (kabupaten//kota)

Besarnya nilai ini bernilai positif dan berkisar antara angka 0 – 1. Semakin besar nilainya, maka dapat diartikan bahwa kesenjangan di wilayah tersebut besar. Sebaliknya, semakin sedikit nilai indeksnya, maka kesenjangan juga semakin rendah. Kesenjangan dikatakan tinggi apabila angka indeksnya 0,50 dan seterusnya. Apabila nilai indeksnya adalah 0, maka dapat dikatakan bahwa di daerah tersebut tidak terjadi kesenjangan,

atau dalam kata lain, daerah tersebut terjadi pemerataan sempurna. Menurut Sjafrizal (2012), Nilai $IW < 0,3$ berarti disparitas pendapatan yang terjadi tergolong rendah, IW antara $0,3 - 0,5$ termasuk kategori sedang, kemudian dikatakan tinggi jika $IW > 0,5$.²⁰

4. Pengujian Asumsi Klasik

a. Uji Normalitas

Uji Normalitas bertujuan untuk menguji apakah dalam model regresi, variabel dependen dan variabel independen keduanya mempunyai distribusi normal atau tidak. Pengambilan kesimpulan dengan Jarque-Bera test atau J-B test. Bila nilai J-B hitung $>$ nilai χ^2 tabel, maka hipotesis yang menyatakan bahwa residual μ berdistribusi normal dapat ditolak. Bila nilai J-B hitung $<$ nilai χ^2 tabel, maka hipotesis yang menyatakan bahwa residual μ berdistribusi normal tidak dapat ditolak.

b. Uji Heteroskedastisitas

Uji heteroskedastisitas bertujuan menguji apakah dalam model regresi terjadi ketidaksamaan variance dari residual satu pengamatan ke pengamatan yang lain. Jika variance dari nilai residual satu pengamatan ke pengamatan yang lain tetap, maka disebut homoskedastisitas dan jika berbeda disebut heteroskedastisitas. Model regresi yang baik adalah model yang

²⁰ Sjafrizal. *EKonomi Wilayah dan Perkotaan*. (Jakarta: PT. Rajagrafindo Persada, 2012) p.87

bersifat homoskedastisitas atau tidak terjadi heteroskedastisitas. Pengujian dilakukan melalui uji Park. Uji Park pada prinsipnya meregres residual yang dikuadratkan dan dilogaritma naturalkan dengan variabel bebas yang dilogaritma naturalkan pada model. Jika t-statistik > ttabel maka ada heteroskedastisitas, jika t-statistik < t-tabel maka tidak ada heteroskedastisitas. atau Jika nilai Prob > 0,05 maka tidak ada heteroskedastisitas, jika nilai Prob < 0,05 maka ada heteroskedastisitas.

5. Pengujian Kriteria Statistik

a. Uji Signifikansi Simultan (Uji F)

Uji F atau uji koefisien regresi secara serentak, yaitu untuk mengetahui pengaruh variabel independen secara serentak terhadap variabel dependen, apakah pengaruhnya signifikan atau tidak.²¹

Hipotesis penelitiannya:

$$H_0 : \beta_1 = \beta_2 = \beta_3 = 0$$

Artinya variabel X1, X2, dan X3 secara serentak tidak berpengaruh terhadap Y.

$$H_a : \beta_1 \neq \beta_2 \neq \beta_3 \neq 0$$

variabel X1, X2, dan X3 secara serentak berpengaruh terhadap Y.

Kriteria pengambilan keputusannya, yaitu:

- a. $F_{hitung} \leq F_{tabel}$, maka H_0 diterima

¹⁶Duwi Priyanto, *SPSS Analisa Korelasi, Regresi dan Multivariate* (Yogyakarta: Gava Media, 2009), p.48

b. $F_{hitung} > F_{tabel}$, maka H_0 ditolak

Nilai F – hitung dapat diperoleh dengan rumus:

$$\frac{R^2/k-1}{(1-R^2) - (n-k)}$$

Keterangan:

R^2 = koefisien determinasi (residual)

K = Jumlah variabel independen ditambah intercept dari suatu model persamaan

N = jumlah sampel

b. Uji Signifikansi parameter Individual (Uji Statistik t)

Uji t untuk mengetahui pengaruh variabel independen secara parsial terhadap variabel dependen, apakah pengaruhnya signifikan atau tidak.²² Selain itu, uji statistik t pada dasarnya menunjukkan seberapa jauh pengaruh satu variabel independen secara individual dalam menerangkan variasi variabel dependen. Dengan uji statistik t maka dapat diketahui apakah pengaruh masing-masing variabel independen terhadap variabel dependen sesuai hipotesis atau tidak.

Hipotesis pengujian:

$$H_0 : \beta_1 \leq 0$$

$$H_1 : \beta_1 > 0$$

¹⁷ Duwi Priyanto, *Op.Cit*, p.50

Kriteria pengujian:

- a) Jika $t_{hitung} > t_{tabel}$, H_0 ditolak, maka salah satu variabel independen mempengaruhi variabel dependen secara signifikan
- b) Jika $t_{hitung} \leq t_{tabel}$, H_0 diterima, maka salah satu variabel independen tidak mempengaruhi variabel dependen secara signifikan

c. Uji Koefisien Determinasi (Uji R²)

R^2 bertujuan untuk mengetahui seberapa jauh variasi variabel independen dapat menerangkan dengan baik variasi variabel dependen. Untuk mengukur kebaikan suatu model (*goodness of fit*), digunakan koefisien determinasi (R²). Koefisien determinasi (R²) merupakan angka yang memberikan proporsi atau persentase variasi total dalam variabel tak bebas (Y) yang dijelaskan oleh variabel bebas (X) (Gujarati, 2003). Koefisien determinasi dirumuskan sebagai berikut :

$$R^2 = \frac{\sum(\bar{Y}_I - \bar{Y})^2}{\sum(Y_I - \bar{Y})^2} = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$$

Dan R^2 *adjusted* dihitung dengan rumus:

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

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 nilai R^2 ketika menilai kebaikan dan kesesuaian suatu model persamaan regresi. Penggunaan *adjusted R²* sudah memperhitungkan jumlah derajat bebas.