

BAB III

METODE PENELITIAN

A. Objek Penelitian

Penelitian ini mencakup 33 provinsi di Indonesia yakni Aceh, Sumatra Utara, Sumatera Barat, Riau, Jambi, Sumatera Selatan, Bengkulu, Lampung, Bangka-Belitung, Kepulauan Riau, Jakarta, Jawa Barat, Jawa Tengah, DIY, Jawa Timur, Banten, Bali, NTB, NTT, Kalbar, Kalteng, Kalsel, Kaltim, Sulawesi Utara, Sulawesi Tengah, Sulawesi Selatan, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat, Papua.

B. Jenis Data

Pada penelitian ini, penulis menggunakan data sekunder. Data sekunder adalah data yang dikumpulkan dari berbagai sumber yang telah ada sehingga penulis hanya menggunakan data tersebut. Data sekunder diperoleh dari sumber publikasi Badan Pusat Statistik (BPS) untuk data yang digunakan dalam penelitian ini berupa Tingkat Partisipasi Angkatan Kerja (TPAK), Produk Domestik Regional Bruto (PDRB), Upah Minimum Provinsi (UMP), Angka Partisipasi Sekolah (APS), dan Indek Pembangunan Manusia (IPM). Penelitian ini mengambil tahun dasar 2010 dan tahun 2015 sebagai periode akhir penelitian karena merupakan data terkini di Badan Pusat Statistik

(BPS). Dengan demikian, penelitian ini menggunakan data panel dengan basis data tahunan dari 33 provinsi di Indonesia pada tahun 2010-2015.

C. Teknik Pengambilan Data

Data dalam penelitian ini didapatkan dari berbagai basis data dan laporan statistik terpercaya, data variabel dependen dan independen dalam penelitian ini diambil melalui situs Badan Pusat Statistik (BPS).

D. Teknik Pengumpulan Data

Dalam penelitian ini, penulis menggunakan teknik dokumentasi, yaitu mengambil data dan informasi terkait dengan meninjau kembali laporan-laporan tertulis berupa angka dan keterangan. Teknik dokumentasi dalam penelitian ini digunakan untuk mencari data jumlah remitansi yang diperoleh Indonesia, Tingkat Partisipasi Angkatan Kerja (TPAK), Produk Domestik Regional Bruto (PDRB), Upah Minimum Provinsi (UMP), Angka Partisipasi Sekolah (APS), dan Indek Pembangunan Manusia (IPM).

E. Definisi Operasional Variabel Penelitian

Penelitian ini terdiri dari satu variabel dependen dan empat variabel independen. Variabel independen adalah variabel bebas atau bisa disebut variabel yang mempengaruhi, sedangkan variabel dependen atau variabel terikat adalah variabel yang dipengaruhi oleh variabel independen. Ada pun,

variabel dependen dan independen yang digunakan dalam penelitian ini meliputi:

1. Tingkat Partisipasi Angkatan Kerja (TPAK)

Variabel TPAK pada penelitian ini dipilih sebagai variabel dependen. TPAK adalah menggambarkan jumlah angkatan kerja dalam suatu kelompok umur sebagai persentase penduduk dalam kelompok umur tersebut yaitu membandingkan angkatan kerja dengan tenaga kerja. Data diambil dari Badan Pusat Statistik (BPS) tahun 2010-2015.

2. Produk Domestik Regional Bruto (PDRB)

Variabel PDRB pada penelitian ini berupa nilai output akhir perekonomian yang ditimbulkan oleh seluruh kegiatan ekonomi di suatu wilayah tertentu (provinsi dan kabupaten/kota), dan dalam satu kurun waktu tertentu. PDRB terbagi menjadi dua jenis yaitu PDRB atas harga berlaku (nominal) dan atas dasar harga konstan (riil). PDRB atas dasar harga konstan digunakan untuk mengetahui pertumbuhan ekonomi dari tahun ke tahun. Sedangkan PDRB atas dasar harga berlaku digunakan untuk menunjukkan kemampuan sumber daya ekonomi yang dihasilkan suatu negara. Data PDRB diperoleh dari Badan Pusat Statistik (BPS) dalam satuan milyar pada tahun 2010-2015.

3. Upah Minimum Provinsi (UMP)

Variabel UMP pada penelitian ini adalah upah minimum yang berlaku untuk seluruh kabupaten/kota di satu provinsi. Data yang

diperoleh mengacu pada Badan Pusat Statistik (BPS) pada tahun 2010-2015 dalam satuan juta Rupiah.

4. Angka Partisipasi Sekolah (APS)

Variabel APS pada penelitian ini didefinisikan sebagai angka partisipasi sekolah yang berlaku untuk seluruh kabupaten/kota di satu provinsi. Data yang diperoleh mengacu pada Badan Pusat Statistik (BPS) pada tahun 2010-2015.

5. Indeks Pembangunan Manusia (IPM)

Variabel IPM pada penelitian ini didefinisikan sebagai indikator pengukuran perbandingan dari harapan hidup, melek huruf, pendidikan dan standar hidup untuk semua negara seluruh dunia. Data yang diambil dari Badan Pusat Statistik (BPS) pada tahun 2010-2015.

F. Metode Analisis Data

Data panel merupakan gabungan data antara data silang (*cross section*) dan runtut waktu (*time series*). Widarjono (2009) bahwa penggunaan data panel dalam sebuah penelitian mempunyai beberapa keuntungan. Pertama. Data panel mampu menyediakan data yang lebih banyak sehingga lebih menghasilkan angka derajat kebebasan (*degree of freedom*) yang lebih besar. Kedua. Data panel dapat mengatasi masalah yang timbul akibat masalah penghilangan variabel (*omitted variable*).

Data panel memiliki beberapa keuntungan sebagai berikut (Wibisono, 2005):

1. Data panel dapat memperhitungkan heterogenitas individu secara eksplisit dengan mengizinkan variabel spesifik individu.
2. Kemampuan dalam mengontrol heterogenitas ini selanjutnya menjadikan data panel dapat digunakan untuk menguji dan membangun model perilaku lebih kompleks.
3. Cocok digunakan sebagai studi penyesuaian dinamis karena didasari oleh observasi *cross section* yang berulang-ulang.
4. Banyaknya jumlah observasi menyajikan data yang lebih informatif, variatif, dan kolinieritas data semakin berkurang dengan *degree of freedom* lebih tinggi sehingga hasil estimasi akan lebih baik.
5. Mempelajari model perilaku yang kompleks.
6. Digunakan untuk meminimalisir bias yang mungkin ditimbulkan oleh agregasi data individu.

Ada pun, model regresi panel dalam penelitian ini sebagai berikut:

$$\log Y_{it} = a + b_1 \log X_{1it} + b_2 \log X_{2it} + b_3 X_{3it} + b_4 X_{4it} + e_{it} \quad (3.1)$$

Keterangan:

Y : Tingkat Partisipasi Angkatan Kerja (TPAK)

a : konstanta

X1 : Produk Domestik Regional Bruto (PDRB)

X2 : Upah Minimum Provinsi (UMP)

X3 : Angka Partisipasi Sekolah (APS)

X4 : Indeks Pembangunan Manusia (IPM)

e : *Error term*

t : Waktu

I : Provinsi

G. Model Estimasi

Model estimasi dengan data panel dapat dilakukan dengan tiga tahap, antara lain:

1. Uji Asumsi Klasik

a. 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 diantara 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).

Multikolinearitas merupakan scenario statistik di mana terdapat hubungan sempurna antara variabel penjelas dan saling bergerak satu sama lain. Di dalam praktiknya, sulit untuk menghasilkan perkiraan yang dapat diandalkan dari masing-masing koefisien individu dan

melihat besarnya kesalahan pada prediksi. Dengan kata lain, hal tersebut akan berakibat pada salahnya kesimpulan tentang hubungan antar variabel. Multikolinearitas meningkatkan varian parameter perkiraan sehingga dapat menyebabkan kurangnya signifikansi variabel penjelas walaupun model yang digunakan benar. Aturan dalam multikolinearitas adalah jika nilai VIF melebihi 5 atau 10, maka hasil regresi mengandung multikolinearitas (Montgomery, 2001).

b. Heteroskedastisitas

Uji heteroskedastisitas bertujuan untuk menguji apakah model regresi mempunyai variabel pengganggu yang tidak konstan atau heteroskedastisitas. Model regresi yang baik mengandung homokedastisitas atau tidak terjadi heteroskedastisitas. Gejala ini lebih sering terjadi pada data *cross section* (Widarjono, 2013). Menurut Widarjono (2013), varian variabel pengganggu yang tidak konstan atau heteroskedastisitas disebabkan oleh residual pada variabel independen di dalam model. Ada pun bentuk fungsi variabel gangguan adalah sebagai berikut:

$$\sigma_i^2 = \sigma_i^2 x_i^2 e^{ui} \quad (3.2)$$

Di mana $e = 2,718$

Uji heteroskedastisitas dapat dilakukan melalui uji White dengan meregresi terhadap residual kuadrat yang prosedurnya sebagai berikut:

H_0 : Tidak ada heteroskedastisitas

H1 : Ada heteroskedastisitas

Jika nilai signifikansi lebih besar dari drajat kepercayaan 0,05, maka dapat dikatakan tidak terjadi heteroskedastisitas. Sedangkan uji autokorelasi dan normalitas sebaiknya tidak dilakukan karena hasilnya tidak akan memberikan makna sama sekali. Hal ini disebabkan karena pada dasarnya uji normalitas digunakan hanya pada data primer dan uji autokorelasi untuk data *time series* dengan periode waktu yang 20 sampai 30 tahun lebih (Baltagi, 2008). Sedangkan pada penelitian ini, penulis menggunakan data sekunder berbasis data panel dengan kurun waktu hanya 10 tahun.

2. Pemilihan Model

a. *Common Effect*

Common effect adalah model data panel yang paling sederhana karena hanya mengombinasikan data *cross section* serta *time series*. Model ini dapat menggunakan pendekatan *Ordinary Least Square* (OLS) atau teknik kuadrat terkecil untuk mengestimasi model pada data panel. Berikut persamaan regresi model *common effect*:

$$Y_{it} = \alpha + X_{it}\beta + \varepsilon_{it} \quad (3.3)$$

Di mana:

i = 33 provinsi di Indonesia

t = 2010 hingga 2015

Proses estimasi dilakukan secara terpisah setiap *cross unit section* yang dapat dilakukan dengan asumsi komponen *error* pada kuadrat terkecil.

b. Fixed Effect

Model ini menjelaskan bahwa antar individu memiliki efek berbeda yang bisa diakomodasikan melalui intersepnya. Dalam model ini, setiap parameter merupakan parameter yang tidak diketahui dan akan diestimasi dengan teknik variabel *dummy* yang dinamakan *Least Square Dummy Variable* (LSDV). LSDV mampu mengakomodasikan efek waktu yang sistematis. Hal ini dilakukan melalui penambahan variabel *dummy* di dalam model.

c. Random Effect

Model ini menjelaskan efek spesifik dari setiap individu sebagai bagian dari komponen *error* yang bersifat acak dan tidak berkorelasi dengan variabel penjelas yang diamati. Model ini disebut dengan *Error Component Model* (ECM). Persamaan dalam model ECM ini dapat dituliskan sebagai berikut:

$$Y_{it} = a + X_{1it}\beta + w_{it} \quad (3.4)$$

i : 33 provinsi di Indonesia

t : Tahun 2010 sampai 2015

Di mana:

$$W_{it} = \varepsilon_{1it} + \mu_1; E(W_{it}) = 0; E(W_{it}^2) = \alpha^2 + \alpha_{\mu^2}; \quad (3.5)$$

$$E(W_{it}^2 W_{jt-1}) = 0; i \neq j; E(\mu_i \varepsilon_{it}) = 0; \quad (3.6)$$

$$E(\varepsilon_i \varepsilon_{is}) = E(\varepsilon_{it} \varepsilon_{jt}) = E(\varepsilon_{it} \varepsilon_{js}) = 0 \quad (3.7)$$

Meskipun komponen error bersifat homoskedastik, nyatanya terdapat kolerasi antara W_t dan W_{t-1} yakni:

$$\text{cross}(W_{it}, W_{i,(t-1)}) = \frac{\alpha_{\mu^2}}{\alpha^2 + \alpha_{\mu^2}} \quad (3.8)$$

Oleh sebab itu, metode OLS tidak dapat digunakan untuk mendapatkan estimator yang efisien bagi model *Random Effects*. Metode yang tepat untuk mengestimasi model *random effects* adalah *Generalized Least Square* (GLS) dengan asumsi homoskedastik dan tidak ada korelasi *cross sectional*.

d. Uji Chow

Uji Chow digunakan untuk menentukan apakah model *Common Effect* atau *Fixed Effect* yang paling tepat digunakan dalam estimasi data panel. Hipotesis dalam uji Chow sebagai berikut:

H_0 : *Common Effect Model* atau *pooled OLS*

H1 : *Fixed Effect Model*

Dasar penolakan terhadap hipotesis di atas ialah dengan membandingkan perhitungan F-statistik dengan F-tabel. Perbandingan digunakan apabila hasil F-statistik lebih kecil dibanding F-tabel, sehingga H0 tidak ditolak yang berarti model yang digunakan ialah *Common Effect Model*. Perhitungan F-statistik didapatkan melalui Uji Chow dengan rumus sebagai berikut (Baltagi, 2005):

$$F = \frac{\frac{(SE_1 - SSE_2)}{(n-1)}}{\frac{SSE_2}{(nt-n-k)}} \quad (3.9)$$

Di mana :

SSE1 : *Sum Square Error* dari model *Common Effect*

SSE2 : *Sum Square Error* dari model *Fixed Effect*

n : Jumlah n (*cross section*)

nt : Jumlah *cross section* dikali jumlah *time series*

k : Jumlah variabel independen

Sedangkan F-tabel didapat dari:

$$F_{tabel} = \{a: df(n-1), nt-n-k\} \quad (3.10)$$

Di mana:

a : Tingkat signifikansi yang dipakai

n : Jumlah unit *cross section*

nt : Jumlah *cross section* dikali *time series*

k : Jumlah variabel independen

e. Uji Hausman

Uji Hausman digunakan untuk memilih antara *fixed effect* atau *random effect*, Uji Hausman didapatkan melalui command `evIEWS` yang terdapat pada direktori panel (Winarno, 2009). Statistik Uji Hausman ini mengikuti distribusi statistik *Chi Square* dengan *degree of freedom* sebanyak k , dimana k adalah jumlah variabel independen. Jika nilai statistik hausman lebih besar dari nilai kritisnya maka model yang tepat adalah model *fixed effect*. Sedangkan sebaliknya bila nilai statistik hausman lebih kecil dari nilai kritisnya maka model yang tepat adalah model *random effect*.

Dasar pengambilan keputusan menggunakan Uji Hausman (*random effect vs fixed effect*), yaitu:

- a. Jika H_0 : diterima, maka model *random effect*.
- b. Jika H_0 : ditolak, maka model *fixed effect*.

3. Uji Signifikansi

Uji signifikansi dilakukan guna melihat apakah hipotesis akan ditolak atau tidak. Terdapat tiga cara dalam uji signifikansi, yaitu:

a. Uji t

Uji t atau biasa dikenal dengan uji parsial digunakan untuk menguji bagaimana pengaruh masing-masing variabel independen dengan variabel dependen. Uji ini dapat dilakukan dengan membandingkan t-tabel atau dengan melihat signifikansi masing-masing t-hitung.

b. Uji F

Uji F dilakukan guna melihat pengaruh variabel independen secara bersama-sama mempengaruhi variabel dependen dengan membandingkan F-tabel. Apabila F-hitung lebih dari F-tabel, maka H_0 ditolak.

c. Koefisien Determinasi

Koefisien determinasi dilakukan untuk mengukur seberapa jauh kemampuan sebuah model menjelaskan variasi dalam variabel dependen. Nilai koefisien determinasi berkisar antara nol dan satu. Apabila nilai R^2 nya mendekati nol berarti variasi variabel dependennya sangat terbatas. Apabila nilainya mendekati satu berarti variabel independennya dapat menjelaskan segala informasi dari variabel dependen.