

BAB III

METODE PENELITIAN

3.1 Ruang Lingkup Penelitian

Penelitian ini dilakukan dengan tujuan untuk menganalisis pengaruh dari faktor korupsi terhadap ketimpangan pendapatan di Indonesia. Penelitian ini menggunakan sampel data 33 Provinsi di Indonesia dan bukan 34 provinsi dikarenakan kurang lengkapnya data yang ada untuk provinsi terbaru, yaitu Provinsi Kalimantan Utara.

Variabel dependen dalam penelitian ini adalah ketimpangan pendapatan, yang dapat dilihat dan diukur dengan menggunakan Indeks Gini. Dalam penelitian ini variabel independen utamanya adalah variabel korupsi yang diukur dengan menggunakan rekapitulasi data perkara tindak pidana korupsi tahap penuntutan oleh Kejaksaan Republik Indonesia di 33 Provinsi di Indonesia tahun 2011-2015. Data ini dijadikan pengukur dikarenakan belum adanya perhitungan untuk Indeks Persepsi Korupsi berdasarkan provinsi di Indonesia.

Variabel independen kontrol lainnya adalah pendidikan yang diambil dari data rata-rata lama sekolah, konsumsi pemerintah, dan pertumbuhan ekonomi yang diambil dari data Produk Domestik Regional Bruto (PDRB). Penelitian ini menggunakan lingkup waktu dari tahun 2011-2015. Penggunaan data yang dimulai dari tahun 2011

dan hanya mencapai tahun 2015 dikarenakan keterbatasan data untuk mencapai tahun 2018 dan juga ketidaklengkapan data yang ada pada tahun-tahun sebelum tahun 2011.

3.2 Metode Analisis

Penelitian ini menggunakan pendekatan deskriptif kuantitatif. Pendekatan deskriptif kuantitatif didasarkan pada data sekunder dari buku dan penelitian-penelitian sebelumnya yang memiliki kaitan dengan variabel-variabel yang ada pada penelitian ini. Hal ini dilakukan untuk membuktikan adanya hubungan antar variabel dan untuk dijadikan referensi pada penelitian. Sedangkan pendekatan kuantitatif dilakukan dengan menggunakan model ekonometrika, untuk menyatakan hasil dan pembahasan penelitian secara matematis. Data dan model ekonometrik dalam analisis kuantitatif ini diolah dengan menggunakan *Microsoft Excel* dan aplikasi STATA.

3.3 Model Penelitian

Model persamaan ekonometrika yang digunakan pada penelitian ini adalah model berbasis data panel yang diadaptasi berdasarkan penelitian yang dilakukan oleh Gyimah-Brempong (2002) dengan judul “*Corruption, Economic Growth, and Income Inequality in Africa*” dengan melakukan beberapa penyesuaian. Model persamaan dalam penelitian ini adalah:

$$Gini_{it} = \beta_0 + \beta_1 Corrupt_{it} + \beta_2 Educ_{it} + \beta_3 LnGovCon_{it} + \beta_4 LnGrowth_{it} + \varepsilon_{it}$$

Keterangan:

i = Provinsi

t = Periode waktu

$Gini_{it}$ = Indeks Gini Ketimpangan Pendapatan

$Corrupt_{it}$ = Korupsi

$Educ_{it}$ = Pendidikan

$Unemp_{it}$ = Tingkat Pengangguran

$LnGovCon_{it}$ = Log Konsumsi Pemerintah

$LnGrowth_{it}$ = Log Pertumbuhan Ekonomi

ε_{it} = Variabel *Error*

3.4 Metode Pengumpulan Data

Penelitian ini menggunakan data panel yang merupakan gabungan dari data *time series* dan data *cross section*. Jenis data panel yang digunakan merupakan data sekunder, yaitu data yang diambil dari sumber kedua. Penelitian ini menggunakan data dari tahun 2011- 2015 dan mencakup 33 provinsi di Indonesia.

Sumber-sumber data yang terlampir pada penelitian ini berasal dari:

- BPS (Badan Pusat Statistik)
- Kejaksaan Agung Republik Indonesia

Jenis data yang digunakan adalah berupa angka dan informasi yang diberikan di dalam penelitian ini berasal dari berbagai jurnal, buku dan penelitian yang berkaitan dengan topik yang dipilih untuk diteliti.

3.5 Operasionalisasi Variabel

Tabel 3.1 Tabel Operasionalisasi Variabel

Variabel	Notasi	Satuan	Operasionalisasi	Sumber Data
<i>Indeks Gini</i>	GINI	Indeks	Indeks yang mengukur tingkat ketimpangan pendapatan secara menyeluruh	Badan Pusat Statistik Indonesia

Korupsi	CORRUPT	Jumlah Tindak Pidana Korupsi pada Tahap Penuntutan	Rekapitulasi Data Perkara Tindak Pidana Korupsi Tahap Penuntutan oleh Kejaksaan Republik Indonesia 33 Provinsi di Indonesia tahun 2011-2015	Laporan Tahunan Kejaksaan Republik Indonesia
Pendidikan	EDUC	Persen	Rata-rata lama sekolah berdasarkan provinsi dari tahun 2011-2015	Badan Pusat Statistik Indonesia.
Konsumsi Pemerintah	LNGOVCON	Juta Rupiah	Realisasi anggaran pengeluaran atau konsumsi pemerintah yang merupakan total dari belanja dan	Statistik Keuangan Badan Pusat Statistik Indonesia

			pembiayaan yang dilakukan pemerintah provinsi Indonesia dari tahun 2011-2015	
Pertumbuhan Ekonomi	LNGROWTH	Ribu Rupiah	Produk Domestik Regional Bruto 33 Provinsi di Indonesia dari tahun 2011-2015 dengan dasar harga konstan tahun 2010	Badan Pusat Statistik Indonesia

3.6 Metode Pengujian

3.6.1 Teknik Analisis Data

Dalam penelitian ini data yang digunakan merupakan data panel. Konsep pada data panel, yaitu data yang merupakan gabungan dari data *time series* dan data *cross section*. Data *time series* adalah data suatu objek yang memiliki runtun waktu lebih dari satu tahun, sedangkan data *cross section* adalah data dengan objek lebih dari satu yang

terdapat pada satu tahun yang sama. Gujarati, (2009) menyatakan secara istilah bahwa data panel adalah data yang terdiri dari beberapa objek dengan beberapa periode waktu. Wibisono (dalam Basuki & Prawoto, 2016) menyatakan, pemilihan penggunaan data panel memiliki beberapa keuntungan, yaitu:

- Data panel dapat memperhitungkan heterogenitas secara jelas dan spesifik
- Data panel dapat digunakan dalam menguji model yang lebih rumit dikarenakan kemampuannya dalam mengontrol heterogenitas
- Karena data panel didasari oleh data *cross section* yang berulang pada beberapa periode waktu membuat metode ini cocok digunakan sebagai *Study of Dynamic Adjustment*, yaitu memungkinkan adanya estimasi terpisah antara karakteristik individu dan karakteristik antar waktu
- Didapatkan data yang lebih informatif dan bervariasi, kurangnya multikolinearitas atau bias antar data, dan *degree of freedom* yang lebih tinggi dikarenakan banyaknya jumlah observasi pada metode ini, sehingga dapat menghasilkan hasil estimasi yang lebih efisien.

3.6.2 Penentuan Model Regresi

Wooldridge (2010) menyatakan terdapat tiga teknik dalam melakukan metode estimasi regresi dengan menggunakan data panel, yaitu:

1. *Common Effects Model*

Model *common effect* adalah teknik pendekatan data panel yang paling

sederhana. Model ini merupakan kombinasi dari data *cross section* dengan data *time series* dalam bentuk *pool*. Dalam model ini dimensi waktu maupun individu tidak terlalu diperhatikan. Data individu diasumsikan sama dalam berbagai periode waktu. Model ini, dalam mengestimasi data panelnya menggunakan metode *pooled least square*. Model ini juga memiliki *intercept* yang tetap.

2. *Fixed Effects Model*

Model ini mengasumsikan adanya *unobservable variables* dan juga adanya perbedaan *intercept* pada setiap subjek ataupun antar periode waktu. Dalam model *fixed effect*, untuk mengestimasi data panel dilakukan dengan menggunakan metode variabel *dummy*. Variabel *dummy* dalam model ini digunakan untuk membedakan antara satu subjek dengan subjek lainnya. Model estimasi ini sering disebut dengan model *Least Squares Dummy Variables* (LSDV). Model estimasi LSDV ini juga dapat digunakan untuk mengakomodasi efek waktu yang sifatnya sistemik dengan cara menambahkan variabel *dummy* waktu di dalam model.

3. *Random Effects Model*

Model ini melakukan pengakomodasian terhadap perbedaan antar individu ataupun waktu ke dalam *error*. Hal ini terjadi karena pada model ini variasi nilai dan hubungan antar individu diasumsikan *random*. Model estimasi ini sering disebut juga dengan *Error Component Model* (ECM). Model *Random Effect* ini digunakan dalam mengatasi kelemahan yang dimiliki oleh model *Fixed Effect* yang menggunakan variabel *dummy*. Metode *Generalized Least Squares* (GLS) dengan asumsi tidak

adanya *cross-sectional correlation* dan bersifat homoskedastik, merupakan metode yang paling tepat dalam melakukan estimasi pada model *random effects* ini.

3.6.3 Uji Hausman

Uji Hausman didefinisikan sebagai penentu dalam pemilihan model yang terbaik yang akan digunakan dalam regresi data panel (Wooldridge, 2010). Pengujian ini ditujukan untuk menentukan model yang paling tepat antara model *Fixed Effects* atau model *Random Effects*.

Hipotesis dalam Uji Hausman adalah sebagai berikut:

- H_0 : *Random Effect Model (REM)*
- H_1 : *Fixed Effect Model (FEM)*

Kriteria dan kesimpulan:

- $P\text{-value} < \alpha = 5\%$, H_0 ditolak, maka model yang tepat adalah *Fixed Effect Model*
- $P\text{-value} > \alpha = 5\%$, H_0 tidak dapat ditolak, maka model yang tepat adalah *Random Effects Model*

3.6.4 Koefisien Determinasi

Koefisien determinasi atau sering disebut uji R^2 (*R-squares*) adalah suatu ukuran untuk menginformasikan seberapa besar kemampuan variabel independen dalam menjelaskan variabel dependen di dalam sebuah model. Apabila nilai dari koefisien determinasi adalah 0, maka variabel dependen dalam model tersebut tidak dapat dijelaskan oleh variabel independen dan apabila nilai dari koefisien determinasi adalah 1, maka artinya variabel dependen secara keseluruhan dapat diterangkan oleh variabel independen. Sehingga dapat disimpulkan, apabila nilai koefisien determinasi mendekati angka 0, artinya adalah variabel independen dalam model tersebut kurang mampu dalam menjelaskan variabel dependen, sehingga dapat dinyatakan bahwa model persamaan tersebut kurang sempurna.

3.6.5 Uji T-Statistik

Uji T-statistik digunakan untuk menguji signifikansi dari pengaruh variabel independen terhadap variabel dependen secara individu (Wooldridge, 2010).

Hipotesis yang digunakan pada Uji T-statistik adalah:

- $H_0 : \beta = 0$, artinya variabel independen tidak mempengaruhi variabel dependen
- $H_1 : \beta \neq 0$, artinya variabel independen mempengaruhi variabel dependen

Kriteria dan kesimpulan:

- T-statistik > T-tabel = H_0 ditolak, secara parsial variabel independen mempengaruhi variabel dependen dengan signifikan

- $T\text{-statistik} < T\text{-tabel} = H_0$ diterima, secara parsial variabel independen mempengaruhi variabel dependen dengan tidak signifikan
- $-T\text{-statistik} < -T\text{-tabel} = H_0$ ditolak, secara parsial variabel independen mempengaruhi variabel dependen dengan signifikan

3.6.6 Uji F-Statistik

Uji F-statistik digunakan untuk menguji signifikansi dari pengaruh seluruh variabel independen terhadap variabel dependen secara bersama-sama.

Hipotesis yang digunakan pada Uji F-statistik adalah:

- $H_0 : \beta_i = 0, i = 1, 2, 3, \dots, n$, artinya secara bersama-sama seluruh variabel independen tidak mempengaruhi variabel dependen
- $H_1 : \beta_i \neq 0, i = 1, 2, 3, \dots, n$, artinya secara bersama-sama seluruh variabel independen mempengaruhi variabel dependen

Kriteria dan kesimpulan:

- $F\text{-statistik} > F\text{-tabel} = H_0$ ditolak, secara bersama-sama variabel independen mempengaruhi variabel dependen dengan signifikan
- $F\text{-statistik} < F\text{-tabel} = H_0$ diterima, secara Bersama-sama variabel independen mempengaruhi variabel dependen dengan tidak signifikan

3.6.7 Pengujian Masalah Asumsi Klasik

3.6.7.1 Uji Autokorelasi

Uji ini dilakukan untuk melihat adanya masalah berupa korelasi antara residual pada suatu variabel dengan residual variabel lainnya yang disebut sebagai masalah autokorelasi. Dampaknya adalah dapat membuat nilai varian untuk masing-masing variabel tidak lagi minimum dan menimbulkan adanya regresi palsu (Gujarati, 2009). Masalah autokorelasi biasanya terjadi pada data *time series*.

Hipotesis yang digunakan pada Uji Autokorelasi adalah:

- H_0 : Adanya masalah autokorelasi
- H_1 : Tidak adanya masalah autokorelasi

Kriteria dan kesimpulan:

- $(\text{Prob} > \text{Chi}^2) < \alpha = H_0$ ditolak, adanya masalah autokorelasi
- $(\text{Prob} > \text{Chi}^2) > \alpha = H_0$ diterima, tidak adanya masalah autokorelasi

Dan dikarenakan autokorelasi biasa terjadi pada data *times series*, pada data *cross section* ataupun data panel, pengujian terhadap masalah autokorelasi tidak akan memiliki pengaruh apapun (Basuki & Prawoto, 2016)

3.6.7.2 Uji Multikolinearitas

Uji ini bertujuan untuk memastikan adanya masalah pada variabel independen, dimana terdapat korelasi antar variabel independen yang mengakibatkan ketidakakuratan pada koefisien-koefisien regresi. Masalah ini disebut sebagai masalah multikolinearitas. Dalam mengidentifikasi masalah multikolinearitas, terdapat beberapa cara untuk mengidentifikasinya, diantaranya adalah:

1. Model memiliki nilai R^2 yang tinggi dengan koefisien parameter signifikansi pada T-statistik yang rendah
2. Terdapat perbedaan antara nilai koefisien variabel dengan hipotesis pada uji yang dilakukan
3. Model memiliki nilai korelasi yang lebih besar dari 0.8 (*rule of thumb*) antar variabel (Gujarati, 2009)

Untuk memperbaiki atau menghilangkan adanya masalah multikolinearitas, dapat dilakukan dengan hal-hal berikut:

- Menghilangkan variabel yang menyebabkan bias
- Menambah data baru
- Melakukan kombinasi antara data *time series* dan data *cross section*

3.6.7.3 Uji Heterokedastisitas

Uji Heterokedastisitas dilakukan untuk melihat adanya masalah heterokedastisitas pada model, ketika residual pada model yang ada tidak bersifat

konstan. Dampak dari adanya heterokedastisitas adalah, proses estimasi yang menjadi tidak efisien diikuti dengan hasil estimasi yang tetap konsisten dan tidak ada bias. Hal ini menyebabkan tidak bergunanya (*miss leanding*) hasil pada Uji T-statistik dan Uji F-statistik.

Hipotesis yang digunakan pada Uji Heterokedastis adalah:

- H_0 : tidak adanya masalah heterokedastisitas
- H_1 : adanya masalah heterokedastisitas

Kriteria dan kesimpulan:

- $(\text{Prob}>\text{Chi}^2) < \alpha = H_0$ ditolak, adanya masalah heterokedastisitas
- $(\text{Prob}>\text{Chi}^2) > \alpha = H_0$ diterima, tidak adanya masalah heterokedastisitas.

3.7 Objek Penelitian

3.7.1 Indeks Gini

Indeks Gini atau Koefisien Gini atau Rasio Gini merupakan indeks yang menunjukkan indikator kesenjangan pendapatan yang biasa digunakan untuk melihat seberapa besar penyimpangan yang terjadi dalam distribusi pendapatan di kalangan rumah tangga. Selain itu, Indeks Gini dapat dipakai untuk mengukur tingkat ketidakmerataan distribusi pendapatan penduduk di berbagai sektor dan juga negara, serta dapat menunjukkan perubahan distribusi pendapatan dalam suatu negara selama periode waktu tertentu, sehingga Indeks Gini dapat menunjukkan peningkatan atau

penurunan dari ketimpangan pendapatan di suatu negara tersebut. Rumus Indeks Gini (Dumairy, 1997):

$$0 < G < 1$$

$$G = 1 - \sum (X_{i+1} - X_i)(Y_i + Y_{i+1})$$

$$G = 1 - \sum f_i(Y_i - Y_{i+1})$$

Keterangan:

G = Gini Rasio

f_i = Proporsi Jumlah Rumah Tangga dalam kelas i

X_i = Proporsi Jumlah Kumulatif rumah tangga dalam kelas i

Y_i = Proporsi Jumlah Kumulatif Pendapatan dalam kelas i

Indeks Gini memiliki nilai berkisar 0 sampai 1. Indeks Gini yang nilainya mendekati angka 0, merupakan koefisien yang tergolong rendah dan menunjukkan adanya kesetaraan distribusi pendapatan atau kekayaan, Sedangkan Indeks Gini yang nilainya mendekati angka 1 merupakan koefisien yang tergolong tinggi dan menunjukkan ketidakmerataan distribusi. Data Indeks Gini pada penelitian ini diambil dari data Indeks Gini pada 33 Provinsi di Indonesia dari tahun 2011-2015 dan didapatkan dari Badan Pusat Statistik Indonesia.

3.7.2 Korupsi

Variabel korupsi merupakan variabel independen utama dalam penelitian ini. Variabel ini dipilih dan didasarkan oleh keempat jurnal dalam kajian empiris di penelitian ini. Variabel Korupsi adalah variabel yang cukup sulit dalam proses pengukurannya. Beberapa penelitian banyak yang menggunakan Indeks Persepsi Korupsi yang dibuat oleh *Transparency International Forum*. Namun dikarenakan tidak lengkapnya data Indeks Persepsi Korupsi untuk 33 Provinsi di Indonesia dan untuk tahun 2011-2015, membuat penelitian ini tidak menggunakan Indeks Persepsi Korupsi sebagai indikator pengukur variabel. Dalam mengukur variabel korupsi, penelitian ini memilih untuk menggunakan alat ukur objektif, yaitu rekapitulasi data perkara tindak pidana korupsi tahap penuntutan oleh Kejaksaan Republik Indonesia di 33 provinsi di Indonesia pada tahun 2011-2015. Data tersebut dapat diambil dari Laporan Tahunan Kejaksaan Republik Indonesia.

3.7.3 Pendidikan

Variabel Pendidikan dipilih berdasarkan variabel pendidikan yang ada pada model di dalam keempat jurnal yang dijadikan acuan dalam penelitian ini. Variabel pendidikan diperoleh dan diukur dari data rata-rata lama sekolah yang diambil dari 33 provinsi di Indonesia dari tahun 2011-2015 dan sumbernya dikeluarkan oleh Badan Pusat Statistik Indonesia.

Rata-rata lama sekolah didefinisikan sebagai jumlah tahun belajar penduduk usia 15 tahun ke atas yang telah menjalani dan menyelesaikan pendidikan formal (tidak termasuk tahun mengulang). Cara menghitung rata-rata lama sekolah adalah sebagai berikut:

$$MYS = \frac{1}{P_{15+}} \sum_{i=1}^{P_{15+}} (\text{Lama sekolah penduduk ke } - i)$$

Penjelasan:

P_{15+} = Jumlah penduduk berusia 15 tahun ke atas

Lama sekolah penduduk ke-i=

- a) Tidak Pernah sekolah = 0
- b) Masih sekolah di SD sampai dengan S1 = konversi ijazah terakhir + kelas terakhir -1
- c) Masih sekoah di S2/S3 = konversi ijazah terakhir + 1
- d) Tidak bersekolah lagi dan tamat di kelas terakhir = konversi ijazah terakhir
- e) Tidak bersekolah lagi dan tidak tamat di kelas terakhir = konversi ijazah terakhir + kelas terakhir -1

3.7.4 Konsumsi Pemerintah

Variabel konsumsi pemerintah dipilih berdasarkan penelitian Gyimah-Brempong (2002) dan Gyimah-Brempong & De Gyimah-Brempong (2006) yang

dijadikan acuan pada penelitian ini. Data yang digunakan untuk variabel konsumsi pemerintah dalam penelitian ini adalah data realisasi Anggaran Pendapatan dan Belanja Daerah (APBD), yaitu total realisasi dari pengeluaran untuk belanja dan pengeluaran untuk pembiayaan. Data konsumsi pemerintah ini diambil dari 33 provinsi di Indonesia dari tahun 2011-2015 dan sumbernya berasal dari Buku Statistik Keuangan yang dikeluarkan oleh Badan Pusat Statistik Indonesia. Data dinyatakan dalam satuan juta dalam mata uang Rupiah.

3.7.5 Pertumbuhan Ekonomi

Variabel pertumbuhan ekonomi adalah variabel yang dipilih berdasarkan keempat jurnal yang dijadikan acuan pada penelitian ini, yang sudah dibahas pada *sub* bagian kajian empiris. Data untuk variabel pertumbuhan ekonomi dalam penelitian ini adalah data Produk Domestik Regional Bruto di 33 Provinsi di Indonesia pada tahun 2011-2015. Produk Domestik Regional Bruto atau disingkat menjadi PDRB merupakan jumlah dari nilai tambah barang dan jasa yang diambil dari seluruh kegiatan yang ada pada seluruh sektor perekonomian di suatu wilayah/daerah. Data Produk Domestik Regional Bruto yang digunakan dalam penelitian ini adalah data PDRB dengan dasar harga konstan. PDRB atas dasar harga konstan dihitung menggunakan harga pada tahun tertentu yang dijadikan sebagai tahun dasarnya. Dalam penelitian ini data yang digunakan adalah data PDRB dengan harga konstan tahun 2010. Data

tersebut didapatkan dari Badan Pusat Statistik Indonesia dengan satuan ribu dalam mata uang Rupiah.