

### **BAB III**

## **METODOLOGI PENELITIAN**

#### **A. Objek Penelitian**

Penelitian ini mencakup 27 negara utama tujuan ekspor dan impor Indonesia, yakni Amerika Serikat, Belanda, Belgia, Inggris, Italia, Jerman, Perancis, Spanyol, Rusia, Turki, Brunei Darussalam, Filipina, Kamboja, Laos, Malaysia, Singapura, Thailand, Vietnam, India, Jepang, Korea Selatan, Pakistan, Republik Rakyat Cina, Saudi Arabia, Australia, Selandia Baru dan Afrika Selatan.

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

Pada penelitian ini, penulis menggunakan data sekunder. Data sekunder merupakan data yang didapatkan dari berbagai sumber dan penulis hanya menggunakan data yang telah ada tersebut. data tersebut diperoleh dari sumber yang telah terpercaya yaitu Bank Indonesia, *World Bank* dan *distancefromto.net* dari periode 2005-2015. Data sekunder tersebut yaitu berupa PDB negara tujuan, PDB negara asal, Jarak antar negara, jumlah populasi negara tujuan, nilai tukar negara asal , serta stabilitas politik di negara tujuan.

Pada penelitian ini menggunakan tahun 2015 menjadi periode akhir penelitian karena merupakan data terkini yang tersedia di Bank Indonesia. Dengan demikian, penelitian ini menggunakan data panel dengan basis data tahunan 27 negara tujuan ekspor dan impor Indonesia pada tahun 2005-2015.

### **C. Teknik Pengambilan Data**

Data dalam penelitian ini didapatkan dari berbagai basis data dan laporan statistik terpercaya, yakni melalui situs resmi Bank Indonesia, *World Bank* dan *distancefromto.net*.

### **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 neraca perdagangan yang diperoleh Indonesia, Produk domestik bruto negara tujuan, produk domestik bruto negara asal, jarak antar Indonesia dengan negara tujuan, jumlah populasi di negara tujuan, nilai tukar di negara asal, serta stabilitas politik negara tujuan.

### **E. Definisi Operasional Variabel Penelitian**

Variabel dalam penelitian ini terdiri dari satu variabel dependen dan enam variabel independen. Yang dimaksud variabel dependen atau variabel terikat adalah suatu variabel yang dipengaruhi oleh variabel bebas/variabel

independen, sedangkan variabel independen yaitu variabel yang menjelaskan atau mempengaruhi variabel lain. Adapun yang meliputi variabel dependen dalam penelitian ini adalah neraca perdagangan Indonesia, sedangkan variabel independen dalam penelitian ini meliputi PDB negara tujuan, PDB negara asal, jarak antar Indonesia dengan negara tujuan, jumlah populasi di negara tujuan, nilai tukar di negara asal dan stabilitas politik di negara tujuan. Berikut adalah penjabaran dari definisi operasional per variabel meliputi :

1. Neraca perdagangan

Variabel dependen yang dipilih dalam penelitian ini adalah variabel neraca perdagangan Indonesia. Neraca perdagangan adalah nilai ekspor suatu negara yang dikurangi dengan nilai impornya. Neraca perdagangan dalam penelitian ini dihitung berdasarkan jumlah nilai ekspor menurut negara tujuan dikurangi dengan nilai impor menurut negara asal yang dinyatakan dalam satuan juta USD. Data diambil dari Bank Indonesia tahun 2005-2015.

2. PDB negara asal

Variabel PDB pada penelitian ini berupa nilai akhir barang dan jasa yang diproduksi di negara Indonesia selama setahun. Data PDB yang digunakan berupa data PDB riil dengan harga konstan pada tahun 2010 yang dikonversi dari mata uang domestik ke dalam USD menggunakan kurs resmi untuk setiap tahunnya. Data PDB diperoleh dari *World Bank* dalam satuan juta USD pada tahun 2005-2015.

### 3. PDB negara tujuan

Variabel PDB pada penelitian ini berupa nilai akhir barang dan jasa yang diproduksi oleh negara tujuan ekspor dan impor selama setahun. Data PDB yang digunakan berupa data PDB riil dengan harga konstan pada tahun 2010 yang dikonversi dari mata uang domestik ke dalam USD menggunakan kurs resmi untuk setiap tahunnya. Data PDB diperoleh dari *World Bank* dalam satuan juta USD pada tahun 2005-2015.

### 4. Jarak

Variabel jarak pada penelitian ini adalah jarak Indonesia dengan negara tujuan ekspor yang dihitung dengan satuan kilometer (km). Jarak dalam hal ini merupakan jarak ibukota kedua negara yang diukur melalui udara. Data yang diperoleh mengacu pada situs *www.distancefromto.net*.

### 5. Jumlah populasi negara tujuan

Jumlah populasi yang dimaksud dalam penelitian ini mencakup total penduduk yang tinggal dalam negara tujuan. Data populasi dalam penelitian ini diperoleh dari *world bank* dengan satuan ribu jiwa dari tahun 2005-2015.

### 6. Nilai Tukar negara asal

Variabel ini adalah nilai tukar Indonesia terhadap USD. Data diperoleh dari *world Bank* pada tahun 2005-2015.

#### 7. Indeks Stabilitas Politik

Yang dimaksud dengan Indeks Stabilitas Politik pada penelitian ini adalah kemungkinan ketidakstabilan politik dan rezim pemerintahan dalam suatu negara dengan agregat nilai berkisar -10 sampai 10. Semakin besar indeksinya maka semakin baik dan stabil keadaan politik dan pemerintahan di negara tersebut. Data ketidakstabilan politik diperoleh dari Polity IV dengan satuan Skor dari tahun 2005-2015.

### **F. Metode Analisis Data**

Data panel merupakan gabungan antara data runtut waktu (*time series*) dan data silang (*cross section*). Data runtut waktu biasanya meliputi satu objek atau individu seperti harga saham, kurs mata uang, SBI, atau tingkat inflasi. Data runtut waktu juga meliputi beberapa periode, baik harian, bulanan, kuartalan, atau tahunan. Data silang terdiri dari banyak objek yang sering dikategorikan kedalam beberapa jenis data, misalnya laba, biaya iklan, laba ditahan, dan tingkat investasi dalam satu periode waktu tertentu.

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 (Baltagi, 2005)

1. Data panel mampu mengakomodasi tingkat heterogenitas variabel-variabel yang tidak dimasukkan dalam model (*unobserved individual heterogeneity*).
2. Data panel mampu mengurangi kolinearitas antar variabel.
3. Data panel dapat meminimalkan bias yang dihasilkan oleh agregasi individu karena unit data lebih banyak.

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

$$\log Y_{it} = a + b_1 \log PDB_{indo_{it}} + b_2 \log PDB_{tujuan_{it}} + b_3 \text{jarak}_{it} + b_4 \text{populasi}_{it} + b_5 \text{nilai tukar}_{it} + b_6 \text{stabilitas politik}_{it} + e_{it} \quad (3.1)$$

Keterangan:

Y : Neraca Perdagangan (Juta USD)

$\alpha$  : Konstanta

X1 : PDB negara asal (Juta USD)

X2 : PDB negara tujuan (Juta USD)

X3 : Jarak (Km)

X4 : Populasi negara tujuan (Ribu Jiwa)

X5 : Nilai Tukar negara asal (USD)

X6 : indeks stabilitas politik (Poin)

e : Error term

t : Waktu

i : Negara

## G. Model Estimasi

Metode 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 multikolienaritas dalam model, estimator masih bersifat *Best Linear Unbieased Estimator* (BLUE) namun estimator mempunyai varian dan kovarian yang besar sehingga sulit didapatkan estimasi yang tepat (Widarjono, 2013). Multikolinearitas merupakan skenario 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:

H0 : Tidak ada heteroskedastisitas

H1 : Ada heteroskedastitas

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*. Metode ini dapat menggunakan pendekatan *Ordinary Least Square* (OLS) atau teknik kuadrat terkecil untuk mengestimasi model pada data panel. Berikut persamaan regresi model *common effect*:

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

Dimana :

$i = 27$  negara tujuan Ekspor dan Impor Indonesia

$t = 2005$  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:

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

$i$  : 27 negara tujuan perdagangan Indonesia

$t$  : Tahun 2005 sampai 2015

Di mana:

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

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

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

Meskipun komponen *error* bersifat homoskedastik, nyatanya terdapat korelasi antara  $W_t$  dan  $W_{it-s}$  yakni:

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

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*

$H_1$  : *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  $H_0$  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:

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

Di mana:

a : Tingkat signifikansi yang dipakai

n : Jumlah unit *cross section*

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

k : Jumlah variabel independen

### e. Pendekatan Mundlak

Pendekatan *Mundlak* digunakan apabila terdapat variabel yang mengandung *time-invariant*. Dalam kata lain yang dimaksud dengan variabel *time-invariant* adalah variabel yang memiliki karakteristik tidak berubah terhadap waktu. Pendekatan ini digunakan sebagai pengganti uji *Hausman* dan juga digunakan apabila model mengalami heteroskedastisitas (Mundlak, 1978). Pendekatan *Mundlak* dapat dituliskan dalam persamaan sebagai berikut:

$$Y_{it} = X_{it} + \alpha_i + \varepsilon_{it} \quad (3.11)$$

Indeks  $i$  menunjukkan individu atau negara sedangkan  $t$  menunjukkan waktu.  $Y_{it}$  adalah variabel dependen dan  $X_{it}$  merupakan variabel independen,  $\varepsilon_{it}$  adalah variabel yang tidak mengandung *time-invariant*, dan  $\alpha_i$  adalah variabel yang mengandung *time-invariant*. Tujuan pendekatan Mundlak ini adalah untuk melihat hubungan antara  $\alpha_i$  terhadap  $X_{it}$ .

$$\alpha_i = \bar{X}_i \theta + V_i \quad (3.12)$$

$$E(\alpha_i | X_i) = \bar{X}_i \theta \quad (3.13)$$

Di mana  $\bar{X}_i$  adalah rata-rata dari panel  $X_{it}$ , sedangkan  $V_i$  adalah *time-invariant* yang tidak berkorelasi dengan variabel dependen.

Seperti dalam regresi, jika  $\theta = 0$ , maka  $\alpha_i$  dan variabel kovarian tidak berkorelasi. Persamaannya dapat ditulis sebagai berikut:

$$Y_{it} = X_{it}\beta + \alpha_i + \varepsilon_{it} \quad (3.14)$$

$$Y_{it} = X_{it}\beta + X_i\theta + V_{it} + \varepsilon_{it} \quad (3.15)$$

$$E(Y_{it}|X_{it}) = X_{it}\beta + X_i\theta \quad (3.16)$$

Persamaan kedua mengganti  $\alpha_i$  dengan  $X_i\theta + V_i$ . Persamaan ketiga bergantung pada fakta bahwa variabel dependen dan variabel timeinvariant bersifat independen, sehingga  $H_0 : \theta = 0$ .

### 3. Uji Signifikansi

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

#### 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-hitung dengan 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-hitung dengan F-tabel. Apabila F-hitung lebih dari 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 berikisar 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.