

### **BAB III**

#### **METODOLOGI PENELITIAN**

##### **A. Objek penelitian**

Penelitian ini difokuskan pada delapan negara penyumbang wisatawan terbanyak yaitu, Singapura, Malaysia, Australia, Amerika Serikat, Inggris Raya, Republik Cina, Jepang, dan Republik Korea. Pemilihan periode penelitian selama delapan tahun yaitu dari tahun 2009 hingga 2016.

##### **B. Jenis Data**

Dalam penelitian ini, penulis menggunakan data sekunder. Data sekunder adalah data yang dikumpulkan oleh berbagai sumber yang telah ada sehingga penulis dapat menggunakan data tersebut. Data dalam penelitian ini dapat diperoleh dari berbagai sumber kredibel seperti *World Bank*, Badan Pusat Statistik, dan Kementerian Pariwisata untuk data jumlah wisatawan mancanegara yang berkunjung ke Indonesia, PDB Indonesia, PDB negara asal wisman, nilai tukar negara asal wisman, jarak, dan populasi negara asal wisman.

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

Data dalam penelitian ini didapat dari berbagai basis data dan laporan statistik terpercaya, yaitu melalui situs resmi *World Bank*, Kementerian Pariwisata, dan Badan Pusat Statistik.

### **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 berupa jumlah wisatawan mancanegara yang berkunjung ke Indonesia, PDB negara asal wisatawan, PDB negara Indonesia, nilai tukar negara asal wisatawan, jarak, nilai tukar negara asal wisatawan dan populasi negara asal wisman.

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

Penelitian ini terdiri dari satu variabel dependen dan lima variabel independen. Variabel independen merupakan variabel bebas atau juga variabel yang mempengaruhi, sedangkan variabel dependen atau variabel terkait merupakan variabel yang dipengaruhi oleh variabel independen. Adapun variabel dependen dan independen yang digunakan dalam penelitian ini berupa:

#### 1. Jumlah Wisatawan Mancanegara

Variabel jumlah wisatawan mancanegara pada penelitian ini dipilih sebagai variabel dependen. Jumlah wisatawan mancanegara adalah jumlah wisatawan mancanegara berdasarkan negara tujuan yang berkunjung ke Indonesia per tahun. Data diperoleh dari Badan Pusat Statistik Indonesia (BPS) dan Kementerian Pariwisata tahun 2009-2016.

#### 2. PDB Per Kapita Negara Indonesia

Variabel PDB dalam penelitian ini berupa nilai akhir barang dan jasa yang diproduksi oleh negara Indonesia. Data PDB per kapita ini dihitung dalam satuan USD yang dikonversi dari mata uang domestik menggunakan *kurs* resmi untuk setiap tahunnya. Data PDB diperoleh dari *World Bank* dalam satuan juta USD pada tahun 2009-2016.

#### 3. PDB Per kapita Negara Asal Wisatawan Mancanegara

Variabel PDB pada penelitian ini berupa nilai akhir barang dan jasa yang diproduksi oleh negara asal wisatawan selama setahun. Data PDB pada penelitian ini dihitung dalam satuan USD yang dikonversi dari mata uang domestik menggunakan *kurs* resmi untuk setiap tahunnya. Data PDB diperoleh dari *World Bank* dalam satuan Juta USD pada tahun 2009-2016.

#### 4. Jarak

Variabel jarak dalam penelitian ini adalah jarak negara Indonesia terhadap negara asal wisman yang diukur dengan satuan

mil. Dalam hal ini jarak merupakan jarak kedua negara diukur melalui udara. Data yang digunakan dalam penelitian ini diperoleh dari situs [www.distancefromto.net](http://www.distancefromto.net).

#### 5. Nilai Tukar Negara Asal wisatawan Mancanegara

Variabel ini diambil dari nilai tukar pada masing-masing negara asal wisatawan dalam USD. Data diperoleh dari *World Bank* pada tahun 2009-2016.

#### 6. Populasi Negara Asal Wisatawan Mancanegara

Variabel populasi adalah jumlah populasi pada masing-masing negara asal wisatawan yang dihitung dengan satuan ribu jiwa. Data ini diperoleh dari *World Bank* tahun 2009-2016.

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

Data panel merupakan gabungan data antara data silang (*cross section*) dan runtut waktu (*time series*). data silang merupakan data yang terdiri dari banyak objek yang dikategorikan dalam beberapa jenis seperti tingkat infasi dalam suatu periode waktu tertentu. Data urut waktu meliputi beberapa periode baik harian, bulanan, kuartalan, dan tahunan. Penggunaan data panel dalam sebuah penelitian mempunyai beberapa keuntungan seperti mampu menyediakan data yang lebih banyak sehingga lebih menghasilkan angka derajat kebebasan (*degree of freedom*) yang lebih besar sehingga data panel dapat mengatasi masalah *omitted variable* atau penghilangan variabel (Widarjono, 2009).

Ada beberapa keuntungan dalam menggunakan data panel yang dikemukakan oleh Wibisono (2005):

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

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

$$\log Y_{it} = a + b_1 \log PDB_{wisman_{it}} + b_2 \log PDB_{bindo_{2it}} + b_3 \log jarak_{3it} + b_4 \log pop_{wisman_{it}} + b_5 \log nilaitukar_{wisman_{5it}} + e_{it} \quad (3.1)$$

Keterangan:

Y : Total Turis Mancanegara (Ribu Jiwa)

$\alpha$  : Konstanta

PDBwisman : Produk Domestik Bruto Negara Indonesia (Juta USD)

PDBindo : Produk Domestik Bruto Negara Asal Wisman (Juta USD)

jarak : Jarak (Km)

popwisman : Populasi Wisman (Jiwa)

nilaitukar : Nilai Tukar Wisman (US Dollar)

e : *Error Term*

t : Waktu

i : Negara

## G. Model Estimasi

Terdapat tiga tahap dalam metode estimasi dengan data panel, antara lain:

### 1. Uji Asumsi Klasik

#### a. Multikolinearitas

Tujuan uji multikolinearitas adalah untuk menguji apakah di dalam model regresi terdapat hubungan antar variabel independen. Jika terjadi multikolinearitas di dalam model, estimator masih bersifat *Best Linear*

*Unbiased Estimator (BLUE)* namun estimator memiliki varian dan kovarian yang besar sehingga sulit didapatkan estimasi yang tepat (Widarjono, 2013).

Skenario statistik multikolinearitas adalah terdapat hubungan yang sempurna antara variabel penjelas dan saling bergerak satu sama lain. Praktiknya, sulit untuk menghasilkan perkiraan yang dapat diandalkan dari masing-masing koefisien individu serta melihat besarnya kesalahan prediksi. Hal ini akan berakibat salahnya kesimpulan hubungan antar variabel. Kurangnya signifikansi variabel penjelas meski model yang digunakan benar adalah disebabkan karena multikolinearitas dapat meningkatkan varian parameter perkiraan. Dalam multikolinearitas jika VIF melebihi 5 atau 10 maka hasil regresi mengandung multikolinearitas (Montgomery, 2012).

#### b. Heterokedastisitas

Tujuan uji Heterokedastisitas adalah untuk menguji apakah dalam model regresi memiliki varian variabel pengganggu yang tidak konstan atau heterokedastisitas. Model regresi yang baik tidak terjadi heterokedastisitas atau mengandung homokedastisitas. Heterokedastisitas ini lebih sering terjadi pada data *cross section*. Varian variabel pengganggu yang tidak konstan atau heterokedastisitas disebabkan oleh residual pada variabel independen dalam model. Ada pun bentuk fungsi variabel pengganggu menurut Widarjono (2013) adalah sebagai berikut.

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

Di mana  $e=2,718$

Uji heterokedastisitas juga dapat dilakukan menggunakan *uji White* dengan meregresi residual kuadrat yang prosedurnya sebagai berikut:

$H_0$  : Tidak ada heterokedastisitas

$H_1$  : Ada heterokedastisitas

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

## 2. Pemilihan Model

### a. *Common Effect*

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

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

Di mana:

$i = 8$  negara asal wisatawan

$t = 2009$  hingga  $2016$

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 antar individu mempunyai efek berbeda yang bisa diakomodasikan melalui intersepnya. Dalam model ini, setiap parameter yang tidak diketahui akan diestimasi dengan teknik variabel dummy yang dinamakan *Least Square Dummy Variable (LSDV)*. LSDV mampu mengakomodasi 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 model ECM ini dapat dituliskan sebagai berikut:

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

$i$  : 8 negara asal wisatawan mancanegara

$t$  : tahun 2009 hingga 2016

Di mana:

$$W_{it} = \varepsilon_{1it} + \mu_1; E(W_{it}) = 0; E(W_{it}^2) = a^2 + a_{\mu_2j}; \quad (3.5)$$

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

$$E(\epsilon_i \epsilon_{is}) = E(\epsilon_{it} \epsilon_{jt}) = E(\epsilon_{it} \epsilon_{js}) = 0 \quad (3.7)$$

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

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

Maka dari itu, metode OLS tidak dapat digunakan untuk mendapatkan estimator yang efisien bagi model *Random Effect*. Metode yang tepat untuk mengestimasi model *random effect* yaitu *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 adalah sebagai berikut:

$H_0$  : *Common Effect Model* atau *Pooled OLS*

$H_1$  : *Fixed Effect Model*

Apabila nilai probabilitas F-statistiknya di bawah 5% maka  $H_0$  ditolak dan  $H_1$  diterima yang artinya model yang tepat adalah *fixed effect*. Sedangkan, apabila F-statistiknya signifikan di atas  $\alpha$  5% maka  $H_1$  ditolak dan  $H_0$  diterima yang artinya model yang tepat digunakan adalah *common effect*. Perhitungan dalam F-statistik didapatkan melalui uji chow dengan rumus dibawah (Baltagi 2008).

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

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 :

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

Di mana:

$\alpha$  : 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 bertujuan untuk membandingkan antara metode *fixed effect* dan metode *random effect*. Uji hausman ini berdasarkan ide bahwa *Least Squares Dummy Variables* (LSDV) pada metode *fixed effect* dan *Generalized Least Squares* (GLS) pada metode *random effect* efisien sedangkan *Ordinary Least Squares* pada metode *common effect* tidak efisien. Dengan dasar hipotesis:

$H_0$  : *Random Effect Model*

$H_1$  : *Fixed Effect Model*

Secara statistik dirumuskan dengan :

$$H = (\beta_{REM} - \beta_{FEM})'(M_{REM})^{-1}(\beta_{REM} - \beta_{FEM}) \sim X^2(k) \quad (3.11)$$

Di mana :

M : Matriks kovarian untuk parameter  $\beta$

k : *Degrees of freedom*

Apabila nilai probabilitasnya lebih kecil dari 5% maka  $H_0$  ditolak. Begitupun sebaliknya, apabila nilai probabilitas menunjukkan angka diatas 5% maka tidak menolak  $H_0$ .

f. Uji lagrange Multiplier (LM)

Uji Lagrange Multiplier bertujuan untuk membandingkan model estimasi *common effect* dan *random effect* (Widarjono, 2007). Uji signifikansi ini dikembangkan oleh *Breusch-Pagan*. Apabila LM hitung lebih besar dari nilai kritis *Chi-Squares* maka hipotesis nul ditolak yang artinya model yang tepat digunakan untuk regresi adalah model *random effect*. Dan sebaliknya jika LM hitung lebih kecil dari nilai kritis *Chi-Squares* maka hipotesis nul ditolak yang artinya model yang tepat digunakan untuk regresi adalah model *common effect*, atau dapat dibuat hipotesis sebagai berikut:

$H_0$  : *Common Effect Model*, jika nilai *p-value* > taraf nyata ( $\alpha$  5%)

$H_1$  : *Random effect Model*, jika nilai *p-value* > taraf nyata ( $\alpha$  5%)

Apabila nilai signifikansi lebih dari 0,05 persen maka *common effect* yang dipilih. Sebaliknya jika kurang dari 0,05 persen maka *random effect* yang dipilih.

### 3. Uji Signifikansi

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

#### a. Uji t

Uji t digunakan untuk menguji bagaimana pengaruh masing-masing variabel independen secara parsial terhadap variabel dependen. Uji ini dapat dilakukan dengan membandingkan probabilitas t-hitung dengan tingkat signifikansi. Jika probabilitas t-hitung lebih dari tingkat signifikansi, maka  $H_0$  ditolak yang berarti variabel independen berpengaruh terhadap variabel dependen.

#### b. Uji F

Uji F dilakukan untuk melihat pengaruh variabel independen secara bersama-sama terhadap variabel dependen dengan cara membandingkan F-hitung dengan tingkat signifikansi. Apabila probabilitas F-hitung lebih tinggi dari signifikansi maka  $H_0$  ditolak.

#### c. Koefisien Determinasi ( $R^2$ )

Koefisien determinasi dilakukan untuk mengukur seberapa jauh kemampuan model dalam menjelaskan variasi dalam variabel dependen. Dikarenakan dalam penelitian ini menggunakan lebih dari satu variabel independen. Nilai koefisien determinasi berkisar antara nol dan satu. Apabila nilai ( $R^2$ ) nya mendekati nol berarti variasi dependennya sangat terbatas. Apabila nilainya mendekati

satu berarti variabel independenya dapat menjelaskan segala variasi dari variabel independen.