

BAB III

METODOLOGI PENELITIAN

A. Objek Penelitian

Penelitian ini mencakup 26 negara utama tujuan migran Indonesia, yakni Malaysia, Singapura, Brunei Darussalam, Hongkong, Korea Selatan, Jepang, Macao, Australia, Arab Saudi, Uni Emirat Arab, Kuwait, Bahrain, Qatar, Oman, Yordania, Mesir, Cyprus, Sudan, Afrika Selatan, Amerika, Belanda, Italia, Jerman, Inggris, Perancis, dan Spanyol.

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 beberapa sumber kredibel, seperti Bank Indonesia, *World Bank*, *ILOSTAT*, *OECD*, *distancefromto.net* untuk data berupa jumlah remitansi, PDB, jumlah Tenaga Kerja Indonesia (TKI), jarak antar negara, populasi, indeks stabilitas politik, dan produktivitas tenaga kerja.

Penelitian ini mengambil tahun dasar 2006 dengan alasan dibentuknya Peraturan Presiden Republik Indonesia Nomor 81 mengenai urgensi pembentukan Badan Nasional Penempatan dan Perlindungan Tenaga Kerja Indonesia (BNP2TKI) dan tahun 2015 sebagai periode akhir penelitian karena merupakan data terkini yang tersedia di Bank Indonesia. Dengan demikian, penelitian ini menggunakan data panel

dengan basis data tahunan dari 26 negara tujuan migran Indonesia pada tahun 2006-2015.

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*, *International Labor Organization (ILOSTAT)*, *Organization for Economic Cooperation and Development (OECD)*, 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 jumlah remitansi yang diperoleh Indonesia, PDB negara tujuan, PDB negara asal, jumlah Tenaga Kerja Indonesia (TKI) di negara tujuan, jarak antar Indonesia dengan negara tujuan, populasi di negara tujuan, Indeks Stabilitas Politik negara tujuan, dan produktivitas tenaga kerja negara tujuan.

E. Definisi Operasional Variabel Penelitian

Penelitian ini terdiri dari satu variabel dependen dan tujuh 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. Remitansi

Variabel remitansi pada penelitian ini dipilih sebagai variabel dependen. Remitansi adalah pengiriman uang oleh tenaga kerja Indonesia di negara tujuan migran. Menurut BNP2TKI, remitansi dihitung berdasarkan jumlah migran yang bekerja di negara tujuan, rata-rata upah di negara tujuan, dan survei pola pengiriman uang migran yang dinyatakan dalam satuan juta USD. Data diambil dari Bank Indonesia tahun 2006-2015.

2. PDB

Variabel PDB pada penelitian ini berupa nilai akhir barang dan jasa yang diproduksi oleh negara tujuan migran 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 2006-2015.

3. Jarak

Variabel jarak pada penelitian ini adalah jarak Indonesia dengan negara tujuan migran 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.

4. Jumlah Tenaga Kerja Indonesia (TKI)

Variabel jumlah TKI pada penelitian ini berupa jumlah TKI yang bekerja di negara tujuan migran. Data diperoleh dari Bank Indonesia dengan satuan ribu jiwa dari tahun 2006-2015.

5. Populasi

Variabel populasi dalam penelitian ini adalah jumlah populasi pada masing-masing negara tujuan migran yang dihitung dengan satuan ribu jiwa. Data diperoleh dari *World Bank* dari tahun 2006-2015.

6. Indeks Stabilitas Politik

Yang dimaksud dengan Indeks Stabilitas Politik pada penelitian ini adalah kemungkinan ketidakstabilan politik dan kejahatan bermotif politik dengan agregat nilai berkisar -2,5 sampai 2,5. Semakin besar indeksnya maka semakin stabil keadaan politik negara tersebut. Data ketidakstabilan politik diperoleh dari *World Bank* dengan satuan poin dari tahun 2006-2015.

7. Produktivitas Tenaga Kerja

Produktivitas tenaga kerja pada penelitian ini adalah total volume output yang diukur dari PDB yang diproduksi per unit tenaga kerja berdasarkan jumlah orang yang dipekerjakan selama kurun waktu tertentu. Data didapatkan dari ILOSTAT dalam satuan USD dari tahun 2006-2015.

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:

$$\log Y_{it} = a + b_1 \log X_{1it} + b_2 \log X_{2it} + b_3 \log X_{3it} + b_4 \log X_{4it} + b_5 \log X_{5it} + b_6 X_{6it} + b_7 \log X_{7it} e_{it} \quad (3.1)$$

Keterangan:

Y : Remitansi (Juta USD)

α : Konstanta

X₁ : Produk Domestik Bruto Negara Tujuan (Juta USD)

X₂ : Produk Domestik Bruto Negara Asal (Juta USD)

X₃ : Jarak (Km)

X₄ : Jumlah TKI (Ribuan Jiwa)

X₅ : Populasi (Ribuan Jiwa)

X₆ : Indeks Stabilitas Politik (Poin)

X₇ : Produktivitas Tenaga Kerja (USD)

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:

H_0 : Tidak ada heteroskedastisitas

H_1 : 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*:

$$Y_{it} = \alpha + X_{it}\hat{\alpha} + \hat{\alpha}_{it} \quad (3.3)$$

Di mana:

$i = 26$ negara tujuan migran 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:

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

i : 26 negara tujuan migran Indonesia

t : Tahun 2006 sampai 2015

Di mana:

$$W_{it} = \hat{a}_{1it} + \hat{a}_{2it}; E(W_{it}) = 0; E(W_{it}^2) = \sigma^2 + \sigma_{\hat{a}_{1it}}^2 + \sigma_{\hat{a}_{2it}}^2; \quad (3.5)$$

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

$$E(\hat{a}_{1it} \hat{a}_{1is}) = E(\hat{a}_{1it} \hat{a}_{1jt}) = E(\hat{a}_{1it} \hat{a}_{1js}) = 0 \quad (3.7)$$

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

$$\text{cross}(W_{it}, W_{i,(t-1)}) = \frac{a_1^2}{a^2 + a_1^2} \quad (3.8)$$

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):

$$F = \frac{\frac{(SSE_1 - SSE_2)}{(n-1)}}{\frac{SSE_2}{(nt-n-k)}} \quad (3.9)$$

Di mana :

SSE_1 : *Sum Square Error* dari model *Common Effect*

SSE_2 : *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. 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*,

(Mundlak, 1978). Pendekatan *Mundlak* dapat dituliskan dalam persamaan sebagai berikut:

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

Indeks i menunjukkan individu atau negara sedangkan t menunjukkan waktu. Y_{it} adalah variabel dependen dan X_{it} merupakan variabel independen, ϵ_{it} adalah variabel yang tidak mengandung *time-invariant*, dan α_i adalah variabel yang mengandung *time-invariant*. Tujuan pendekatan *Mundlak* ini adalah untuk melihat hubungan antara α_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 α_i dan variabel kovarian tidak berkorelasi. Persamaannya dapat ditulis sebagai berikut:

$$Y_{it} = X_{it} \hat{\alpha} + \alpha_i + \epsilon_{it} \quad (3.14)$$

$$Y_{it} = X_{it} \hat{\alpha} + \bar{X}_i \theta + V_i + \epsilon_{it} \quad (3.15)$$

$$E(Y_{it} | X_{it}) = X_{it} \hat{\alpha} + \bar{X}_i \theta \quad (3.16)$$

Persamaan 3.15 mengganti α_i dengan $X_i\beta + V_i$. Persamaan ketiga bergantung pada fakta bahwa variabel dependen dan variabel *time-invariant* bersifat independen, sehingga $H_0: \beta = 0$.

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 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.