

BAB III

METODOLOGI PENELITIAN

A. Objek Penelitian

Penelitian ini mencakup 26 negara utama tujuan migran Indonesia, yaitu 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

Penelitian ini menggunakan metode pendekatan kuantitatif. Data dalam penelitian ini diperoleh dari beberapa sumber terpercaya, seperti Bank Indonesia, *World Bank*, *oic-oci.org*, *distancefromto.net* untuk data berupa Tenaga Kerja Indonesia (TKI), jarak antar negara, pendapatan per kapita, jumlah populasi negara tujuan, dan agama mayoritas di negara tujuan.

Penelitian ini mengambil tahun 2011 dengan alasan bahwa pada tahun 2010 Presiden Republik Indonesia mengesahkan undang-undang mengenai pelaksanaan penempatan dan perlindungan tenaga kerja Indonesia di luar negeri dan undang-undang tersebut efektif diberlakukan pada tahun 2011. Kemudian, 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 dari 26 negara tujuan migran Indonesia pada tahun 2011-2015.

C. Teknik Pengambilan Data

Data dalam penelitian ini didapatkan dari berbagai basis data dan laporan statistik terpercaya, yaitu situs resmi Bank Indonesia, *World Bank*, *oic-oci.org*, 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 tenaga kerja Indonesia (TKI), jarak antara Indonesia dengan negara tujuan, pendapatan per kapita, jumlah populasi negara tujuan, dan agama mayoritas di negara tujuan.

E. Definisi Operasional Variabel Penelitian

Penelitian ini terdiri dari satu variabel dependen dan lima 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. Tenaga Kerja Indonesia (TKI)

Variabel TKI pada penelitian ini dipilih sebagai variabel dependen. TKI atau tenaga kerja Indonesia adalah setiap warga negara Indonesia yang memenuhi syarat untuk bekerja di luar negeri dalam hubungan kerja untuk jangka waktu tertentu dengan menerima upah. TKI yang dimaksud dalam penelitian ini mencakup seluruh TKI yang bekerja di luar negeri, baik laki-laki ataupun wanita. Data diambil dari Bank Indonesia tahun 2006-2015. Data TKI dalam penelitian ini diukur dalam satuan ribu jiwa.

2. 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*.

3. Pendapatan per kapita negara asal

Variabel pendapatan per kapita negara asal pada penelitian ini berupa besarnya pendapatan penduduk di Indonesia. Pendapatan per kapita didapatkan dari hasil pembagian pendapatan nasional suatu negara dengan

jumlah penduduk negara tersebut. Data pendapatan per kapita yang digunakan berupa data riil dengan harga konstan pada tahun 2010 yang dikonversi dari mata uang domestik ke dalam USD menggunakan kurs resmi untuk setiap tahunnya. Data pendapatan per kapita negara asal diperoleh dari *World Bank* dalam satuan juta USD pada tahun 2011-2015.

4. Pendapatan Per kapita negara tujuan

Variabel pendapatan per kapita negara tujuan pada penelitian ini berupa besarnya pendapatan penduduk setiap negara tujuan. Pendapatan per kapita didapatkan dari hasil pembagian pendapatan nasional suatu negara dengan jumlah penduduk negara tersebut. Data pendapatan per kapita yang digunakan berupa data riil dengan harga konstan pada tahun 2010 yang dikonversi dari mata uang domestik ke dalam USD menggunakan kurs resmi untuk setiap tahunnya. Data pendapatan per kapita negara tujuan diperoleh dari *World Bank* dalam satuan juta USD pada tahun 2011-2015.

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 2011-2015.

6. Agama

Variabel agama pada penelitian ini adalah agama mayoritas pemeluk Islam di negara asal dan negara tujuan migrasi. Dalam penelitian ini, variabel agama dijadikan sebagai variabel *dummy* yaitu di mana negara dengan mayoritas pemeluk agama islam dihitung dengan angka 1 kemudian negara dengan mayoritas pemeluk agama non islam dihitung dengan angka 0. Data yang diperoleh mengacu pada situs www.oic-oci.org

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 ke dalam beberapa jenis data, misalnya laba, biaya iklan, laba ditahan, dan tingkat investasi dalam suatu periode waktu tertentu.

Ada beberapa keuntungan yang diperoleh dengan menggunakan data panel. Pertama, data panel merupakan gabungan *time series* dan *cross section* yang mampu menyediakan data yang lebih banyak sehingga akan menghasilkan *degree of freedom* yang lebih besar. Kedua, menggabungkan informasi dari data *time series* dan *cross section* dapat mengatasi masalah yang timbul ketika ada masalah penghilangan variabel (*omitted-variable*).

Ada pun, 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. Digunakan untuk meminimalisir bias yang mungkin ditimbulkan oleh agregasi data individu.

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

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

Keterangan:

Y : Tenaga Kerja Indonesia (Ribuan jiwa)

α : Konstanta

X₁ : Jarak (Km)

X₂ : PDB per kapita negara asal (Juta USD)

X_3 : PDB per kapita negara tujuan (Juta USD)

X_4 : Populasi negara tujuan (Ribu Jiwa)

X_5 : Kesamaan mayoritas agama negara asal dan negara tujuan
(*variabel boneka*)

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 dalam model regresi ditemukan adanya hubungan antar variabel bebas atau independen. Model regresi yang baik seharusnya tidak terjadi korelasi di antara 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*. 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 residual kuadrat yang prosedurnya sebagai berikut:

H_0 : Tidak terdapat heteroskedastisitas

H_1 : Terdapat heteroskedastitas

Jika nilai signifikansi lebih besar dari derajat kepercayaan 0,05, maka dapat dikatakan tidak terdapat 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 dalam penelitian ini, penulis menggunakan data sekunder berbasis data panel dengan kurun waktu hanya lima 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*:

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

Di mana:

$i = 26$ negara tujuan migran Indonesia

$t = 2011$ 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 : 26 negara tujuan migran Indonesia

t : Tahun 2011 sampai 2015

Di mana:

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

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

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

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

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

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 :

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. Uji 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:

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

Indeks i menunjukkan individu atau negara sedangkan t menunjukkan waktu. Y_{it} adalah variabel dependen dan X_{it} merupakan variabel independen, 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}

$$\begin{aligned} \alpha_i &= \bar{X}_i \theta + V_i \\ E(\alpha_i | X_i) &= \bar{X}_i \theta \end{aligned} \quad (3.12)$$

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:

$$\begin{aligned}
Y_{it} &= X_{it}\beta + \alpha_i + \varepsilon_{it} \\
Y_{it} &= X_{it}\beta + \bar{X}_i\theta + V_{it} + \varepsilon_{it} \\
E(Y_{it}|X_{it}) &= X_{it}\beta + \bar{X}_i\theta
\end{aligned} \tag{3.13}$$

Persamaan kedua mengganti α_i dengan $X_i\theta + V_i$. Persamaan ketiga bergantung pada fakta bahwa variabel dependen dan variabel *time-invariant* 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 probabilitas t-hitung dengan tingkat signifikansi. Apabila probabilitas t-hitung lebih dari tingkat signifikansi, maka H_0 ditolak yang artinya variabel independen negara berpengaruh terhadap variabel dependen.

b. Uji F

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

c. Koefisien Determinasi

Koefisien determinasi dilakukan untuk mengukur seberapa jauh kemampuan 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.