

BAB III

METODOLOGI PENELITIAN

A. Objek Penelitian

Penelitian ini mencakup empat negara ASEAN yang memiliki tingkat ekspor paling tinggi ke Indonesia atau Indonesia memiliki tingkat impor tertinggi dari keempat negara ASEAN tersebut, yaitu Malaysia, Singapore, Thailand, dan Vietnam.

B. Jenis Data

Pada penelitian ini, penulis menggunakan data sekunder. Data sekunder adalah data yang dikumpulkan dari berbagai sumber yang telah tersedia sehingga penulis hanya menggunakan data tersebut. Data sekunder didapat dari berbagai sumber kredible, seperti UN Comtrade, *distancefromto.net*, World Bank untuk data berupa PDB negara asal, PDB negara mitra dagang Indonesia, jarak Indonesia dengan negara mitra dagang, jumlah impor, serta populasi Indonesia.

C. Teknik Pengambilan Data

Data dalam penelitian ini didapat dari berbagai basis data dan laporan statistik terpercaya, yakni melalui situs *UN Comtrade*, *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 total impor Indonesia, PDB negara asal, PDB negara mitra dagang Indonesia, jarak Indonesia dengan negara mitra dagang dan populasi Indonesia.

E. Definisi Operasional Variabel Penelitian

Penelitian ini terdiri dari satu variabel dependen dan tiga variabel independen. Variabel independen merupakan variabel bebas atau bisa disebut juga variabel yang mempengaruhi, sedangkan variabel dependen atau variabel terikat merupakan variabel yang dipengaruhi oleh variabel independen. Ada pun, variabel dependen dan independen yang digunakan dalam penelitian ini meliputi:

1. Impor

Variabel impor pada penelitian ini dipilih sebagai variabel dependen. Impor adalah barang atau jasa yang dihasilkan di negara lain yang masuk ke suatu negara. Impor suatu negara ditentukan oleh beberapa faktor diantaranya daya saing negara. Data dihitung dalam satuan USD pada tahun 2006-2017 yang diperoleh dari *UN Comtrade*.

2. Jarak

Variabel jarak pada penelitian ini adalah jarak Indonesia dengan mitra dagang yang dihitung dengan satuan mil. jarak dalam hal ini merupakan jarak ibukota kedua negara yang diukur melalui udara. Data yang diperoleh mengacu pada situs www.distancefromto.net.

3. PDB Negara Tujuan Impor

Variabel PDB pada penelitian ini berupa nilai akhir barang dan jasa yang diproduksi Indonesia dalam periode satu tahun. Data PDB pada penelitian ini dihitung dalam satuan USD. Data PDB diperoleh dari *World Bank* dalam satuan USD pada tahun 2006-2017.

4. PDB Negara Asal Impor

Variabel PDB pada penelitian ini berupa nilai akhir barang dan jasa yang diproduksi oleh negara mitra dagang selama setahun. Data PDB pada penelitian ini dihitung dalam satuan USD. Data PDB diperoleh dari *World Bank* dalam satuan USD pada tahun 2006-2017.

5. Populasi Negara Tujuan Impor

Jumlah populasi yang dimaksud dalam penelitian ini mencakup total penduduk yang tinggal dalam negara pengimpor dalam hal ini Indonesia. Data populasi dalam penelitian ini diperoleh dari *World Bank* dengan satuan juta jiwa dari tahun 2006-2017.

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. Widarjono (2009) menjelaskan 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*).

Ada pun, data panel memiliki beberapa keuntungan sebagai berikut (Wibisono, 2005):

1. Data panel dapat memperhitungkan heterogenitas individu secara eksplisit dengan mengiinkan variabel spesifik individu.
2. Kemampuan dalam mengontrol heterogenitas ini selanjutnya menjadikan data panel dapat digunakan untuk menguji dan membangun model perilaku yang 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 akan 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.
6. Mempelajari model perilaku kompleks.

Adapun model regresi panel dalam penelitian ini sebagai berikut:

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

Keterangan:

- Y : Impor Total (USD)
 α : Konstanta
X₁ : Jarak (Km)
X₂ : PDB (Juta USD)
X₃ : PDB (Juta USD)
X₄ : Populasi (Ribuan Jiwa)
e : *Error term*
t : Waktu
i : Negara

G. Model Estimasi

Model estimasi dengan data panel dapat dilakukan dengan tiga tahap, antara lain:

1. *The Poisson Pseudo-Maximum Likelihood Estimator*

The Poisson Pseudo-Maximum Likelihood Estimator (PPML) dalam konteks model gravitasi telah menjadi populer dalam perdagangan internasional dan studi literatur migrasi. Diperkenalkan oleh Gourieroux, Monfort dan Trognon pada tahun 1984. Kemudian Santos Silva dan Tenreyro pada 2006 menunjukkan bahwa PPML adalah metode yang sederhana namun kuat untuk memperkirakan parameter hambatan bilateral dari model gravitasi. Dua penelitian di atas menjadi salah satu alat standar dalam literatur ekonomi internasional, banyak digunakan untuk menjelaskan perdagangan bahkan belakangan ini digunakan untuk arus perdagangan.

Dua penelitian di atas dijadikan acuan untuk diaplikasikan dalam penelitian ini karena penelitian ini membahas tentang perdagangan internasional khususnya impor dengan pendekatan model gravitasi maka estimasi yang paling cocok digunakan adalah PPML. Dalam penelitian Santos dan Tenreyro (2006) menyatakan bahwa estimasi OLS dapat mengakibatkan bias karena dalam kerjasama perdagangan antar negara dapat memiliki volume nol atau biasa disebut dengan (*Zero Trade Volume*). Ketika terjadi kasus *Zero Trade Volume* atau aliran perdagangan nol maka data akan dikeluarkan (*Zero trade is dropped*). Untuk mengatasi

permasalahan bias akibat, heteroskedastisitas, dan hilangnya nilai nol, estimasi yang dapat dilakukan dengan teknik PPML (*Poisson Pseudo Maximum Likelihood*) dengan simulasi Monte Carlo (Sattayanuwat, 2014). Pertimbangan bentuk nonlinier dari model gravitasi Anderson dan Van Wincoop dengan istilah *error term*:

$$X_{ij}^k = \frac{Y_i^k E_j^k}{Y^k} \left(\frac{\tau_{ij}^k}{\pi_i^k P_j^k} \right) (1 - \sigma_k) e_{ij}^k \quad (3.2)$$

Mengambil logaritma memberikan model gravitasi standar dalam bentuk linier, akan tetapi juga memperjelas *error term*:

$$\log X_{ij}^k = \log Y_i^k + \log E_j^k - Y^k + (1 - \sigma_k) [\log \tau_j^k - \log \pi_i^k - \log P_j^k] + \log e_{ij}^k \quad (3.3)$$

Nilai tengah tergantung pada momen yang lebih tinggi, sehingga termasuk variannya. Jika heteroskedastik, yang sangat mungkin dalam praktiknya maka nilai yang diharapkan dari istilah kesalahan tergantung pada satu atau lebih dari variabel penjelas karena itu termasuk istilah varians. Ini melanggar asumsi pertama OLS dan menunjukkan bahwa estimator mungkin bias dan tidak konsisten. Penting untuk dicatat bahwa jenis heteroskedastisitas ini tidak dapat mengatasi dengan hanya menerapkan penaksir matriks kovarian yang kuat, karena hal tersebut mempengaruhi estimasi parameter selain kesalahan standar.

2. *Maximum Likelihood Estimation*

Maximum Likelihood Estimation (MLE) adalah teknik yang digunakan untuk mencari titik tertentu untuk memaksimalkan sebuah

fungsi, teknik ini sangat luas dipakai dalam penaksiran suatu parameter distribusi data dan tetap dominan dipakai dalam pengembangan uji-uji yang baru. Misalkan saja terdapat populasi dengan fungsi probabilitas $P(Y_i, \theta)$ memiliki satu parameter maka apabila ditunjukkan hasil observasi Y_1, \dots, Y_n fungsi joint probabilitas dari sampel :

$$g(Y_1, \dots, Y_n) = \prod_{i=1}^n P(Y_i, \theta) \quad (3.4)$$

Apabila fungsi diatas dianggap fungsi dengan parameter theta (θ), maka selanjutnya akan disebut fungsi *Likelihood* :

$$L(\theta) = \prod_{i=1}^n P(Y_i, \theta) \quad (3.5)$$

Apabila fungsi *Likelihood* dengan parameter θ diatas bisa dimaksimalkan akan diperoleh penduga θ yang optimal. Dengan tujuan memaksimalkan fungsi *likelihood* ini bisa dilakukan dengan beberapa jalan yaitu diferensial biasa, bisa juga dengan diferensial *Newton-Raphson* (Neter, 1996).

$$L_c(\beta, \phi^2) = constant - \frac{NT}{2} \log\{d^{[Q+\phi^2(P-J_{NT})]}d\} + \frac{N}{2} \log\phi^2 \quad (3.6)$$

3. Uji Asumsi Klasik

Penelitian ini tidak menggunakan uji asumsi klasik. Jika terdapat multikolinieritas, Stata akan menghilangkan satu atau lebih variabel untuk menghilangkannya. Apabila ingin melihat lebih detail dapat dengan melakukan –estat vif-. Jika model Anda

memberi Anda kesalahan standar besar yang tidak masuk akal untuk beberapa variabel yang memiliki VIF tinggi, sebenarnya tidak ada yang bisa dilakukan dengan cara apa pun kecuali ada cara untuk mendapatkan lebih banyak data.

4. 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.

