

BAB III

METODOLOGI PENELITIAN

A. Tujuan Penelitian

Berdasarkan masalah-masalah yang telah peneliti rumuskan, maka tujuan penelitian ini adalah untuk mengetahui:

1. Apakah investasi mempengaruhi kesempatan kerja pada sektor Industri alat angkut di Indonesia?
2. Apakah tingkat pendidikan dasar mempengaruhi kesempatan kerja pada Industri alat angkut di Indonesia?
3. Apakah tingkat pendidikan menengah mempengaruhi kesempatan kerja pada Industri alat angkut di Indonesia?
4. Apakah tingkat pendidikan tinggi mempengaruhi kesempatan kerja pada Industri alat angkut di Indonesia?
5. Apakah investasi dan tingkat pendidikan dasar menengah tinggi mempengaruhi kesempatan kerja pada sektor Industri alat angkut di Indonesia?

B. Obyek dan Ruang Lingkup Penelitian

Objek dan ruang lingkup penelitian dari penelitian ini adalah investasi, dan tingkat pendidikan terhadap kesempatan kerja di Indonesia dengan menggunakan data-data statistik dari Badan Pusat Statistik (BPS) kementerian ketenagakerjaan, kementerian Perindustrian dan Badan Koordinasi dan Penanaman Modal.

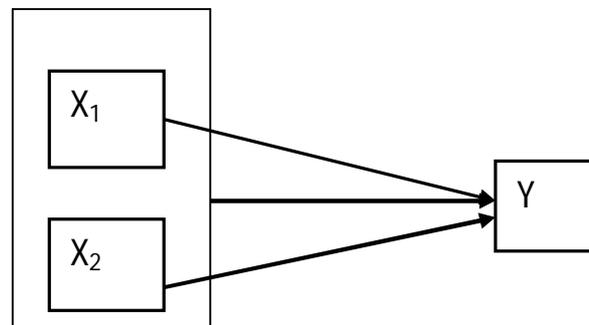
Penelitian ini dilakukan pada bulan Januari-Juni 2015 karena merupakan waktu yang efektif bagi peneliti untuk melaksanakan penelitian sehingga peneliti dapat fokus pada saat penelitian dan keterbatasan peneliti dalam waktu, tenaga, dan materi. Ruang lingkup penelitian ini adalah mengkaji hubungan antara investasi, dan tingkat pendidikan terhadap kesempatan kerja sektor industri alat angkut di Indonesia.

C. Metode Penelitian

Metode penelitian yang digunakan dalam penelitian ini adalah metode *Ekspos Facto* dengan pendekatan korelasional. *Ekspos Facto* adalah meneliti peristiwa yang telah terjadi dan kemudian menuntut ke belakang untuk mengetahui faktor-faktor yang menimbulkan kejadian tersebut. Metode ini dipilih karena sesuai untuk mendapatkan informasi yang bersangkutan dengan status gejala pada saat penelitian dilakukan. Pendekatan korelasional yang dilakukan adalah dengan menggunakan korelasi ganda. Korelasi ganda dipilih karena dapat menunjukkan arah pengaruh faktor-faktor penentu.

Dalam penelitian ini terdapat tiga variabel yang menjadi objek penelitian dimana kesempatan kerja di sektor industri alat angkut merupakan variabel terikat (Y). Sedangkan variabel bebas adalah investasi (X1), dan tingkat pendidikan (X2). Konstelasi pengaruh antar variabel di atas dapat digambarkan sebagai berikut:

Konstelasi hubungan antar variabel



Keterangan:

X_1 = investasi (variabel bebas)

X_2 = tingkat pendidikan (variabel bebas)

Y = kesempatan kerja di sektor industri alat angkut (variabel terikat)

→ = arah pengaruh

D. Jenis dan Sumber Data

Jenis data yang digunakan dalam penelitian ini adalah data sekunder yang bersifat kuantitatif, yaitu data yang telah tersedia dalam bentuk angka. Sedangkan data yang digunakan dalam penelitian ini adalah data runtut waktu (*time series*) dan data deret lintang (*cross section*). Data *time series* adalah data yang dikumpulkan dari waktu ke waktu terhadap suatu individu, sedangkan data *cross section* adalah data yang dikumpulkan dalam satu waktu terhadap banyak individu.²⁶ Data *time series* sebanyak lima tahun dari tahun 2010 sampai 2014 dan data *cross section* di Indonesia. Data sekunder tersebut diperoleh dari sumber-sumber seperti catatan atau laporan yang dipublikasikan oleh Badan Pusat Statistik (BPS)

²⁶ Nachrowi, *Pendekatan Populer dan Praktis Ekonometrika untuk Analisis Ekonomi dan Keuangan*, (Jakarta: LPFE UI, 2006), p. 309.

E. Operasionalisasi Variabel Penelitian

Operasionalisasi variabel penelitian ini diperlukan untuk memenuhi jenis dan indikator dari variabel-variabel yang terkait dalam penelitian ini. Selain itu, proses ini dimaksudkan untuk menentukan skala pengukuran dari masing-masing variabel sehingga pengujian hipotesis dengan alat bantu statistik dapat dilakukan secara luas.

a. Kesempatan kerja

1. Definisi Konseptual

Kesempatan Kerja yaitu jumlah tenaga kerja yang terserap pada lapangan kerja yang tersedia dalam masyarakat, dengan asumsi tidak ada yang kosong atau banyaknya orang yang bekerja pada suatu lapangan pekerjaan.

2. Definisi Operasional

Kesempatan kerja adalah jumlah tenaga kerja yang terserap. Dalam penelitian ini kesempatan kerja diukur dari jumlah tenaga kerja yang terserap pada industri alat angkut di Indonesia dari tahun 2010-2014

b. Investasi

1. Definisi Konseptual

Investasi adalah penanaman modal yang dimiliki dan biasanya berjangka waktu lama dengan harapan mendapatkan keuntungan di masa-masa yang akan datang.

2. Definisi Operasional

Data investasi diambil dari realisasi total Penanaman Modal Dalam Negeri (PMDN) dan Penanaman Modal Asing (PMA) industri alat angkut di Indonesia dari tahun 2010-2014.

c. Tingkat Pendidikan

1. Definisi Konseptual

Tingkat pendidikan adalah jenjang yang ditempuh oleh seseorang guna memperoleh pekerjaan yang lebih layak.

2. Definisi Operasional

Tingkat pendidikan ini dilihat dari banyaknya tenaga kerja yang bekerja di industri alat angkut berdasarkan tingkat pendidikan dasar menengah dan tinggi di Indonesia periode 2010-2014.

F. Teknik Analisis Data

1. Analisis Data Panel

Data yang digunakan dalam analisis ekonometrika dapat berupa data *time series*, data *cross section*, atau data panel. Data panel (*panel pooled data*) merupakan gabungan data *time series* dan data *cross section*. Dengan kata lain, data panel merupakan unit-unit individu yang sama yang diamati dalam kurun waktu tertentu. Jika kita memiliki T periode waktu ($t= 1,2,\dots,T$) dan N jumlah individu ($i= 1,2,\dots,N$), maka dengan data panel kita akan memiliki total unit observasi sebanyak NT . Jika jumlah unit waktu sama untuk setiap individu, maka data disebut *balanced panel*. Jika sebaliknya, yakni jumlah unit waktu berbeda untuk setiap individu, maka disebut

unbalanced panel. Penggunaan data panel pada dasarnya merupakan solusi akan ketidakterediaan data *time series* yang cukup panjang untuk kepentingan analisis ekonometrika.

Gujarati (2001) berdasarkan uraian dari Baltagi, keunggulan penggunaan data panel dibanding data runtun waktu dan data lintas sektor adalah:

1. Estimasi data panel dapat menunjukkan adanya heterogenitas dalam tiap unit.
2. Dengan data panel, data lebih informatif, mengurangi kolinieritas antara variabel, meningkatkan derajat kebebasan dan lebih efisien.
3. Data panel cocok digunakan untuk menggambarkan adanya dinamika perubahan.
4. Data panel dapat lebih mampu mendeteksi dan mengukur dampak.
5. Data panel bisa digunakan untuk studi dengan model yang lebih lengkap.
6. Data panel dapat meminimumkan bias yang mungkin dihasilkan regresi.

a) Estimasi Model Regresi Panel

Estimasi model regresi data panel terdapat tiga spesifikasi model yang mungkin digunakan, yakni model *common effect*, *fixed effect*, dan *random effect*.

1. Model Common Effect

Model *common effect* atau *pooled regression* merupakan model regresi data panel yang paling sederhana. Model ini pada dasarnya mengabaikan struktur panel dari data, sehingga diasumsikan bahwa

perilaku antar individu sama dalam berbagai kurun waktu atau dengan kata lain pengaruh spesifik dari masing-masing individu diabaikan atau dianggap tidak ada. Dengan demikian, akan dihasilkan sebuah persamaan regresi yang sama untuk setiap unit *cross section*. Persamaan regresi untuk model *common effect* dapat dituliskan sebagai berikut

$$Y_{it} = \alpha + \beta X_{it} + u_{it} \quad i = 1, 2, \dots, N \quad t = 1, 2, \dots, T \quad (1)$$

Keterangan:

Y : variabel dependen

α : koefisien regresi

X : variabel independen

β : estimasi parameter (koefisien)

u : *error term*

N : jumlah (individu)

T : jumlah periode waktu

Berdasarkan asumsi struktur matriks varians-kovarians residual, maka pada model *common effect* metode yang dapat digunakan, yaitu *Ordinary Least Square*(OLS), jika struktur matrik varians-kovarians residualnya diasumsikan bersifat homoskedastik dan tidak ada *cross sectional correlation*.

2. Model *Fixed Effect*

Jika model *common effect* cenderung mengabaikan struktur panel dari data dan pengaruh spesifik masing-masing individu, maka model *fixed effect* adalah sebaliknya. Pada model ini, terdapat efek spesifik individu α_i dan diasumsikan berkorelasi dengan variabel penjelas yang teramati X_{it} . Ekananda (2005) menyatakan bahwa berdasarkan asumsi

struktur matriks varians-kovarians residual, maka pada model *fixed effect* metode estimasi yang dapat digunakan, yaitu *Ordinary Least Square* (OLS/LSDV), jika struktur matriks varians-kovarians residualnya diasumsikan bersifat homoskedastik dan tidak ada *cross sectional correlation*.

3. Model *Random Effect*

Pada model *random effect*, efek spesifik dari masing-masing individu α_i diperlakukan sebagai bagian dari komponen *error* yang bersifat acak dan tidak berkorelasi dengan variabel penjelas yang teramati X_{it} . Dengan demikian, persamaan model *random effect* dapat dituliskan sebagai berikut:

$$Y_{it} = \alpha + \beta X_{it} + w_{it}; i = 1, \dots, N; t = 1, \dots, T \quad (2)$$

Dimana

$$w_{it} = \alpha + u_{it}; E(w_{it}) = 0; E(w_{it}^2) = \sigma^2 + \sigma_u^2; E(w_{it}w_{jt-1}) = 0; i \neq j$$

Meskipun komponen *error* w_{it} bersifat homoskedastik, nyatanya terdapat korelasi antara w_{it} dan w_{it-s} (*equicorrelation*). Karena itu, metode OLS tidak bisa digunakan untuk mendapatkan estimator yang efisien bagi model *random effect*. Metode yang tepat untuk mengestimasi model *random effect* adalah *Generalized Least Square* (GLS) dengan asumsi homoskedastik dan tidak ada *cross sectional correlation*.

Analisis regresi dalam penelitian dilakukan tiga kali, yaitu pengaruh investasi terhadap kesempatan kerja, pengaruh tingkat pendidikan terhadap kesempatan kerja Adapun model regresi yang diuji sebagai berikut:

$$Y = \alpha + \beta X_1 + u \quad (3)$$

$$Y = \alpha + \beta X_2 + u \quad (4)$$

Keterangan :

Y : Kesempatan kerja

X₁ : Investasi

X₂ : Tingkat Pendidikan

α : konstanta

β : parameter (koefisien)

u : *error term*

2. Penyeleksian Model Estimasi Data Panel

Untuk mengetahui model estimasi data panel terbaik, diperlukan pengujian signifikansi antar model sebagai berikut :

Tabel III.1
Pengujian Signifikansi Model Panel Terbaik

No	Pengujian Signifikansi Model	Hipotesis Pengujian	Rumus Uji	Ket
A	CE atau FE	H ₀ : CE lebih baik dari FE	Uji F	Tolak H ₀ F _{hit} > F _{tab}
		H _i : FE lebih baik dari CE		
B	CE atau RE	H ₀ : CE lebih baik dari RE	Uji LM	Tolak H ₀ LM > chi-sq _{tab}
		H _i : RE lebih baik dari CE		
C	FE atau RE	H ₀ : RE lebih baik dari FE	Uji Hausman	Tolak H ₀ chi-sq _{hit} > chi-sq _{tab}
		H _i : FE lebih baik dari RE		

Sumber: Wing W. Winarno, *Analisis Ekonometrika dan Statistika*, 2011.

Keterangan:

CE: *Common Effect*

FE: *Fixed Effect*

RE: *Random Effect*

a). Pengujian Signifikansi *Common Effect* atau *Fixed Effect*

Pengujian antara *common effect* atau *fixed effect* dapat dilakukan dengan uji statistik F untuk mengetahui apakah model *fixed effect* lebih baik dengan melihat *Residual Sum Squares* (RSS) dengan derajat bebas sebank $(n - 1)$ untuk numerator dan $(nT - n - k)$ untuk denominator.

$$F = \frac{(RSS_1 - RSS_2)/(n-1)}{RSS_2/(nT-n-k)} \quad (5)$$

b). Pengujian Signifikansi *Common Effect* atau *Random Effect*

Pengujian signifikansi *common effect* atau *random effect* melalui pengujian *Lagrange Multiplier* untuk mengetahui signifikansi dari *random effect* berdasarkan residual dari OLS (*common effect*). LM mengikuti sebaran *chi-square* dengan derajat bebas satu. Secara matematis, statistik uji untuk *LMtest* (*Lagrange Multiplier*) dapat dituliskan sebagai berikut:

$$LM = \frac{nT}{2(T-1)} \left[\frac{\sum_{t=1}^n (\sum_{i=1}^n e_{it})^2}{\sum_{t=1}^n \sum_{i=1}^n e_{it}^2} - 1 \right]^2 \quad (6)$$

c). Pengujian Signifikansi *Fixed Effect* atau *Random Effect*

Setelah menguji signifikansi antara *common effect* atau *fixed effect* serta *common effect* atau *random effect*, maka selanjutnya jika terbukti *fixed effect* dan *random effect* sama-sama lebih baik dari *common*

effect adalah melakukan pengujian signifikansi *fixed effect* atau *random effect*. Uji ini dilakukan dengan membandingkan dan untuk *subset* dari koefisien variabel-variabel yang bervariasi antar unit waktu (*time-varying variables*). Secara sistematis dengan menggunakan notasi matriks, statistik uji Hausman (H) dapat dituliskan sebagai berikut:

$$H = (\hat{\beta}_{FE} - \hat{\beta}_{RE})[\text{var}(\hat{\beta}_{FE}) - \text{var}(\hat{\beta}_{RE})]^{-1}(\hat{\beta}_{FE} - \hat{\beta}_{RE}) \quad (7)$$

Di bawah hipotesis nol, statistik uji ini mengikuti sebaran *chi-square* dengan derajat bebas M, di mana M adalah jumlah variabel penjelas yang nilainya bervariasi antar unit waktu di dalam model.

Sementara itu, Judge *et.al.* dalam Gujarati (2003) memberikan sejumlah pertimbangan terkait pilihan, apakah menggunakan model *fixed effect* (FE) atau model *random effect* (RE). Pertimbangan-pertimbangan itu adalah sebagai berikut:

1. Jika jumlah data *time series* (T) besar dan jumlah data *cross section* (N) kecil, ada kemungkinan perbedaan nilai parameter yang diestimasi dengan FE dan RE cukup kecil. Karena itu, pilihan ditentukan berdasarkan kemudahan perhitungan. Dalam hal ini adalah model FE.
2. Ketika N besar dan T kecil, estimasi kedua metode dapat berbeda secara signifikan. Pada kondisi seperti ini, pilihan ditentukan berdasarkan keyakinan apakah individu yang diobservasi merupakan sampel acak yang diambil dari populasi tertentu atau tidak. Jika

observasi bukan merupakan sampel acak, maka digunakan model FE. Jika sebaliknya, maka digunakan model RE.

3. Jika efek individu tidak teramati α_i berkorelasi dengan satu atau lebih variabel bebas, maka estimasi dengan RE bias, sedangkan estimasi dengan FE tidak bias.
4. Jika N besar T kecil, serta semua asumsi yang disyaratkan oleh model RE terpenuhi, maka estimasi dengan menggunakan RE lebih efisien dibanding estimasi dengan FE.

3. Pengujian Asumsi Klasik

Untuk membangun persamaan regresi panel yang terbaik dari kriteria ekonometrika, perlu dilakukan penyelidikan dan penanganan adanya masalah-masalah yang berkaitan dengan pelanggaran asumsi dasar. Berikut ini adalah asumsi-asumsi yang diperlukan dalam analisis regresi:

a) Normalitas

Pengujian normalitas dilakukan untuk mengetahui apakah residual berdistribusi normal atau tidak. Hal tersebut didasarkan pada asumsi bahwa faktor kesalahan (residual) didistribusikan secara normal. Salah satu metode yang dapat digunakan untuk menguji normalitas adalah *Jarque-Bera test*. Uji statistik ini dapat dihitung dengan rumus berikut:

$$JB = n \left[\frac{\mu_3^2}{6\mu_2^3} + \frac{(\mu_4 - 3)^2}{24} \right] \quad (11)$$

Keterangan:

n : jumlah sampel

μ_2 : varians

μ_3 : skewness

μ_4 : kurtosis

Jarque-Bera test mempunyai distribusi *chi square* dengan derajat bebas dua. Jika hasil *Jarque-Bera test* lebih besar dari nilai *chi square* pada $\alpha = 5\%$, maka tolak hipotesis nol yang berarti tidak berdistribusi normal. Jika hasil *Jarque-Bera test* lebih kecil dari nilai *chi square* pada $\alpha = 5\%$, maka terima hipotesis nol yang berarti *error term* berdistribusi normal.

b) Heteroskedastisitas

Beberapa asumsi dalam model regresi adalah (1) nilai residual memiliki nilai rata-rata nol, (2) residual memiliki varian yang konstan, dan (3) residual suatu observasi tidak saling berhubungan dengan residual lainnya, sehingga menghasilkan estimator yang BLUE (*Best Linier Unbiased Estimator*).

Apabila asumsi (1) tidak terpenuhi, yang terpengaruh hanya *slope* estimator dan tidak membawa konsekuensi serius dalam analisis ekonometris. Sedangkan apabila asumsi (2) dan (3) dilanggar, maka akan membawa dampak serius bagi prediksi dengan model yang dibangun.

Dalam kenyataannya, nilai residual sulit memiliki varian yang konstan. Hal ini sering terjadi pada data yang bersifat *cross section*

dibanding data *time series*. Untuk mengidentifikasi ada tidaknya masalah heteroskedastisitas, secara sistematis statistik uji yang digunakan dapat dirumuskan sebagai berikut:

$$\mathbf{LM} = \frac{\mathbf{T}}{2} \sum_{i=1}^n \left[\frac{\widehat{\sigma}_i^2}{\widehat{\sigma}^2} - 1 \right]^2 \quad (12)$$

Hasil uji LM harus dibandingkan dengan nilai *chi-square* tabel dengan derajat bebas $(n - 1)$ dengan ketentuan tolak H_0 jika nilai LM lebih besar dari *chi-square* yang berarti model yang terbentuk mengandung masalah heteroskedastisitas.

c) Uji Multikolinieritas

Multikolinieritas adalah keadaan dimana kedua variabel independen atau lebih pada model regresi terjadi hubungan linear yang sempurna atau mendekati sempurna. Model regresi yang baik mensyaratkan tidak adanya masalah multikolinieritas. Apabila koefisien korelasi lebih besar dari rule of thumb 0,9 maka tidak ada masalah multikolinieritas antar variabel independen.

4. Uji Hipotesis

Untuk mengetahui keberartian model regresi yang dihasilkan, hal pertama yang harus diperhatikan adalah melihat kesesuaian tanda dan nilai koefisien estimasi. Jika kriteria terpenuhi, maka selanjutnya keberartian

model regresi yang diperoleh dapat dinilai dengan memperhatikan hal-hal berikut:

a). Uji Keberartian Koefisien Regresi (Uji t)

Uji t digunakan untuk mengetahui apakah variabel bebas secara parsial berpengaruh signifikan terhadap variabel tak bebasnya. Hipotesis pengujian: $H_0: \beta_i = 0, H_1: \beta_i \neq 0$.

Statistik uji yang digunakan adalah statistik uji t -student. Adapun rumusnya adalah sebagai berikut:

$$t_{hitung} = \frac{\hat{\beta}_i}{se(\hat{\beta}_i)} \quad (8)$$

Hipotesis nol ditolak jika $t_{hitung} > t_{\alpha/2; (nT-n-k-1)}$. Keputusan ini dapat juga didasarkan pada perbandingan nilai p -value dengan tingkat signifikansinya (α). Hipotesis nol ditolak jika nilai p -value lebih kecil dari (α). Hal ini berarti secara parsial variabel bebasnya dengan tingkat kepercayaan sebesar $(1 - \alpha) \times 100$ persen.

b). Uji keberartian Regresi (Uji F)

Uji F digunakan untuk menguji apakah variabel independen secara bersama-sama berpengaruh secara signifikan terhadap variabel dependen. Metode yang digunakan dalam uji ini adalah dengan cara membandingkan antara F_{hitung} dengan F_{tabel} atau $F_{(\alpha; n+k-1; nT-n-k)}$ pada tingkat kesalahan 5% dengan hipotesis:

$$H_0: \beta_1 + \beta_2 = 0$$

$$H_1: \beta_1 + \beta_2 \neq 0$$

Hipotesis nol ditolak jika $F_{hitung} > F_{tabel}$, maka seluruh variabel independen berpengaruh signifikan terhadap variabel dependen secara simultan dan sebaliknya. Untuk menguji kedua hipotesis ini digunakan nilai statistik F yang dihitung dengan rumus sebagai berikut:

$$F = \frac{R^2 / (k - 1)}{(1 - R^2) / (n - k)}$$

Keterangan:

R^2 : koefisien determinasi

k : jumlah variabel bebas

n : jumlah data

c). Perhitungan Koefisien Determinasi (R^2)

R^2 digunakan untuk mengukur kebaikan atau kesesuaian suatu model persamaan regresi. Besaran R^2 dihitung dengan rumus:

$$R^2 = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2} = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS} \quad (9)$$

Sedangkan $R^2_{adjusted}$ dihitung dengan rumus:

$$\bar{R} = 1 - (1 - R^2) \frac{nT - 1}{nT - n - k} \quad (10)$$

Keterangan:

ESS : jumlah kuadrat yang dijelaskan

RSS : jumlah kuadrat residual

TSS : jumlah kuadrat total

n : jumlah observasi

T : jumlah periode waktu

k : banyaknya variabel bebas tanpa intersep

Adjusted R² digunakan karena sudah menghilangkan pengaruh penambahan variabel bebas dalam model, karena R^2 akan terus naik seiring dengan penambahan variabel bebas. Penggunaan *adjusted R²* sudah memperhitungkan jumlah derajat bebas.