

BAB III METODOLOGI PENELITIAN

3.1 Teknik Pengumpulan Data

Tahapan metodologi terdiri atas 6 (enam) tahapan, *pertama*, dengan mengacu kepada teori, kita mengajukan suatu hipotesis atau pertanyaan. *Kedua*, untuk menjawab pertanyaan atau hipotesis yang diajukan pada tahap pertama, kita mengajukan model ekonometri yang dapat digunakan untuk mengetes hipotesis kita. *Ketiga*, setelah modelnya sudah terbangun, parameter dari model tersebut kita estimasi dengan suatu software computer. *Keempat*, hasil dari estimasi parameter perlu kita verifikasi terlebih dahulu apakah hasilnya sesuai dengan model atau tidak. *Kelima*, jika dari hasil verifikasi mengatakan model yang telah terestimasi sudah layak, maka model tersebut kita gunakan untuk memprediksi pergerakan atau memprediksi nilai suatu variabel. *Keenam*, akhirnya, prediksi tersebut dapat kita gunakan sebagai bahan pertimbangan dalam pengambilan suatu keputusan atau suatu kebijakan. (Nachrowi, 2006).

Dalam melihat hubungan faktor-faktor yang mempengaruhi kebijakan upah buruh di Indonesia terutama Upah Minimum Propinsi (UMP) maka data-data yang digunakan adalah data sekunder yaitu data yang bersumber pada instansi pemerintah yang telah dipublikasikan dan data yang diolah kembali dari data sekunder yang diterbitkan oleh instansi pemerintah seperti Badan Pusat Statistik (BPS) lembaga pemerintah yang diakui dan mempunyai legalitas dalam menerbitkan data statistik di Indonesia maupun Dinas Tenaga Kerja dan Badan Perencanaan Daerah.

Data-data sekunder yang dikumpulkan disini adalah data-data yang diduga sebagai variabel yang mempunyai pengaruh terhadap penetapan Upah Minimum Propinsi (UMP) di Indonesia. Dengan menggunakan data panel. Data panel adalah data yang dikumpulkan secara *cross section* dan diikuti pada periode waktu tertentu. Karena data panel merupakan gabungan dari data *cross-section* dan data *time series*, jumlah pengamatan menjadi sangat banyak. Hal ini bisa merupakan keuntungan (data banyak) tetapi model yang menggunakan data

ini menjadi lebih kompleks (parameternya banyak). Maka modelnya ditulis dengan:

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

Dimana:

N = banyaknya observasi

T = banyaknya waktu

N x T = banyaknya data panel

Metode yang digunakan diperoleh dari prosedur untuk mendapatkan hasil pendugaan parameter yang memiliki sifat tak bias linier terbaik (*Best Linear Unbiased Estimator/BLUE*). Secara ringkas, sifat BLUE mengandung arti bahwa pendugaan parameter yang dihasilkan akan memiliki varian yang minimum dan tidak berarti pendugaan dari masing-masing sample akan sama dengan populasinya. Metode pendugaan untuk mendapatkan estimasi yang bersifat BLUE mensyaratkan sejumlah asumsi yaitu: (i) hubungan antara variable independent dengan variable dependennya bersifat linier, (ii) variable dependennya adalah non stokastik, (iii) tidak ada hubungan linier antar variable dependen (*no multicollinearity*), (iv) galat memiliki nilai harapan nol (0), (v) galat memiliki varians yang konstan untuk semua observasi (*no heteroscedasticity*), (vi) galat ϵ_i adalah independent secara statistik (*no serial autocorrelation*).

Untuk metode analisis datanya menggunakan metode analisis deskriptif, yakni suatu cara analisa langsung dari hasil Stata dan tabel dengan memanfaatkan data-data yang tersedia dalam tabel seperti rata-rata, persentase, elastisitas dan ukuran statistik lainnya.

Untuk menganalisis faktor-faktor yang mempengaruhi kebijakan upah buruh di Indonesia alat analisa yang digunakan adalah Stata versi 8 dengan alat ini didapatkan nilai koefisien parameter masing-masing variabel, nilai R-square, uji signifikansi baik untuk uji masing-masing variabel (uji t) maupun untuk uji variabel secara keseluruhan (uji F). Dengan Stata juga dapat diketahui ada tidaknya pelanggaran asumsi baik itu heteroskedastisitas, multikolinearitas maupun autokorelasi.

Untuk melihat pengaruh Kebutuhan Hidup Minimum (KHM), Inflasi (IHK), dan tingkat perkembangan perekonomian (PDRB) di 28 propinsi di Indonesia, maka penulis menggunakan metode analisis regresi berganda (*multiple regression*) terhadap faktor-faktor yang mempengaruhi Upah Minimum Propinsi (UMP) di (28) dua puluh delapan propinsi di Indonesia.

Pada regresi berganda peubah tak bebas Y (*variable dependen*). Tergantung kepada dua atau lebih peubah bebas (*variable independent*). Garis regresi merupakan garis yang menghubungkan rata-rata distribusi Y dengan seluruh kemungkinan nilai-nilai X . Variabel bebas (X) adalah variabel yang nilainya dapat ditentukan dan ditulis pada ruas kiri persamaan. Variabel terikat (Y) adalah suatu variabel sebagai akibat dari perubahan yang terjadi pada variabel bebas (X). Untuk menyatakan kuat tidaknya hubungan linear antara X dan Y dapat diukur koefisien korelasi (*Coefficient Correlation*) atau r . Dan untuk dapat mengetahuinya besarnya sumbangan (pengaruh) dari X terhadap perubahan Y dapat dilihat dari koefisien determinasi (*Coeffisien of Determination*) atau R^2 .

Bentuk persamaan regresi linear berganda adalah sebagai berikut:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \epsilon_i \quad (1)$$

Bentuk persamaan 1 (satu) diatas dianggap fungsi model merupakan persamaan linear dimana hubungannya bersifat linear berupa garis lurus. Alternatif persamaan yang mungkin digunakan adalah model Logaritma Natural (\ln) bentuk persamaan linear yang diasumsikan adalah sebagai berikut:

$$\ln Y_i = \beta_0 + \beta_1 \ln X_{1i} + \beta_2 \ln X_{2i} + \beta_3 \ln X_{3i} + \epsilon_i \quad (2)$$

Persamaan dua (2) menunjukkan hubungan logaritma atau hubungan elastisitas variabel-variabel bebas terhadap variabel terikat atau sensitifitas variabel bebas terhadap variabel terikat yang menunjukkan persentase pengaruh variabel independen (bebas) terhadap variabel dependen (terikat).

Dimana:

$$i = 1, 2, 3, \dots, n$$

$$Y_i = \text{Upah Minimum Propinsi}$$

$$X_{1i} = \text{Kebutuhan Hidup Minimum (KHM)}$$

$$X_{2i} = \text{Indeks Harga Konsumen (IHK)}$$

$$X_{3i} = \text{Tingkat perkembangan perekonomian (PDRB)}$$

β_0 = Intersept

$\beta_1, \beta_2, \beta_3$, merupakan penduga (koefisien regresi) model persamaan dan ε_1 adalah besaran yang membuat nilai Y menyimpang dari garis regresinya.

3.2 Metode Estimasi

Didalam teori ekonometri, suatu model yang menyatukan antara data deret waktu (*time series*) dan data kerat lintang (*cross section*) menghasilkan data yang disebut *pooled data* atau data panel atau *longitudinal data*, sehingga dalam data panel jumlah observasi merupakan hasil kali observasi deret waktu ($T > 1$) dengan observasi kerat lintang ($N > 1$).

Adanya keragaman data antar daerah yang begitu besar dalam kualitas SDM dan kinerja ekonomi, maka metode data panel dapat digunakan untuk mengendalikan heterogenitas individu antar-daerah tersebut. Keuntungan menggunakan teknik panel data menurut Baltagi (1995) adalah

- (1) Dapat mengendalikan heterogenitas individu;
- (2) Dengan mengkombinasikan observasi berdasarkan deret waktu dan kerat lintang, maka data panel memberikan informasi yang lebih lengkap, bervariasi, kolienaritas antar variabel menjadi berkurang, serta memperbesar derajat kebebasan, sehingga lebih efisien;
- (3) Dapat meneliti karakteristik individu yang mencerminkan dinamika antar-waktu dari masing-masing variabel bebas, sehingga analisis lebih komprehensif dan mencakup hal-hal yang mendekati realitas;
- (4) Data panel dapat digunakan dalam membangun dan menguji model perilaku yang lebih kompleks.

Disamping memiliki keuntungan, model data panel juga memiliki beberapa kekurangan (Baltagi, 1995), yaitu:

- (1) Masalah koleksi data dan efisien;
- (2) Kemungkinan distorsi dari kesalahan pengukuran;
- (3) Dimensi seri waktu yang lebih pendek.

Dalam estimasi model data panel terdapat tiga pilihan yang dapat dilakukan yaitu: *Common Effect*, *Fixed Effect*, dan *Random Effect*. *Common Effect* merupakan teknik estimasi data panel yang paling sederhana yaitu dengan cara mengkombinasikan data *time series* dan *cross section* dengan metode OLS.

Pendekatan ini tidak memperhatikan dimensi individu maupun waktu, sehingga intersep dan slope dianggap sama (konstan). Model *Common Effect* dapat ditulis:

$$y_{it} = a + \beta X_{it} + \varepsilon_{it}$$

Dimana: $i = 1, 2, \dots, N$ (jumlah data kerat lintang atau *cross section*)

$t = 1, 2, \dots, T$ (jumlah data runtun waktu atau *time series*)

Fixed effect sudah memperhatikan keragaman atau heterogenitas individu yakni dengan mengasumsikan bahwa intersep antar kelompok individu berbeda, sedangkan slope-nya dianggap sama. Pengertian *Fixed effect* didasarkan adanya perbedaan intersep antara individu namun sama antar waktu (*time invariant*), sedangkan koefisien regresi (*slope*) dianggap tetap baik antar kelompok individu maupun antar-waktu. Dalam model *Fixed effect*, generalisasi secara umum sering dilakukan dengan cara memberikan variabel boneka (*dummy variable*). Tujuannya adalah untuk mengizinkan terjadinya perbedaan nilai parameter yang berbeda-beda baik lintas unit *cross section* maupun antar waktu. Model *Fixed effect* dapat ditulis:

$$y_{it} = a_i + \beta X_{it} + \gamma_i \sum D_i + \varepsilon_{it}$$

atau dalam bentuk covariance model dapat ditulis:

$$y_{it} = a_i + \beta X_{it} + \gamma_2 W_2 + \gamma_3 W_3 + \dots + \gamma_N W_{n1} + \delta_2 Z_{i2} + \delta_3 Z_{i3} + \dots + \delta_T Z_{iT} + \varepsilon_{it}$$

dimana : $W_{it} = 1$; untuk unit individu ke i , $i = 2, \dots, N$;

$W_{it} = 0$; lainnya;

$Z_{it} = 1$; untuk periode waktu ke t , $t = 2, \dots, T$;

$Z_{it} = 0$; lainnya.

Keputusan untuk memasukan variabel boneka dalam model efek tetap tak dapat dipungkiri akan menimbulkan konsekuensi, yaitu mengurangi banyaknya derajat kebebasan (*degree of freedom*), sehingga akan mengurangi efisiensi parameter yang diestimasi. Berkaitan dengan hal ini, dalam model data panel dikenal pendekatan ke tiga yaitu model acak (*Random Effect*). Dalam *Random Effect*, parameter-parameter yang berbeda antar daerah maupun antar waktu dimasukan kedalam error. Oleh karena inilah, model acak sering juga disebut model komponen error (*Error Component Model*). diasumsikan pula bahwa error secara individu (U_i) tidak saling berkorelasi, begitu juga dengan error kombinasinya (ε_{it}). Model *Random Effect* dapat ditulis:

$$y_{it} = a + \beta X_{it} + U_i + \varepsilon_{it}$$

Untuk memilih salah satu model estimasi yang dianggap paling tepat dari tiga jenis model data panel, maka perlu dilakukan serangkaian uji, yaitu: (1) Uji F statistik untuk menentukan perlu tidaknya memakai metode estimasi dengan individual effect atau memilih antara *common effect* vs *fixed effect*; (2) Uji Lagrange Multiflier (LM) untuk memilih menggunakan struktur heteroskedastik atau homokedastik, atau untuk memilih antara *common effect* vs *random effect*; dan (3) Uji Hausman untuk menentukan pilihan metode estimasi antara *fixed effect* vs *random effect*.

(1) Uji F Statistik

Uji F statistik merupakan uji perbedaan dua regresi dalam hal ini regresi data panel dengan asumsi intersep dan *slope* sama (*common effect*). Uji ini dilakukan dengan membandingkan residual *sum of squares* (RSS) dari kedua hasil regresi tersebut. Rumusnya adalah :

$$F = \frac{(RSS_1 - RSS_2) / m}{(RSS_2) / (n - k)}$$

Dimana :

RSS_1 : *residual sum of squares* dengan *common effect*;

RSS_2 : *residual sum of squares* dengan *fixed effect*;

M : numerator (yaitu jumlah restriksi atau pembatasan dalam model *common effect* atau jumlah kelompok individu dikurangi 1;

$(n - k)$: denominator; n : jumlah observasi; k: jumlah parameter dalam model *fixed effect*

Hipotesis nolnya adalah *intersep* dan *slope* sama (*common effect*). Nilai statistik hitung hingga akan mengikuti distribusi statistik F dengan derajat bebas (df) sebanyak m untuk numerator dan $(n - k)$ untuk denominator.

(2) Uji Lagrange Multiflier (LM)

Uji LM digunakan untuk memilih model *random effect* atau *common effect*. Uji bisa juga dinamakan uji signifikansi *random effect* yang dikembangkan oleh Bruesch–Pagan (1980). Uji LM Bruesch–Pagan ini didasarkan pada nilai residual dari metode *common effect*. Nilai LM dihitung dengan rumus:

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

Dimana: n = jumlah individu; T = jumlah periode waktu;
 e = residual metode *common effect*

Hipotesis nolnya adalah intersep dan slope sama (*common effect*). Uji LM ini didasarkan pada distribusi *chi-square* dengan *degree of freedom* sebesar jumlah variabel independen. Jika nilai LM statistik lebih besar dari nilai kritis statistik *chi-square* maka kita menolak hipotesis nol, berarti estimasi yang lebih tepat dari regresi data panel adalah model *random effect*. Sebaliknya jika nilai LM statistik lebih kecil dari nilai kritis statistik *chi-square* maka kita menerima hipotesis nol yang berarti model *common effect* lebih baik digunakan dalam regresi.

(3) Uji Hausman

Untuk menentukan salah satu model diantara dua model *fixed effect* atau *random effect*, Hausman (1978) telah mengembangkan suatu uji yang didasarkan pada ide bahwa *Least Square Dummy Variable* (LSDV) dalam metode *Fixed Effect* dan *Generalized Least Square* (GLS) pada *Random Effect* adalah efisien, sedangkan metode *Ordinary Least Square* (OLS) tidak efisien, dilain pihak alternatifnya adalah metode OLS efisien dan GLS tidak efisien. Oleh karena itu, hipotesis nolnya adalah hasil estimasi keduanya tidak berbeda sehingga uji Hausman bisa dilakukan berdasarkan perbedaan estimasi tersebut. Unsur penting untuk uji ini adalah

kovarian matrik dari perbedaan vektor $\left[\hat{\beta} - \hat{\beta}_{GLS} \right]$:

$$\text{Var} \left[\hat{\beta} - \hat{\beta}_{GLS} \right] = \text{Var} \left[\hat{\beta} \right] + \text{Var} \left[\hat{\beta}_{GLS} \right] - \text{Cov} \left[\hat{\beta}, \hat{\beta} \right] - \text{Cov} \left[\hat{\beta}, \hat{\beta}_{GLS} \right]^1$$

Hasil metode Hausman adalah bahwa perbedaan kovarian dari estimator yang efisien dengan estimator yang tidak efisien adalah nol, sehingga

$$\text{Cov} \left[\hat{\beta} - \hat{\beta}_{GLS}, \hat{\beta}_{GLS} \right] \text{Cov} \left[\hat{\beta}, \hat{\beta}_{GLS} \right] = \text{Var} \left[\hat{\beta}_{GLS} \right] = 0$$

$$\text{Cov} \left[\hat{\beta}, \hat{\beta}_{GLS} \right] = \text{Var} \left[\hat{\beta}_{GLS} \right]$$

Kemudian kita masukkan ke dalam persamaan akan menghasilkan kovarian matriks sebagai berikut:

$$\text{Var}\left[\hat{\beta} - \hat{\beta}_{\text{GLS}}\right] = \text{Var}\left[\hat{\beta}\right] - \text{Var}\left[\hat{\beta}_{\text{GLS}}\right] = \text{Var}\left[\hat{q}\right]$$

Selanjutnya mengikuti kriteria Wad, uji Hausman ini akan mengikuti distribusi *chi-square* sebagai berikut:

$$W = X^2(k) = \hat{q}' \text{Var}\left[\hat{q}\right]^{-1} \hat{q}$$

$$\text{Dimana } \hat{q} = \left[\hat{\beta} - \hat{\beta}_{\text{GLS}}\right] \text{ dan } \text{Var}\left[\hat{q}\right] = \text{Var}\left[\hat{\beta}\right] - \text{Var}\left[\hat{\beta}_{\text{GLS}}\right]$$

Statistik uji Hausman mengikuti distribusi statistik *chi-square* dengan *degree of freedom* sebanyak k , dimana k adalah jumlah variabel independen. Jika nilai statistik Hausman lebih besar dari nilai kritisnya, maka model yang tepat adalah model *fixed Effect*, sedangkan sebaliknya bila nilai statistik uji Hausman lebih kecil dari nilai kritisnya maka model yang tepat adalah model *Random Effect*.

3.3 Uji Asumsi Klasik

Sebagai upaya untuk menghasilkan model yang efisien, fisibel dan konsisten, maka perlu pendeteksian terhadap pelanggaran asumsi model yaitu gangguan antara waktu (*time-related disturbance*), gangguan antar individu (*cross sectional disturbance*), dan gangguan akibat keduanya.

(1) Uji Otokerasi

Otokorelasi merupakan hubungan yang terjadi antara galat dari suatu observasi dengan galat observasi lainnya. Otokorelasi biasanya terjadi pada data deret waktu, yaitu apabila galat suatu observasi mempengaruhi galat observasi pada periode berikutnya. Otokorelasi biasanya tidak muncul pada data kerat lintang, karena hanya menunjukkan satu titik waktu saja.

Dampak yang ditimbulkan akibat otokorelasi adalah varians dari taksiran OLS akan *underestimate*, walaupun taksiran OLS tidak bias. Keberadaan otokorelasi dapat dideteksi melalui *Durbin Watson Test* yang membandingkan nilai DW hitung dengan nilai batas bawah (d_1) dan batas atas (d_u) dari tabel Durbin Watson berdasarkan jumlah observasi dan variabel bebas. Adapun persamaan uji Durbin Watson adalah:

$$d = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n (e_i)^2}$$

Selang kepercayaan untuk menganalisis otokorelasi dibagi menjadi 5 daerah yaitu:

- Daerah A, yaitu untuk nilai uji DW < d1 (korelasi positif)
- Daerah B, yaitu untuk d1 < nilai uji DW < du (tidak dapat disimpulkan ada atau tidaknya otokorelasi)
- Daerah C, yaitu du < nilai uji DW < 4 – du (tidak ada korelasi)
- Daerah D, yaitu 4 – du < nilai uji DW < 4 – d1 (tidak dapat disimpulkan ada atau tidaknya otokorelasi)
- Daerah E, yaitu 4 – d1 < DW < 4 (korelasi positif)

Masalah otokorelasi dapat diatasi dengan cara memasukan variabel autogressive-nya. Hipotesis yang digunakan dalam menentukan ada atau tidaknya autokorelasi dalam model persamaan regresi adalah sebagai berikut:

Ho : tidak ada autokorelasi

H1 : ada autokorelasi

Akibat dari adanya autokorelasi maka estimasi koefisien regresi tidak unbiased, tetapi standar error koefisien regresi terlalu rendah, akibatnya pengujian F dan t menjadi tidak valid (cenderung signifikan).

Untuk membuktikan ada atau tidaknya autokorelasi dalam model persamaan ini maka digunakan *Breusch-Godfrey Serial Correlation LM Test* yang telah disediakan dalam program Stata. Untuk melakukan uji signifikan adanya autokorelasi dapat dilihat dari hasil output Stata.

Jika $\alpha >$ probabilitas, maka tolak Ho pada tingkat kepercayaan 95%

Jika $\alpha <$ probabilitas maka terima Ho pada tingkat kepercayaan 95%.

(2) Uji Heteroskedastisitas

Heteroskedastisitas terjadi jika varians pengganggu berbeda antar observasi, sehingga tiap-tiap observasi memiliki nilai reliabilitas yang berbeda. Metode estimasi OLS memberikan angka penimbang yang lebih besar pada observasi yang memiliki yang lebih besar dengan maksud meminimalkan jumlah kuadrat dari galat.

Heteroskedastisitas menimbulkan dampak tidak efisiennya proses estimasi, namun hasil estimasi tetap konsisten dan tidak bias. Masalah Heteroskedastisitas menjadikan uji t dan uji F menjadi menyesatkan. Pada umumnya, Heteroskedastisitas terjadi pada data kerang lintang. Untuk menguji ada tidak variasi error yang berpola atau yang disebut dengan heteroskedastisitas, hipotesisnya adalah sebagai berikut:

Ho : Homokedastisitas

H1 : Heteroskedastisitas

Akibat adanya heteroskedastisitas maka nilai koefisien tidak berbias, tetapi varian estimasi koefisien regresi tidak minimal lagi. Akibatnya pengujian F dan t cenderung tidak signifikan, dan ini berarti akan terjadi kesalahan dalam pengambilan kesimpulan. Keberadaan Heteroskedastisitas dapat diuji dengan *Park test*, *Goldfelt-Quandt test*, *Breusch-Pagan-Godfrey Test*, dan *White General Heteroscedasticity* (Gujarati, 2003).

Untuk membuktikan adanya heteroskedastisitas, jika $\alpha >$ probabilitas, maka tolak Ho; pada tingkat kepercayaan 95% (α 5 %), dan jika $\alpha <$ probabilitas, maka terima Ho; pada tingkat kepercayaan 95% (α 5 %).

(3) Uji Multikolinearitas

Metode estimasi yang menghasilkan pendugaan yang memiliki ciri *Best Linier Unbiased Estimation (BLUE)* mensyaratkan tidak adanya hubungan linear antar variabel independen atau tidak ada kolinearitas jamak. Sebaliknya jika diantara variabel independen memiliki korelasi linier yang tinggi, maka model pendugaan tersebut dikatakan terdapat kolinearitas jamak yang serius. Kolinearitas jamak yang serius akan berdampak pada:

- kesulitan untuk memisahkan efek suatu variable independent terhadap variable dependen dengan efek dari variable independent yang lain.
- distribusi parameter regresi menjadi sangat sensitive terhadap korelasi yang terjadi antar variable independent dan galat baku regresi. Kondisi ini muncul dalam bentuk varians dan galat baku parameter yang tinggi yang berdampak pada nilai t statistik menjadi lebih kecil sehingga variable independent tersebut menjadi tidak signifikan pengaruhnya. Pengaruh lebih lanjutnya adalah bahwa koefisien regresi yang dihasilkan tidak mencerminkan

nilai sebenarnya dimana sebahagian koefisien cenderung overestimate dan yang lainnya underestimate.

2. Multikolinearitas adalah hubungan antara variabel bebas, yaitu suatu kondisi adanya korelasi yang kuat antara variabel bebas $X_1=f(X_2)$ atau $X_2=f(X_3)$ atau sebaliknya. Untuk menentukan adanya multikolinearitas dapat ditentukan melalui matrik korelasi atau meregresi antar variabel bebas dalam model persamaan. Dalam makalah ini alat yang digunakan untuk mendeteksi adanya multikolinearitas adalah dengan menggunakan hasil output Stata yakni matriks multikolinearitas.

Akibat adanya multikolinearitas ini maka standar error koefisien regresi yang diduga akan besar. Akibatnya nilai uji t menjadi rendah, sehingga variabel yang seharusnya signifikan dapat menjadi tidak signifikan. Lebih jauh lagi, tidak hanya variabel tidak signifikan, tetapi juga mempunyai tanda koefisien yang salah, akibatnya akan bertentangan dengan teori yang melandasinya.

Multikolinearitas muncul jika variabel bebas memiliki korelasi yang tinggi, sehingga sulit memisahkan efek satu variabel bebas terhadap variabel terikat dari efek variabel bebas lainnya.

Indikasi-indikasi terjadinya Multikolinearitas:

- Jika ditemukan nilai R^2 yang tinggi dan nilai statistik F yang signifikan tetapi sebagian besar nilai statistik t tidak signifikan.
- Korelasi sederhana yang relatif tinggi (0,8 atau lebih) antara satu atau lebih pasang variabel bebas. Jika koefisien korelasi kurang dari 0,8 berarti masalah tidak terlalu serius, belum terjadi Multikolinearitas.
- Regresi bantuan (*Auxiliary Regression*), dengan cara meregresi masing-masing variabel bebas pada variabel bebas lainnya. Apabila nilai R^2 nya tinggi maka ada indikasi ketergantungan linier yang hampir pasti diantara variabel-variabel bebas.

Dan untuk mengatasinya cukup sederhana yakni mengeluarkan variabel bebas (*independen*) yang diperkirakan mempunyai korelasi cukup tinggi dengan variabel lain, akan tetapi perlu dipertimbangkan dulu landasan teorinya sebelum mengeluarkan variabel yang mengandung korelasi tersebut. Atau dapat juga dengan mengubah bentuk model dan menambah data atau memilih sampel baru.

3.4 Kerangka Berfikir

