

BAB III

METODE PENELITIAN

3.1 Jenis dan Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yaitu data yang bukan diusahakan sendiri pengumpulannya oleh peneliti. Data sekunder yang digunakan dalam penelitian ini adalah penggabungan data *cross section* dari 35 Kabupaten/Kota di Provinsi Jawa Tengah dan *Times Series* dari tahun 2011 sampai tahun 2015. Metode pengumpulan data yang digunakan dalam penelitian ini adalah metode dokumentasi, penulis mengumpulkan data dari data-data yang dipublikasikan oleh BPS Provinsi Jawa Tengah dalam berbagai tahun publikasi seperti Tengah Dalam Angka, statistik Sosial dan Kependudukan Jawa Tengah Hasil SUSENAS, Indikator Kesejahteraan Rakyat Jawa Tengah, Indikator Sosial Politik dan Keamanan Jawa Tengah.

3.2 Definisi Operasional Variabel

Variabel yang penelitian yang digunakan ada 4 variabel bebas (independen) dan variabel terikat (dependen). Variabel independen dalam penelitian ini adalah, tingkat tingkat upah minimum, tingkat pengangguran terbuka, jumlah penduduk, dan Angka Meleak Huruf (AMH), sedangkan variabel terikat dalam penelitian ini adalah tingkat kemiskinan. Untuk memperjelas dan mempermudah pemahaman terhadap variabel-variabel yang akan dianalisis

dalam penelitian ini, maka perlu dirumuskan definisi operasional variable, yaitu sebagai berikut :

1. Kemiskinan

Kemiskinan merupakan suatu kondisi ketidakmampuan memenuhi standar minimum kebutuhan dasar yang meliputi kebutuhan makan maupun non makan. Data kemiskinan yang dipakai dalam penelitian ini adalah data kemiskinan Provinsi Jawa Tengah tahun 2011-2015 (Dalam Persen).

2. Upah minimum

Upah minimum adalah upah bulanan terendah yang terdiri dari upah pokok termasuk tunjangan tetap. Data yang dipakai dalam penelitian ini adalah upah minimum Provinsi Jawa Tengah tahun 2011-2015 (dalam rupiah).

3. Pengangguran

Pengangguran sendiri merupakan suatu keadaan di mana seseorang yang tergolong dalam angkatan kerja ingin mendapatkan pekerjaan tetapi belum dapat memperolehnya (Sukirno, 2011). Data yang dipakai dalam penelitian ini adalah tingkat pengangguran terbuka Provinsi Jawa Tengah tahun 2011-2015 (dalam persen).

4. Jumlah Penduduk

Jumlah penduduk merupakan jumlah semua orang yang berdomisili di wilayah geografis Republik Indonesia selama enam bulan atau lebih dan atau mereka yang berdomisili kurang dari enam

bulan tetapi bertujuan untuk menetap. Data yang dipakai dalam penelitian ini adalah jumlah penduduk Provinsi Jawa Tengah tahun 2011-2015 (dalam jiwa).

5. Angka Melek Huruf

Angka Melek Huruf merupakan proporsi penduduk berusia 15 tahun keatas yang memiliki kemampuan membaca dan menulis kalimat sederhana terhadap penduduk usia 15 tahun keatas. Data yang dipakai dalam penelitian ini adalah angka melek huruf Provinsi Jawa Tengah tahun 2011-2015 (dalam persen).

3.3 Metode Analisis

Penelitian ini menggunakan metode penelitian kuantitatif dengan menggunakan data panel (*pooled data*) sebagai alat pengolahan data serta dalam analisisnya menggunakan program *Eviews* 8. Data panel merupakan data gabungan antara data (*time series*) dan (*cross section*). Analisis ekonometrika dengan menggunakan data panel akan memberikan beberapa manfaat terkait dengan adanya kelemahan pada dua jenis data lainnya, yaitu mengatasi heterogenitas data, keterbatasan data, dan sekaligus mengatasi variasi data sebagai akibat perubahan waktu (*time variant*). Adapun persamaan model dalam bentuk *cross section* dapat dituliskan dalam model berikut :

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i ; i = 1, 2, \dots, n$$

Dimana β_0 adalah intersep atau konstanta, β_1 adalah koefisien regresi, dan ϵ_i adalah variabel gangguan (*error*). Sedangkan persamaan model regresi *Times series* adalah sebagai berikut :

$$Y_t = \beta_0 + \beta_1 X_t + \epsilon_t ; t = 1, 2, \dots, t$$

Dimana t menunjukkan banyaknya periode waktu data *time series*. Mengingat data panel merupakan gabungan dari data *cross section* dan data *time series*, maka model regresi data panel dapat dituliskan sebagai berikut :

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \epsilon_{it}$$

$$n = 1, 2, \dots, n ; t = 1, 2, \dots, t$$

Dimana n adalah banyaknya variable bebas, i adalah jumlah unit observasi, t adalah banyaknya periode waktu, sehingga $(n \times t)$ menunjukkan banyaknya data panel yang akan di analisis. Maka bentuk regresi data panel untuk penelitian ini adalah sebagai berikut :

$$POV = \beta_0 + \beta_1 UMK_{1it} + \beta_2 TPT_{2it} + \beta_3 JP_{3it} + \beta_4 MH_{4it} + \epsilon$$

Dimana :

POV = Persentase jumlah penduduk miskin Kabupaten i tahun t
(persen)

β_0 = Konstanta

$\beta_1 \beta_2 \beta_3$ = Koefisien variable independen

UMP_{1it} = Upah Minimum Kabupaten Kabupaten i tahun t (rupiah)

TPT_{2it} = Tingkat Pengangguran Terbuka Kabupaten i tahun t (persen)

JP_{3it} = Jumlah Penduduk Kabupaten i tahun t (jiwa)

MH_{4it} = Melek Huruf Kabupaten i tahun t (persen)

Karena terdapat perbedaan dalam satuan dan besaran variabel bebas dalam persamaan regresi harus dibuat dengan model logaritma natural. Alasan pemilihan model logaritma natural menurut Ghozali (2005) adalah sebagai berikut :

- a. Menghindari adanya heteroskedastisitas
- b. Mengetahui koefisien yang menunjukkan elastisitas dan,
- c. Mendekatkan skala data.

Dalam model penelitian ini logaritma yang digunakan adalah dalam bentuk semilog linear (semi-log). Dimana semi-log mempunyai beberapa keuntungan diantaranya :

- a. Koefisien – koefisien model semi-log mempunyai interpretasi yang sederhana
- b. Model semi-log sering mengurangi masalah statistik umum yang dikenal sebagai heteroskedastisitas,

c. Model semi-log mudah dihitung .

Sehingga persamaan penelitian dengan model semi-log adalah sebagai berikut :

$$POV = \beta_0 + \beta_1 \text{LogUMK}_{1it} + \beta_2 \text{TPT}_{2it} + \beta_3 \text{LogJP}_{3it} + \beta_4 \text{MH}_{4it} + \varepsilon$$

Keunggulan regresi data panel menurut Wibisono (2005) antara lain:

1. Panel data mampu memperhitungkan heterogenitas individu eksplisit dengan mengizinkan variable spesifik individu
2. Kemampuan mengontrol heterogenitas ini selanjutnya menjadikan data panel dapat digunakan untuk menguji dan membangun model perilaku lebih kompleks
3. Data panel mendasarkan diri pada observasi *cross-section* yang berulang-ulang (*Times Series*), sehingga metode data panel cocok digunakan sebagai *study of dynamic adjustment*
4. Tingginya jumlah observasi memiliki implikasi pada data yang lebih informative, lebih variatif, dan kolinearitas (multikolinieritas) antara data semakin berkurang, dan derajat kebebasan (*degree of freedom*) lebih tinggi sehingga dapat diperoleh hasil estimasi yang lebih efisien.
5. Data panel dapat digunakan untuk mempelajari model-model perilaku yang lebih kompleks
6. Dapat digunakan untuk meminimalkan bias yang mungkin ditimbulkan oleh agregasi data individu

3.3.1 Metode Analisis Data Panel

Dalam mengestimasi data panel menggunakan tiga macam pendekatan, yaitu :

1. Pendekatan Pendekatan pooled Least Square (OLS) atau *common effect*:

Estimasi model common effect merupakan salah satu model analisis regresi data panel yang paling sederhana. Dalam model analisis ini diansumsikan bahwa intersep dan slope selalu tetap baik antar waktu maupun antar individu. Sistematis model ini menggabungkan antara data *time series* dan data *cross section* ke dalam data panel (pooled data) yang dapat diregresi menggunakan metode OLS.

2. Pendekatan *Fixed Effect*

Model estimasi ini merupakan suatu pendekatan yang mengansumsikan bahwa perbedaan antar individu dapat diakomodasi dari perbedaan intersepnya. Untuk mengestimasi data panel model ini menggunakan variable dummy untuk menangkap perbedaan intersep antar individu. Namun demikian, slopenya sama antar individu.

3. *Random effect Model*

Model ini mengestimasi data panel di mana variable gangguan mungkin saling berhubungan antar waktu atau antar individu. Pada model ini intersep diakomodasi oleh *error terms* masing-masing individu, model ini juga disebut dengan *Error Component Model (ECM)* atau teknik *Generalied Least Square (GLS)*.

3.3.2 Pemilihan Model Dalam Pengolahan Data

Menurut Sriyana (2014) untuk mengestimasi regresi data panel dengan menggunakan metode *common effect*, *fixed effect*, dan *random effect*, beberapa model yang digunakan adalah:

1. Uji F Statistik (Chow) , untuk memilih model yang lebih baik antara model *common effect* (tanpa variable dummy) atau dengan model *fixed effect*.
2. Uji Hausman untuk membandingkan antara model *fixed effect* atau *random effect* yang lebih baik untuk digunakan.
3. Uji Langrange Multipler (LM) atau lengkapnya The Breusch-Pagan LM test digunakan untuk memilih model *common effect* (tanpa variable dummy) atau dengan model *random effect*.

3.3.2.1 Uji Signifikansi Fixed Effect (Chow-Test)

Uji Signifikansi *Fixed Effect* merupakan tahap pertama pengujian pemilihan model. Pengujian ini digunakan untuk memilih model antara model *common effect* atau model *fixed effect*. Pengujian signifikansi *fixed effect* dapat dilakukan dengan melakukan uji statistic F. Cara ini dilakukan dengan melihat nilai *residual sum of squares* (RSS) dari kedua model regresi tersebut.

$$F_{\text{Hitung}} = \frac{(RSS1 - RSS2) / (i - t)}{(RSS2) / (it - i - k)}$$

Dimana, RSS1 dan RSS2 masing-masing adalah *residual sum of squares* model *common effect* tanpa variable *dummy* dan model *Fixed Effect* dengan variable *dummy*, i = merupakan jumlah individu, t = jumlah periode waktu, k = banyaknya parameter dalam model *Fixed Effect*.

Nilai F hitung yang diperoleh selanjutnya diuji berdasarkan nilai F tabel dengan df sebesar m untuk numerator dan $n-k$ untuk denominator. Nilai statistic F akan mengikuti nilai distribusi statistic F dengan derajat kebebasan (df) sebesar $n-1$ untuk numerator dan sebesar $(i.t)-k$ untuk denominator.

Kesimpulannya adalah jika nilai statistic F yang dihasilkan lebih besar dibandingkan nilai F-tabel pada tingkat signifikansi yang digunakan maka regresi data panel dengan menggunakan model *Fixed Effect* lebih baik jika dibandingkan dengan model regresi data panel menggunakan *common effect*.

Untuk melihat model mana yang baik digunakan antara *fixed effect* dan *common effect* dapat juga diketahui dengan membandingkan antara nilai probabilitas F-statistik dengan tingkat signifikansi α . Apabila $p < \alpha$, maka model *Fixed Effect* lebih baik digunakan dibandingkan menggunakan model *common effect*.

3.3.2.2 Uji Hausman

Untuk melihat apakah model mengikuti *random effect* atau *fixed effect* digunakan uji hausman. Unsur penting untuk uji ini adalah covarian matriks dari perbedaan vector $(\beta - \beta_{GLS})$:

$$\text{var}(\beta - \beta_{GLS}) = \text{Var}(\beta) + \text{Var}(\beta_{GLS}) - \text{Cov}(\beta, \beta_{GLS}) - \text{Cov}(\beta_{GLS}, \beta) \dots \dots (1)$$

Hasil metode hausman adalah bahwa perbedaan kovarian dari estimator yang efisien dengan estimator yang tidak efisien adalah nol sehingga;

$$\text{Cov}((\beta - \beta_{\text{GLS}}), \beta_{\text{GLS}}) = \text{Cov}(\beta - \beta_{\text{GLS}}) - \text{Var}(\beta_{\text{GLS}}) = 0 \dots \dots \dots (2)$$

$$\text{Cov}(\beta - \beta_{\text{GLS}}) = \text{Var}(\beta_{\text{GLS}})$$

Kemudian memasukkan persamaan (1) akan menghasilkan kovarian matriks sebagai berikut ;

$$\text{Var}(\beta - \beta_{\text{GLS}}) = \text{Var}(\beta) - \text{Var}(\beta_{\text{GLS}}) = \text{Var}(q) \dots \dots \dots (3)$$

Selanjutnya mengikuti distribusi chi-square ;

$$M = q' \text{Vaar}(q)^{-1} q$$

Dimana : $q = (\beta - \beta_{\text{GLS}})$

$$\text{Var}(q) = \text{Var}(\beta) - \text{Var}(\beta_{\text{GLS}}),$$

- a. Statistic hausman > nilai kritisnya → model yang tepat *fixed effect*
- b. Statistik hausman < nilai kritisnya → model yang tepat *random effect*

3.3.2.3 Uji Langrange Multipler (LM)

Menurut Widarjono (2009) uji LM merupakan uji yang secara umum digunakan untuk mendeteksi masalah autokorelasi. Bisa juga digunakan untuk mendeteksi ada tidaknya autokorelasi dalam model autoregresif .Adapun nilai statistic LM dihitung berdasarkan formula :

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

Dimana : n = Jumlah individu

T = Jumlah periode waktu

e = Residual metode OLS

Jika :

1. Nilai statistic LM > nilai kritis statistic chi-square (memilih *random effect*)
2. Nilai statistic LM < nilai kritis statistic chi-square (memilih *common effect*)

3.3.3 Pengujian Statistik

3.3.3.1 Koefisien Determinasi (Uji R²)

Sriyana (2014), uji koefisien determinasi dilakukan untuk menjelaskan seberapa besar garis regresi menjelaskan perilaku datanya. Adapun rumus dari koefisien determinasi sebagai berikut :

$$\begin{aligned}R^2 &= ESS / TSS \\ &= 1 - RSS/TSS \\ &= 1 - \frac{(\sum \hat{e}_i^2)}{\sum y_i^2} \\ &= 1 - \frac{(\sum \hat{e}_i^2)}{\sum (Y_i - \bar{Y})^2}\end{aligned}$$

Dari rumus tersebut tampak bahwa koefisien determinasi akan meningkat sesuai dengan jumlah variable bebasnya. Artinya koefisien determinasi akan semakin besar jika kita terus menambah variable independen dalam model. Para

ahli ekonometrik mengembangkan alternative lain agar R^2 tidak merupakan fungsi darivariable independen. Alternatifnya digunakan R^2 yang disesuaikan ($R^2_{adjusted} = \bar{R}^2$) dengan rumus sebagai berikut :

$$\bar{R}^2 = 1 - \frac{(\sum \hat{e}_i^2) / (n-k)}{\sum (Y_i - \bar{Y})^2 / (n-1)}$$

Dimana :

k = jumlah koefisien regresi, termasuk intersep

n = jumlah observasi

koefisien determinasi memiliki nilai antara 0 sampai dengan 1. Semakin tinggi nilainya maka menunjukkan semakin eratnya hubungan antara variable bebas dengan varabel terikat.

3.3.3.2 Uji t (penguji variable secara individu)

Uji ini digunakan untuk mengetahui signifikan atau tidak signifikannya variable independen terhadap variable dependen. Adapun hipotesis untuk uji t sebagai berikut :

1. Jika hipotesis signifikan positif

a. $H_0 : \beta_i \leq 0$

b. $H_1 : \beta_i > 0$

2. Jika hipotesis signifikan negatif

a. $H_0 : \beta_i \geq 0$

b. $H_1 : \beta_i < 0$

Menentukan tingkat signifikan alfa (α) yaitu sebesar 5 %

Kriteria pengujian :

1. Jika nilai probabilitas t- statistic $> 0,05$, maka H_0 diterima , artinya variable bebas tidak mempengaruhi variable terikat secara signifikan
2. Jika nilai probabilitas t – Statistik $< 0,05$, maka H_0 ditolak, artinya variable bebas mempunyai pengaruh signifikan terhadap variable terikat.

3.3.3.3 Uji F

Uji F digunakan untuk mengetahui significant ataupun tidak signifikan variable independen terhadap variable dependen secara menyeluruh. Uji F ini menggunakan langkah – langkah sebagai berikut :

1. $H_0 : \beta_1 = \beta_2 = \beta_3 = 0$, berarti tidak ada pengaruh signifikan dari variable independen terhadap variable dependen.
 $H_1 : \beta_1 \neq \beta_2 \neq \beta_3 \neq 0$, berarti ada pengaruh signifikan dari variable independen secara bersama-sama terhadap variable dependen.
2. Menentukan besarnya nilai F hitung dan signifikansi F (sig-F)
3. Menentukan tingkat signifikan (α) alfa yaitu sebesar 5 %
4. Kriteria pengujian :
 - a. Jika nilai sig. F $> 0,05$, maka H_0 diterima, artinya variable bebas secara serentak tidak mempengaruhi variable terikat secara signifikan
 - b. Jika nilai sig. F $< 0,05$, maka H_0 ditolak, artinya variable bebas secara serentak mempengaruhi variable terikat secara signifikan.