BAB III

METODE PENELITIAN

F. Tujuan Penelitian

1. Menganalisis pengaruh Produk Domestik Regional Bruto (PDRB) terhadap kemiskinan di Jawa Timur.
2. Menganalisis pengaruh Pendidikan terhadap kemiskinan di Jawa Timur.

B. Objek dan Ruang Lingkup Penelitian

Objek penelitian ini adalah angka kemisikinan di Jawa Timur. angka Kemisikinan di Jawa Timur dipengaruhi oleh PDRB atas harga konstan dan tingkat pendidikan

Ruang lingkup dalam penelitian ini mencakup data Provinsi Jawa Timur, seperti data tingkat kemisikinan di Jawa Timur, data kependudukan, data PDRB , dan data pendidikan. Penelitian ini mengambil data tahunan antara tahun 2010-2015. Waktu penelitian tersebut dipilih karena pada rentang tahun tersebut terjadi transisi pemerintahan dari presiden sebelumnya ke presiden yang sekarang. Selain itu peneliti ingin membandingkan hasil kerja kedua pemerintah dalam hal pengentasan kemiskinan melalui variabel-variabel yang djadikan objek penelitian

C. Variabel Operasional Penelitian

1. Kemiskinan (KM)

a. Definisi Konseptual

Kemiskinan berarti sejumlah penduduk yang tidak dapat memenuhi kebutuhan dasar hidup yang telah ditetapkan oleh suatu badan atau orang tertentu dan perhitungan yang dilakukan oleh suatu badan atau orang tertentu dan perhitungan yang dilakukan oleh badan atau organisasi tersebut digunakan sebagai standar perhitungan untuk menentukan jumlah kemiskinan yang ada di suatu daerah.

b. Definisi Operasional

Kemiskinan berarti sejumlah penduduk yang tidak dapat memenuhi kebutuhan dasar hidup yang telah ditetapkan oleh suatu badan atau orang tertentu dan perhitungan yang dilakukan oleh suatu badan atau orang tertentu dan perhitungan yang dilakukan oleh badan atau organisasi tersebut digunakan sebagai standar perhitungan untuk menentukan jumlah kemiskinan yang ada di suatu daerah, yaitu Provinsi Jawa Timur. Data persentase penduduk miskin yang digunakan adalah garis kemiskinan yang ditetapkan Badan Pusat Statistik (BPS). Dalam data ini, data yang digunakan adalah persentase penduduk miskin tahun 2011 – 2015 (dalam satuan persen).

2. PDRB (PDRB)

a. Definisi Konseptual

PDRB adalah keseluruhan nilai barang dan jasa yang diproduksi didalam suatu daerah tertentu dalam satu tahun tertentu.

b. Definisi Operasional

PDRB adalah keseluruhan nilai barang dan jasa yang diproduksi didalam suatu daerah tertentu dalam satu tahun tertentu, sementara PDRB atas dasar harga konstan adalah jumlah nilai produksi atau pendapatan atau pengeluaran yang dinilai atas dasar harga tetap (harga pada tahun dasar) yang digunakan selama satu tahun. Dalam variabel ini data yang akan digunakan adalah data laju PDRB atas dasar harga konstan yang ditetapkan Badan Pusat Statistik (BPS) dan terjadi di Provinsi Jawa Timur miskin tahun 2011 – 2015.

3. Pendidikan (PP)

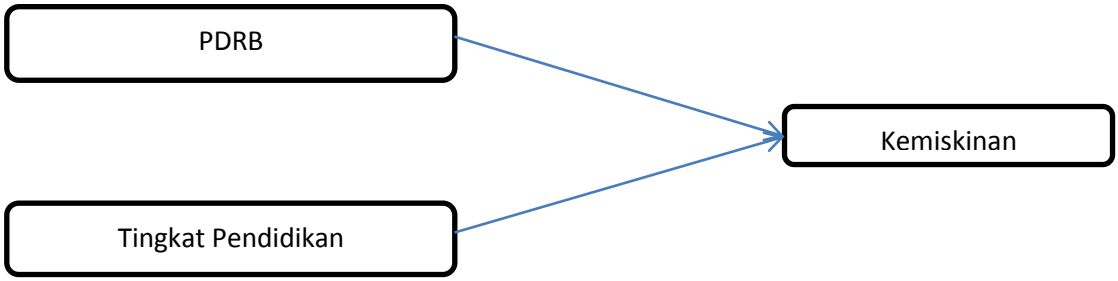
a. Definisi Konseptual

Pendidikan adalah sebuah usaha yang di lakukan secara sadar dan terencana untuk mewujudkan suasana belajar dan proses pembelajaran agar peserta didik secara aktif mengembangkan potensi dirinya untuk memiliki kekuatan spiritual keagamaaan, membangun kepribadian, pengendalian diri, kecerdasan, akhlak mulia, serta ketrampilan yang diperlukan dirinya, masyarakat, bangsa, dan Negara.

b. **Definisi Operasional**

Indikator tingkat pendidikan terdiri dari jenjang pendidikan dan kesesuaian jurusan. Jenjang pendidikan adalah tahapan pendidikan yang ditetapkan berdasarkan tingkat perkembangan peserta didik, tujuan yang akan dicapai, dan kemampuan yang dikembangkan, yaitu terdiri dari pendidikan dasar dan pendidikan tinggi. Data yang digunakan dalam penelitian ini adalah tingkat pendidikan masyarakat Jawa Timur pada jenjang SMA sederajat hingga tingkat universitas pada tahun 2011 – 2015 (dalam satuan persen).

Konstelasi pengaruh antar variabel dapat digambarkan sebagai berikut



Gambar 3.1 Konstelasi Pengaruh Antar Variabel

Keterangan :

PDRB	: Variabel Independen X1
Tingkat Pendidikan	: Variabel Independen X3
Kemiskinan	: Variabel Dependen Y
⟶	: Arah Pengaruh

D. Jenis Dan Sumber Data

Data yang dipergunakan dalam penelitian ini adalah data sekunder. Menurut Istijanto yang dimaksud dengan data skunder adalah data yang telah dikumpulkan oleh pihak lain bukan oleh periset itu sendiri, yang digunakan untuk tujuan yang lain.[27] Data yang digunakan adalah data panel.[28] Pemilihan periode ini disebabkan karena kemiskinan mengalami fluktuasi dan terjadinya peningkatan PDRB dan diiringi dengan fluktuasi tingkat pendidikan masyarakat Jawa Timur pada tahun 2011-2015. Selain itu, adanya peralihan kekuasaan Indonesia membuat penelitian pada periode tersebut menarik untuk diamati serta data tersedia pada tahun tersebut.

[27] Istijanto, *Aplikasi Praktis Riset Pemasaran: Cara Praktis Meneliti Konsumen dan Pesaing*, (Jakarta: Gramedia Pustaka Utama, 2009), p.38
[28] Damodar N. Gujarati, *Basic Econometrics Edisi Ke-4*, (New York: McGraw-Hill Inc, 2004), p.636

Data sekunder yaitu data yang bukan diusahakan sendiri pengumpulannya oleh peneliti, misalnya diambil dari Badan Pusat Statistik, dokumen Perusahaan atau organisasi, surat kabar dan majalah, ataupun publikasi lainnya[29] Periode data yang digunakan adalah data tahun 2011 – 2015 untuk masing-masing kabupaten/kota di Jawa Timur. Data yang diperlukan adalah:

1. Data jumlah penduduk daerah untuk masing-masing kabupaten/kota di Provinsi Jawa Timur tahun 2011 – 2015.
2. Data laju Produk Domestik Regional Bruto atas dasar harga konstan untuk masing-masing kabupaten/kota Jawa Timur tahun 2011 – 2015.
3. Data pendidikan yang diproksi dengan tingkat pendidikan untuk masing-masing kabupaten/kota di Jawa Timur tahun 2011- 2015.
4. Data penduduk miskin untuk masing-masing kabupaten/kota di Jawa Timur tahun 2011 – 2015.

Adapun sumber data tersebut diatas diperoleh dari:

1. Data jumlah penduduk daerah untuk masing-masing kabupaten/kota di Provinsi Jawa Timur tahun 2011 – 2015. Diperoleh dari Badan Pusat Statistik dalam terbitan "Jawa Timur Dalam Angka" dalam berbagai tahun terbitan.
2. Data laju Produk Domestik Regional Bruto atas dasar harga konstan untuk masing-masing kabupaten/kota Jawa Timur tahun 2011 – 2015. Diperoleh dari Badan Pusat Statistik dalam terbitan "Produk Domestik Regional Bruto Kabupaten/Kota Di Indonesia 2011-2015".
3. Data pendidikan yang diproksi dengan tingkat pendidikan untuk masing-masing kabupaten/kota di Jawa Timur tahun 2011- 2015. Diperoleh dari Badan Pusat Statistik

[29] Marzuki, *Metodologi Riset*, Ekonisia Kampus Fakultas Ekonomi UII, Yogyakarta, 2005

dalam terbitan “Laporan Eksekutif Pendidikan Provinsi Jawa Timur” terbitan tahun 2011 - 2015

4. Data penduduk miskin untuk masing-masing kabupaten/kota di Jawa Timur tahun 2011 – 2015. Diperoleh dari Badan Pusat Statistik dalam terbitan “Data dan Informasi Kemiskinan Kebupaten Kota” dalam berbagai tahun terbitan.

E. Metode Pengumpulan Data

Metode pengumpulan data adalah cara atau langkah yang digunakan oleh peneliti untuk mengumpulkan data yang dibutuhkan dalam melakukan penelitian. Metode pengumpulan data yang digunakan dalam penelitian ini adalah metode *ekspos facto*. *Ekspos facto* adalah pencarian empiris yang sistematis Dimana peneliti tidak dapat mengendalikan variabel bebasnya Karena peristiwa ini telah terjadi atau sifatnya tidak dapat dimanipulasi. Cara menerapkan metode penelitian ini dengan menganalisis peritiwa-peristiwa yang terjadi dari tahun-tahun sebelumnya untuk mengetahui faktor-faktor yang dapat menimbulkan kejadian tersebut[30]

Metode ini bermanfaat untuk mencari dan menggambarkan hubungan antara dua variabel atau lebih serta mengukur seberapa besar hubungan antar variabel yang dipilih untuk diteliti. Metode ini dipilih karena sesuai untuk mendapatkan informasi yang bersangkutan dengan status gejala saat peneltian dilakukan.

F. Teknik Analisis Data

1. Model Estimasi Regresi Data Panel

Model estimasi yang digunakan dalam penelitian ini adalah model estimasi regresi dengan menggunakan data panel. Data panel merupakan kombinasi dari data *cross-section dan time series*. Sehingga, secara umum persamaan regresi dengan menggunakan data panel dapat dituliskan sebagai berikut:

$$Y_{it} = \beta_1 + \beta_2.X_{2it} + \beta_3.X_{3it} + \beta_4.X_{4it} + \mu_{it}$$

[30] Husein Umar, *Metode Penelitian untuk Skripsi dan Tesis Bisnis Edisi Ke-2*, (Jakarta: PT. Raja Grafindo Persada, 2009), p. 28

Dimana Y_i merupakan variabel dependen yang nilainya dipengaruhi oleh tiga variabel *expalanatory* yang dinoatsikan dengan $X_{2it}, X_{3it}, X_{4it}$. Penggunaan notas i pada variabel dependen dan variabel *explanatory* menunjukkan adanya nilai tersendiri dari setiap unit *cross-section*. Sedangkan notasi t menunjukkan adanya nilai tersendiri dari setiap unit *time series*. $\beta_1, \beta_{2,}\, \beta_3, dan\, \beta_4$ merupakan parameter tetap yang nilainya tidak diketahui. Untuk memperoleh nilai dari keempat parameter tersebut perlu dilakukan perhitungan terlebih dahulu berdasarkan data-data yang diperoleh dari variabel dependen dan variabel *explanatory*. Dalam literature istilah yang digunakan untuk menyebut β_1 adalah *intercept* dan $\beta_2, \beta_3,$ dan β_4 adalah *slope coefficient*. Besarnya nilai *intercept* sama dengan Y_{it} ketika X_{it} sama dengan nol. Selain itu, variabel dependent juga bisa dipengaruhi oleh variabel lain yang nilainya tidak diketahui positif atau negatif dan berada di luar sistem. Variabel tersebut secara teknis dinamakan sebagai dengan *sthocastic disturbance* atau s*tochastic error term.*[31] Pada persamaan tersebut, variabel semacam ini dinotasikan dengan μ_{it}, yang tidak diketahui nilainya dan merupakan variabel acak yang membawa pengaruh positif atau negarif terhadap Y_{it}.

Berdasarkan persamaan di atas dapat dikatakan bahwa terdapat dua komponen yang dapat mempengaruhi besarnya nilai Y_{it}. Pertama, $\beta_1 + \beta_2.X_{2it} + \beta_3.X_{3it} + \beta_4.X_{4it}$, yang berarti besarnya nilai Y_{it} ditentukan oleh X_{it}. Kompenen pertama ini disebut sebagai komponen sistemtik atau deterministik. Kedua, μ_{it} yang merupakan variabel acak yang bersifat *disturbance (pengganggu)* dan nilainya pun tidak diketahui positif atau negatif. Selain itu variabel ini berada di luar sistem persamaan dan tidak dapat diteilti atau diobservasi.

Menurut Gujarati terdapat beberapa model yang digunakan untuk mengmodel estimasi regressi data panel. Alat untuk mengestimasi tersebut didasarkan pada asumsi berdasarkan

[31] Damodar N. Gujarati , op. cit, hal 44

intercept, *slope coefficient,* dan *error term*. Sehingga diperoleh beberapa kemungkinan di antaranya adalah:

1. Diasumsikan bahwa *intercept* dan *slope coefficient* konstan antar *time series* dan *cross-section*, serta *error term* meliputi perbedaan baik dalam waktu *time series* dan *cross-section*.
2. Diasumsikan bahwa *slope coefficient* konstan, akan tetapi *intercept* berbeda untuk setiap *cross-section*.
3. Diasumsikan bahwa *slope coefficient* konstan, akan tetapi intercept berbeda untuk setiao *cross-section* antar waktu.
4. Diasumsikan bahwa semua koefisien baik *intercept* dan *slope coefficient* berbeda untuk setiap *cross-section*.
5. Diasumsikan bahwa semua koefisien baik *intercept* dan *slope coefficient* berbeda untuk setiap *cross-section* antar *time series*.

Untuk mengmodel estimasi regresi dengan menggunakan data panel dapat dilakukan dengan tiga pendekatan, antara lain:

a. Model *Common Effects*

Metode paling sederhana yang digunakan untuk mengmodel estimasi regresi dengan menggunakan data panel adalah model *common effect*s. Pada dasarnya model *common effect*s sama dengan model estimasi *Ordinary Least Square* (OLS), yaitu estimasi yang dilakukan dengan menggunakan pendekatan kuadrat terkecil. Namn data yang digunakan bukan data *time series* atau *cross -section* saja, melainkan data panel yang diterapkan dalam bentuk *pooled* (kombinasi antara *cross-section* dan *time series*). Pada model estimasi regresi data panel ini,

semua koefisien diasumsikan konstan, baik itu *intercept* ataupun *slope coefficient*-nya pada setiap unit *cross section* yang dijadikan sampel. Adapun persamaan regresi dalam model *common effects* dapat ditulis sebagai berikut:

$$Y_{it} = \beta_1 + \beta_2.X_{2it} + \beta_3.X_{3it} + \beta_4.X_{4it} + \mu_{it}$$

Dimana i menunjukkan *cross-section* (individu) dan t menunjukkan periode waktunya. Dengan asumsi komponen *error* dalam pengolahan kuadrat terkecil biasa, proses estimasi secara terpisah untuk setiap unit *cross-section* dapat dilakukan.

b. Model *Fixed Effects*

Model estimasi regresi data panel ini memiliki asumsi bahwa *intercept* berbeda dari setiap *cross-section* dan konstan dari setiap *time series*. Sedangkan *coefficient slope*-nya konstan dari setiap *cross-section* dan *time series*. Untuk menjelaskan asumsi tersebut kita dapat menuliskan model sebagai berikut:

$$Y_{it} = \beta_{1i} + \beta_2.X_{2it} + \beta_3.X_{3it} + \beta_4.X_{4it} + \mu_{it}$$

Untuk *intercept* ditambahkan dengan notasi i untuk menggambarkan bahwa nilai *interept* dari setiap *cross-scetion* berbeda-beda. Perbedaan tersebut dapat mengacu pada faktor-faktor lain yang mempengaruhi besarnya nilai dari Y_{it} ketika variabel *expalanatory* sama dengan nol. Sebagai contoh, di Indonesia penerimaan pajak bumi dan bangunan (PBB) bersifat otonom, yang besarnya tidak tergantung pada fluktuasi pendapatan. Namun hal ini dapat berbeda bagi negara lain. Dalam beberapa literatur model estimasi ini dikenal sebagai model *fixed effects*. Istilah *fixed effect*s mengacu pada fakta bahwa, meskipun *intercept* berbeda pada setiap *cross-section*, namun tidak bagi *time series*. Artinya *time series* dalam model ini bersifat *invariant*. Disisi lain, model

fixed effects berasumsi bahwa *slope coefficient* tidaklah berbeda pada setiap cross-section atau *time series*. Bagaimana sebenarnya kita dapat membiarkan *intercept* berbeda-beda dari setiap *cross-section*? Kita dapat dengan mudah melakukan hal tersebut dengan menggunakan teknik variabel *dummy*. Penggunaan variabel *dummy* dalam mengmodel estimasi regresi data panel ini meneyebabkan model ini sering disebut sebagai *Least Square Dummy Variable* (LSDV). Dengan penggunaan variabel *dummy* dalam mengmodel estimasi regresi ini, kita dapat menuliskan persamaan regresi sebagai berikut:

$$Y_{it} = \alpha_1 + a_{2.}D_{2i} + a_3 D_{3i} + a_{4.}D_{4i} + \beta_2.X_{2it} + \beta_3.X_{3it} + \beta_4.X_{4it} + \mu_{it}$$

Dimana variabel *dummy* pada persamaan tersebut dinotasikan dengan *D* dan tambahan notasi i menggambarkan variasi nilai dari setiap *cross-section*. Nilai untuk variabel *dummy* berupa angka 0 dan 1. Angka 0 menggambarkan mengindikasikan apa yang tidak dimiliki dari suatu atribut. Sedangkan angka 1 mengindikasikan apa yang dimiliki dari suatu atribut.

c. Model *Random Effect*

Keputusan untuk memasukan variabel *dummy* dalam model *fixed effects* memiliki konsekuensi berkurangnya *degree of freedom* yang akhirnya dapat mengurangi efisiensi dari parameter yang diestimasi. Oleh karena itu, dalam model data panel dikenal pendekatan yang ketiga yaitu model *random effects*.[32] model *random effects* disebut juga dengan model *error component*. karena di dalam model ini parameter yang berbeda antar unit *cross-section* maupun antar *timer series* dimasukkan ke dalam *error term*. Dengan menggunakan model *random effects*, maka dapat menghemat pemakaian derajat kebebasan dan tidak mengurangi jumlahnya seperti yang dilakukan oleh mode *fixed effects*. Hal ini berimplikasi pada parameter yang merupakan

[32] B.H. Baltagi, *Econometerics Analysis of Panel Data Edisi Ke-3*, Chiester: John Wiley & Sons Ltd, 2005

hasil estimasi akan menjadi semakin efisien dan model yang dihasilkan semakin baik. Untuk persamaan regresi dari model *random effects* dapat dimulai dari persamaan berikut:

$$Y_{it} = \beta_{1i} + \beta_2.X_{2it} + \beta_3.X_{3it} + \beta_4.X_{4it} + \mu_{it}$$

Dengan memperlakukan β_{1i} sebagai *fixed*, kemudian diasumsikan bahwa *intercept* memiliki nilai rata-rata sebesar β_1. Sedangkan nilai rata-rata dari setiap *cross-section* dapat dituliskan sebagai barikut:

$$\beta_{1i} = \beta_1 + \varepsilon_i \qquad i = 1,2,\dots,N$$

Dimana ε_i adalah *random error term* dengan nilai rata-rata sama dengan nol dan merupakan variasi dari σ_ε^2. Secara esensial, dapat dikatakan bahwa semua *cross-section* memiliki nilai rata-rata yang sama untuk *intercept*, yaitu sebesar β_1. Sedangkan perbedaan nilai *itercept* dari setiap unit *cross-section* direfleksikan dalam *errot term* ε_i. Apabila dua sebelumnya persamaan disubstitusikan, maka akan diperoleh persamaan regresi sebagai berikut:

$$\begin{aligned} Y_{it} &= \beta_1 + \beta_2.X_{2it} + \beta_3.X_{3it} + \beta_4.X_{4it} + \mu_{it} + \varepsilon_i \\ &= \beta_1 + \beta_2.X_{2it} + \beta_3.X_{3it} + \beta_4.X_{4it} + \omega_{it} \\ \omega_{it} &= \mu_{it} + \varepsilon_i \end{aligned}$$

Berdasarkan persamaan di atas *error term* kini dinotasikan dengan ω_{it}, yang terdiri dari dua komponen, yaitu ε_i, yang merupakan *cross-section error component*, artinya pada komponen ε_i ini terdapat perbedaan nilai *intercept* dari setiap unit *cross-section* yang direfleksikan oleh komoponen ε_i. Sedangkan komponen μ_{it} merupakan kombinasi antara *time series* dan *cross-section* dari *error component*, artinya terdapat perbedaan nilai *intercept* dari setiap unit *time series* dan *cross-section* yang direfleksikan oleh komponen μ_{it}.

Perbedaan utama antara model *fixed effects* dan model *random effects* adalah pada perlakuan *intercept*. Pada model *fixed effects* setiap unit cross-*section* memiliki nilai intercept tersendiri yang fixed. Sedangkan pada model *random effects* setiap unit *cross-section* memiliki nilai *intercept* tersendiri yang direfleksikan oleh *error term* ε_i. Sedangkan nilai *intercept* rata-rata dari seluruh *cross-section* direfleksikan oleh β_1.

Menurut Gujarati[33] dasar pemilihan antara model *fixed effects* dan *random effects* adalah sebagai berikut:

b. Jika *T* (jumlah data *time series*) besar dan *N* (jumlah data dari *cross-section*) kecil, maka akan menunjukkan bahwa tidak ada perbedaan nilai parameter yang diestimasi oleh model *fixed effects* dan *random effects*. Pemilihan model terbaik dilakukan berdasarkan kemudahan penghitungan. Maka dalam hal ini, model *fixed effects* lebih baik daripada *random effects*.
c. Ketika *N* besar dan *T* kecil, estimasi yang diperoleh dari kedua metode akan memiliki perbedaan yang signifikan. Jadi, apabila kita meyakini bahwa unit *cross-section* yang kita pilih dalam penelitian diambil secara acak maka model *random effects* harus digunakan. Sebaliknya, apabila kita meyakini bahwa unit *cross-section* yang kita pilih dalam penelitian tidak diambil secara acak maka kita harus menggunakan model *fixed effects*.
d. Jika *error component* ε_i berkorelasi dengan variabel independen, maka parameter yang diperoleh dengan model *random effects* akan bias sementara parameter yang diperoleh dengan menggunakan model *fixed effects* tidak bias.

[33] Damodar N. Gujarati, 2004, op cit, hal 650 - 651

e. Apabila *N* besar dan *T* kecil, dan apabila asumsi yang mendasari *random effects* dapat terpenuhi, maka model *random effects* akan lebih efisien dari model *fixed effects*.

Untuk memilih model mana yang paling tepat digunakan untuk pengolahan data panel, maka terdapat beberapa pengujian yang dapat dilakukan, antara lain:

a. *Chow Test*

Chow Test adalah pengujian untuk memilih apakah model yang digunakan model *common effects* atau *fixed effects*. Dalam pengujian ini dilakukan dengan hipotesis sebagai berikut:[34]

H0: $a_1 = a_2 = \cdots = a_n = 0$ (efek dari unit *cross-section* secara keseluruhan tidak berarti)

Ha: $a_i \neq 0; i = 1,2, \ldots, n$ (efek dari salah satu atau lebih unit *cross-section* berarti)

Statistik uji yang digunakan merupakan uji F, yaitu:

$$F - statistik = \frac{[RRSS - URSS]/(n-1)}{URSS/(nT - n - k)}$$

Keterangan:

n = Jumlah unit *cross-section*

T = Jumlah perode waktu (*time series*)

K = Jumlah variabel independen

RSS = restricted residual sums of squares yang berasal dari model *common effects*

URSS = unrestricted residual sums of squares yang berasal dari model *fixed effects*

Sedangkn F-tabel diperoleh dari:

$$F - tabel = \{a: df(n-1, nt - n - k)$$

Keterangan:

[34] B.H. Baltagi, *Econometerics Analysis of Panel Data Edisi Ke-3*, Chiester: John Wiley & Sons Ltd, 2005

α : Tingkat signifikansi yang dipakai (alfa)

n : Jumlah unit *cross-section*

nt : Jumlah unit *cross-section* dikali jumlah *time series*

k : Jumlah variabel independen

Dasar penolakan terhadap hipotesis di atas adalah dengan membandingkan perhitungan F-statistik dan F-tabel. Apa bila hasil F-statistik lebih besar dari F-tabel, maka H_0 ditolak, yang berarti model *fixed effects* yang paling baik untuk digunakan dalam mengmodel estimasi regresi data panel. Sebailnya, apabila F-statistik lebih kecil dari F-tabel, maka H_0 diterima, yang berarti model *common effects* yang paling baik untuk digunakan dalam mengmodel estimasi regresi data panel.[35] Selain dengan membadingkatn F-tabel dan F-statistik, dapat juga dilakukan dengan membandingkan antar nilai probabilitas dari F-statistik dan *alpha* (0,05). Apabila nilai probabilitas dari F-statistik > 0,05, maka H_0 diterima yang artinya model *common effects* yang paling baik untuk digunakan. Jika sebaliknya, maka H_0 ditolak yang artinya model *fixed effects* yang paling baik digunakan.

b. *Hausman Test*

Hausman Test adalah pengujian statistik sebagai dasar pertimbangan kita dalam memilih apakah menggunakan *model fixed effects* atau *random effects*. Uji ini bekerja dengan menguji apalah terdapat hubungan antara *error component* demam satu atau lebih variabel independen dalam suatu model. Hipotesis awalnya adalah tidak terdapat hubungan antara *error component*

[35] Wing Wahyu Winarno, *Analisis Ekonometrika Dan Statistika Dengan Eviews. Edisi Kedua*, Yogyakarta: UPP STIM YKPN, 2009

dengan variabel independen. Menurut Baltagi hipotesis dari uji Hausman adalah sebagai berikut:[36]

H0: Korelasi (X_{it}, μ_{it}) = 0 (efek *cross-section* tidak berhubungan dengan *error component*)

Ha: Korelasi (X_{it}, μ_{it}) ≠ 0 (efek *cross-section* berhubungan dengan *error component*)

Statistik uji yang digunakan adalah uji *chi square*. Jika nilai *chi square*-statistik > *chi square*-tabel (*a,k*) atau nilia *p-value* kurang dari taraf signifikansi yang ditentukan, maaka hipotesis awal (H_0) ditolak sehingga model yang terpilih adalah model *fixed effects*. Hal ini dilakukan untuk mengetahui apakah terdapat efek random di dalam data panel.[37]

Dalam perhitungan uji Hausman diperlukan asumsi bahwa banyaknya kategori *cross-section* lebih besar dibandingkan jumlah variabel independen (termasuk konstanta) dalam model. Lebih lanjut, dalam estimasi uji Hausman diperlukan estimasi variansi *cross-section* yang positif, yang tidak selalu dapat dipenuhi oleh model. Apabila kondisi-kondisi ini tidak dapat dipenuhi, maka hanya dapt digunakan model *fixed effects*.

c. *Lagrange Multiplier Test*

Lagrange Multiplier Test digunakan untuk menguji model apakah yang terbaik untuk digunakan dalam penelitian, yaitu untuk menguji model *common effects* dan model *random effects*. Hipotesis yang digunakan dalam uji Lagrange Multiplier adalah sebagai berikut:

H_0 = Model *Random Effects*

[36] B.H. Baltagi, op cit, 310
[37] Dedi Rosad, *Analisis Ekonometrika dan Runtun Waktu Terapan dengan R*, (Yogyakarta: Andi Offset i, 2011), hal 264

H_a = Model *Common Effects*

Untuk dapat menentukan jawaban dari hipotesis di atas, maka diperlukanlah perhitungaan LM-statistik nya. Perhitungan LM-statistik dapat dituliskan dengan rumus sebagai berikut:

$$\text{LM} - \text{statistik} = \frac{nT}{2(T-1)}\left[\frac{T^2 \sum \bar{e}^2}{\sum e^2} - 1\right]^2$$

Keterangan:

n = jumlah *cross-section*

T = jumlah *time-series*

$\sum \bar{e}^2$ = jumlah rata-rata kuadrat residual

$\sum e^2$ = jumlah residual kuadrat

Nilai LM-statistik akan dibandingkan dengan nilai *Chi Square*-tabel dengan derajat kebebasan (*degree of freedom*) sebanyak jumlah variabel independen (bebas) dan alpha atau tingkat signifikansi sebesar 5%. Apabila nilai LM-statistik > *Chi Square*-tabel, maka H_0 di terima, yang artinya model yang dipilih adalah model *random effects*, jika sebaliknya maka H_0 ditolak, yang artinya model yang dipilih adalah model *common effects*.

2. **Pengujian Asumsi Klasik**

Kelebihan penelitian menggunakan data panel adalah data yang digunakan menjadi lebih informatif, variabilitasnya lebih besar, kolineariti yang lebih rendah diantara variabel dan banyak derajat bebas (degree of freedom) dan lebih efisien. Panel data dapat mendeteksi dan mengukur dampak dengan lebih baik dimana hal ini tidak bisa dilakukan dengan metode cross section

maupun time series. Panel data memungkinkan mempelajari lebih kompleks mengenai perilaku yang ada dalam model sehingga pengujian data panel tidak memerlukan uji asumsi klasik. Dengan keunggulan regresi data panel maka implikasinya tidak harus dilakukannya pengujian asumsi klasik dalam model data panel[38]

a. Deteksi Heterokedastisitas

Deteksi heterokedasitas bertujuan untuk menguji apakah dalam model regresi terjadi ketidaksamaan varians dari satu pengamatan ke pengamatan lainnya. Model regresi dikatakan baik apabila tidak terjadi heterokedasitas, artinya adanya ketetapan atau konstan antara varians dari satu pengamatan ke pengamatan lainnya (Homokedastisitas). Hipotesis yang digunakan untuk mendeteksi heterokedastistias berdasarkan uji *White* adalah sebagai berikut:

H_0 = (struktur *variance-covariance residual* homokedastik)

H_a = (struktur *variance-covariance* residual heterokedastik)

Berdasarkan hipotesis tersebut, maka kriteria pengambilan kesimpulan yakni jika nilai probabilitas (*p-value*) dari Chi Square > 0,05, maka H0 diterima, artinya varians error bersifat homokedastik. Jika sebaliknya, maka H0 ditolak, yang berarti varians error bersifat heterokedastik.

b. Deteksi Multikolinieritas

Deteksi multikolineritas bertujuan untuk mendeteksi apakah antara variabel independen (variabel bebas) terdapat korelasi. Sehingga sulit untuk memisahkan pengaruh antara variabel-variabel itu secara individu terhadap varaibel terikat. Model regresi dikatakan baik apabila tidak

[38] *Damodar N. Gujarati,* Essentials of Econometrics. McGraw-Hill, New York, 1992

ada korelasi antar variabel independen. Keberadaan multikolinieritas menyebabkan standar error cenderung semakin besar. Meningkatnya tingkat korelasi antar variabel, menyebabkan standar error semakin sensitif terhadap perubahan data.

Menurut Gujarati[39] tingginya koefisien korelasi antar variabel bebas merupakan salah satu indikator dari adanya multikolinearitas antar variabel bebas. Jika terjadi koefisien korelasi lebih dari 0,80 maka dapat dipastikan terdapat multikolinearitas antar variabel bebas.

3 Pengujian Kriteria Statistik

Gujarati menyatakan bahwa uji signifikansi merupakan prosedur yang digunakan untuk menguji kebenaran atau kesalahan dari hasil hipotesis nol dari sampel. Ide dasar yang melatarbelakangi pengujian signifikansi adalah uji statistik (estimator) dari distribusi sampel dari suatu statistik dibawah hipotesis nol. Keputusan untuk mengolah Ho dibuat berdasarkan nilai uji statistik yang diperoleh dari data yang ada[40].

Uji statistik terdiri dari pengujian koefisien regresi parsial (uji t), pengujian koefisien regresi secara bersama-sama (uji F), dan pengujian koefisien determinasi (uji-R^2).

[39] Damodar N. Gujarati, op. cit, hal 359
[40] Damodar N. Gujarati, *Basic Econometric*, New York: McGraw-Hill Inc, 2004

a. Uji Parsial (Uji – t)

Uji parsial digunakan untuk mengetahui pengaruh masing-masing variabel independen terhadap variabel dependen secara parsial.[41] Pengujian dapat dilakukan dengan menyusun hipotesis sebagai berikut:

1. H0 : b1 = 0 tidak ada pengaruh antara variabel PDRB dengan kemiskinan.

   H1 : b1 < 0 ada pengaruh negatif antara variabel PDRB dengan kemiskinan.
2. H0 : b2 = 0 tidak ada pengaruh antara variabel Tingkat Pendidikan dengan kemiskinan.

   H1 : b2 < 0 ada pengaruh negatif antara variabel Tingkat Pendidikan dengan kemiskinan.
3. H0 : b3 = 0 tidak ada pengaruh antara variabel tingkat Jumlah Penduduk dengan kemiskinan.

   H1 : b3 > 0 ada pengaruh positif antara variabel tingkat Jumlah Penduduk dengan kemiskinan.

Dasar pengambilan keputusan, apabila angka probabilitas signifikansi > 0.05 maka H0 diterima dan H_a ditolak. Artinya variaebel independen secara parsial tidak berpengaruh terhadap variabel dependen. Namun, apabila angka probabilitas signifikansi < 0,05 maka H_0 ditolak dan H_a diterima. Artinya variabel independen secara parsial berpengaruh terhadap variabel dependen. Dasar pengambilan keputusan juga dapat dilakukan dengan membandingkan nilai t-statistik

[41] Imam Ghazali, *Op.cit*, p.98

dengan t-tabel. H0 diterima jika t-tabel > t-statistik dan ditolak jika t-tabel < t-statistik. Nilai t-statistik dapat dicari dengan menggunakan rumus sebagai berikut

$$t - statistik = \frac{\beta_i - \beta_0}{SE(\beta_i)}$$

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

0 = Hipotesis awal = nol

Keteranagan:

β_i = nilai parameter (*intercept* dan *slope coefficient*)

β_0 = Hipotesis awal yang diuji nilainya sama dengan nol

SE = Standar eror setiap parameter (*intercept* dan *slope coefficient*)

b. Uji Signifikansi Simultan (Uji F)

Untuk menguji keberartian regresi dalam penelitian ini digunakan Uji statistik F dengan Tabel Anova. Uji statistik F pada umumnya menunjukkan apakah semua variabel independen atau bebas yang dimasukkan ke dalam model mempunyai pengaruh bersama-sama terhadap variabel dependen atau terikat. Pengujian dapat dilakukan dengan menyusun hipotesis terlebih dahulu sebagai berikut:

H_0 : ß$_2$ = ß$_3$ = ß$_4$ = 0

H_a : ß$_2$ ≠ ß$_3$ ≠ ß$_4$ ≠ 0

Kriteria pengujian, apabila nilai signifikansi <0,05 maka H_0 ditolak dan H_a diterima. Artinya semua variabel independe atau bebas secara simultan berpengaruh terhadap variabel dependen atau terikat. Begitu juga sebaliknya, apabila nilai signifikansi >0,05 maka H_0 diterima dan H_a ditolak. Artinya semua variabel independen atau bebeas secara simultan tidak

berpengaruh terhadap variabel dependen atau terikat. Selain itu dapat digunakan kriteria lain pada pengujian keberartian regeresi, yaitu apabila F-tabel > F-statistik maka H_0 diterima dan apabila F-tabel < F-statistik maka H_0 ditolak. Nilai dari F-statistik datang dicari dengan menggunakan rumus sebagai berikut:

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

Keterangan:

R^2 = koefisien determinasi (residual)

k = jumlah variabel independen ditambah intercept dari suatu model persamaan

n = jumlah sampel

c. Koefisien Determinasi (R^2)

Koefisien determinasi (*Goodness of fit*) dilakukan untuk mengukur seberapa jauh kemampuan model dalam menerangkan variasi variabel dependen.[42] Nilai R^2 menunjukan seberapa baik model yang dibuat mendekati fenomena dependen seharusnya. Rumus menghitungnya adalah dengan terlebih dahulu mencari nilai R atau koefisien korelasi:

$$R^2 = \frac{\beta_1 \sum X_1 Y + \beta_2 \sum X_2 Y + \beta_3 \sum X_3 Y}{\sum Y^2}$$

Nilai dari koefisien determinan adalah 0 sampai 1. Jika $R^2 = 0$, hal tersebut menunjukkan variasi dari variabel terikat tidak dapat diterangkan oleh variabel bebas. Namun jika $R^2 = 1$, maka variasi dari variabel terikat dapat dijelaskan oleh variabel bebas.

[42] Ghozali, Imam, op. cit, hal 97

Kelemahan mendasar pada koefisien determinasi yaitu bias terhadap jumlah variabel independen yang masuk ke dalam model. Setiap penambahan satu variabel independen yang belum tentu berpengaruh signifikan atau tidak terhadap variabel dependen, maka nilai R^2 pasti akan meningkat. Oleh sebab itu, digunakan nilai *adjusted* R^2 yang dapat naik turun apabila ada penambahan variabel independen ke dalam model.