

BAB III

METODOLOGI PENELITIAN

A. Tujuan Penelitian

Berdasarkan masalah-masalah yang peneliti rumuskan, maka tujuan penelitian ini, yaitu:

1. Untuk mengetahui adanya pengaruh pendidikan terhadap produktivitas tenaga kerja di lima negara ASEAN (Indonesia, Thailand, Filipina, Laos, Kamboja) pada tahun 2006-2015.
2. Untuk mengetahui adanya pengaruh kesehatan terhadap produktivitas tenaga kerja di lima negara ASEAN (Indonesia, Thailand, Filipina, Laos, Kamboja) pada tahun 2006-2015.
3. Untuk mengetahui adanya pengaruh teknologi terhadap produktivitas tenaga kerja di lima negara ASEAN (Indonesia, Thailand, Filipina, Laos, Kamboja) pada tahun 2006-2015.
4. Untuk mengetahui adanya pengaruh upah terhadap produktivitas tenaga kerja di lima negara ASEAN (Indonesia, Thailand, Filipina, Laos, Kamboja) pada tahun 2006-2015.
5. Untuk mengetahui adanya pengaruh pendidikan, kesehatan, teknologi, dan upah terhadap produktivitas tenaga kerja negara ASEAN dalam 10 tahun, yaitu dari tahun 2006-2015 untuk negara ASEAN (Indonesia, Thailand, Filipina, Laos, Kamboja)

B. Objek dan Ruang Lingkup Penelitian

Objek penelitian ini adalah pendidikan, kesehatan, teknologi, dan upah terhadap produktivitas tenaga kerja negara di lima negara ASEAN Indonesia, Thailand, Filipina, Kamboja dan Laos. Data yang digunakan dalam penelitian ini adalah data sekunder yang tersedia di Bank Dunia, UNESCO, ILO, APO, Badan Pusat Statistik, dan lembaga statistik lainnya. Setiap variabel dari masing-masing negara digunakan data dalam jangka waktu 10 tahun, yaitu dari tahun 2006-2015 untuk negara ASEAN (Indonesia, Thailand, Filipina, Kamboja dan Laos)

Ruang lingkup penelitian ini adalah untuk mengkaji pengaruh pendidikan, kesehatan, teknologi, dan upah terhadap produktivitas tenaga kerja negara di 10 tahun, yaitu dari tahun 2006-2015 untuk negara ASEAN (Indonesia, Thailand, Filipina, Kamboja dan Laos) karena merupakan waktu yang paling efektif bagi peneliti untuk melaksanakan penelitian, sehingga peneliti dapat fokus pada saat penelitian. Selain itu, peneliti juga memiliki keterbatasan waktu sebagaimana yang telah ditetapkan dalam jadwal akademik. Tenaga dan materi yang terbatas juga merupakan salah satu keterbatasan yang dimiliki oleh peneliti.

a) Operasionalisasi Variabel Penelitian

Variabel operasional penelitian ini diperlukan untuk memahami 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 komprehensif.

1. Pendidikan

a. Definisi Konseptual

Pendidikan adalah suatu bentuk kesadaran dan terencana untuk mengembangkan potensi dirinya memiliki spiritual keagamaan, pengendalian diri, kepribadian, kecerdasan, akhlak mulia, serta keterampilan yang diperlukan dirinya, masyarakat serta bangsa dan negara yang dibutuhkan untuk mencapai peradaban yang tertinggi.

b. Definisi Operasional

Peningkatan produktivitas tenaga kerja yang bertumpu pada pendidikan pada dasarnya bertujuan untuk meningkatkan kemampuan kerja manusia dan untuk meningkatkan taraf hidup manusianya. Pendidikan suatu negara dapat diukur dengan rata rata lama sekolah, kemudian dibagi dengan total produktivitas tenaga kerja rata rata lama sekolah dari situs resmi Bank Dunia. Data rata rata lama sekolah dalam penelitian ini memiliki jangka waktu 10 tahun, yaitu dari tahun 2006-2015 untuk negara ASEAN (Indonesia, Thailand, Filipina, Kamboja dan Laos)

2. Kesehatan

a. Definisi Konseptual

Kesehatan adalah suatu proses perubahan pada diri manusia yang ada hubungannya dengan tercapainya tujuan kesehatan Nyaswander (2002 : 5).

b. Definisi Operasional

Kesehatan merupakan modal manusia yang dapat berhubungan dengan produktivitas dalam aspek mikro maupun makro. Serta banyak

faktor yang mempengaruhi kesehatan diantaranya faktor perilaku dan faktor lingkungan. Kedua faktor tersebut memiliki kontribusi besar . Kesehatan suatu negara dapat diukur dengan menjumlahkan angka harapan hidup, kemudian dibagi dengan total produktivitas tenaga kerja. Angka harapan hidup dari situs resmi Bank Dunia. Data Angka harapan hidup dalam penelitian ini memiliki jangka waktu 10 tahun, yaitu dari tahun 2006-2015 untuk negara ASEAN (Indonesia, Thailand, Filipina, Kamboja dan Laos)

3. Teknologi

a. Definisi Konseptual

Teknologi adalah sebuah proses, metode dan pengetahuan yang diterapkan untuk melaksanakan suatu pekerjaan Schaum (2004).

a. Definisi Operasional

Teknologi merupakan sumber daya yang dapat menciptakan pertumbuhan dari berbagai aspek melalui peningkatan produktivitas. Teknologi suatu negara dapat diukur dengan menjumlahkan angka faktor total produktivitas ,kemudian dibagi dengan total produktivitas tenaga kerja. faktor total produktivitas diperoleh dari situs resmi Bank Dunia. Data faktor total produktivitas dalam penelitian ini memiliki jangka 10 tahun 2006-2015 di lima negara ASEAN Indonesia, Thailand, ,Filipina, Kamboja dan Laos

4. Upah

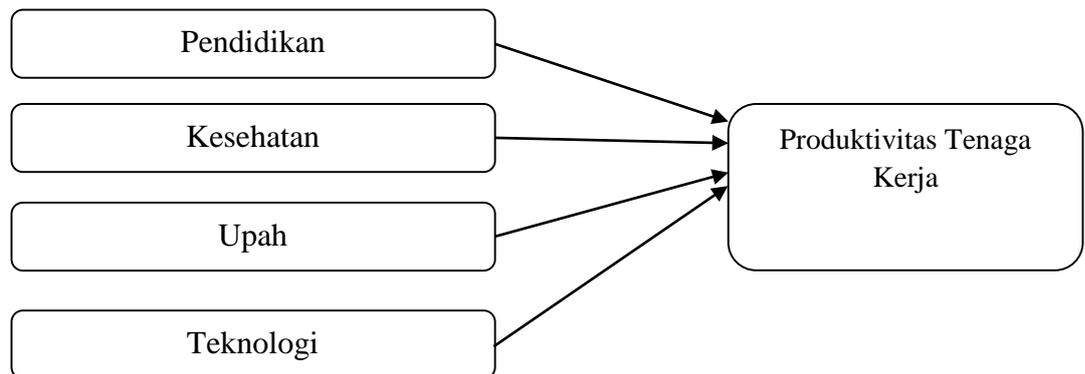
a. Definisi Konseptual

Upah adalah imbalan pembayaran atas jasa yang diberikan baik fisik maupun mental. Upah dapat dinilai dalam bentuk uang atau kesepakatan yang terjadi antara pihak perusahaan maupun tenaga kerja.

b. Definisi Operasional

Seorang dalam melakukan aktivitas tertentu terkhusus dalam bidang ekonomi akan berorientasi pada upah. Karena upah merupakan komponen penting yang mendukung kegiatan ekonomi berjalan dengan optimal. Upah suatu negara dapat diukur dengan menjumlahkan angka kelayakan hidup, kemudian dibagi dengan total produktivitas tenaga kerja. Data angka kelayakan hidup ekonomi diperoleh dari situs resmi Bank Dunia. Data tingkat kelayakan hidup dalam penelitian ini memiliki jangka waktu 10 tahun 2006-2015 di lima negara ASEAN Indonesia, Thailand, Filipina, Kamboja dan Laos

Konstelasi pengaruh antar variable dapat digambarkan sebagai berikut:



Keterangan:

Pendidikan : Variabel Independen X1

Kesehatan : Variabel Independen X2

Upah : Variabel Independen X3

Teknologi : Variabel Independen X4

Produktivitas tenaga kerja : Variabel Dependen Y

—————> : Arah Pengaruh

C. Jenis dan Sumber Data

Penelitian ini menggunakan data skunder dari setiap variabel, yaitu variabel Pendidikan, kesehatan, teknologi, upah dan produktivitas tenaga kerja . Data yang digunakan adalah data panel, yaitu kombinasi antara *cross section* dan *time series*. Data panel yang dijadikan sampel dalam penelitian ini terdiri dari lima negara ASEAN Indonesia, Thailand, Filipina, Kamboja dan Laos. Sedangkan rentang waktu yang digunakan dalam penelitian ini adalah 10 tahun, yaitu mulai dari tahun 2006-2015. Data yang digunakan dalam penelitian ini diperoleh dari berbagai sumber, yaitu situs resmi dari Badan Pusat Statistik, Bank Indonesia, Bank Dunia, UNESCO, Kementerian Keuangan, dan OECD, Serta situs-situs resmi lainnya.

D. 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 peristiwa-peristiwa yang terjadi dari tahun-tahun sebelumnya untuk mengetahui faktor-faktor yang dapat menimbulkan kejadian tersebut.

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 penelitian dilakukan.

E. Teknik Analisis Data

Analisis yang digunakan dalam penelitian ini adalah analisis regresi linear berganda dengan menggunakan data panel, yaitu analisis regresi yang digunakan untuk mengestimasi nilai dari variabel dependen yang dipengaruhi oleh beberapa variabel independen dengan menggunakan data panel. Langkah pertama yang dilakukan dalam teknik analisis regresi ini adalah menentukan model estimasi yang terbaik dengan melakukan beberapa uji, yaitu uji Hausman, uji Chow . Kemudian mendeteksi gejala asumsi klasik untuk mengetahui model estimasi yang digunakan dapat menjadi estimator yang baik atau tidak. Deteksi gejala asumsi yang dilakukan adalah deteksi normalitas, heteroskedastisitas, multikolinearitas, dan autokorelasi.

Selanjutnya, dilakukanlah uji hipotesis dengan menggunakan uji t dan uji F. Uji t dilakukan untuk mengetahui apakah terdapat pengaruh dari variabel independen terhadap variabel dependen secara parsial. Sedangkan uji F dilakukan untuk mengetahui apakah terdapat pengaruh dari variabel independen terhadap variabel dependen secara simultan. Terakhir, melakukan analisis koefisien determinasi (R^2). Langkah terakhir ini dilakukan untuk mengetahui seberapa besar kemampuan dari variabel independen dalam menjelaskan nilai dari variabel dependen. Analisis

regresi yang dilakukan dalam penelitian ini menggunakan bantuan aplikasi Eviews8 dan Mcs.Excel 2013.

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*. Menurut Gujarati dalam (Gujarati 2014) terdapat beberapa model yang digunakan untuk mengmodel estimasi regresi data panel. Alat untuk mengestimasi tersebut didasarkan pada asumsi berdasarkan *intercept*, *slope coefficient*, dan *error term*. Sehingga secara umum, diperoleh tiga model estimasi regresi data panel, yang terdiri dari model *common effects*, *fixed effects*, dan *random effects*. Berikut ini adalah penjelasan dari tiga model tersebut:

a. Model *Common Effects*

Model estimasi regresi dengan menggunakan data panel yang paling sederhana adalah model *common effects*. Pada dasarnya model *common effects* sama dengan model estimasi *Ordinary Least Square* (OLS). Namun 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 negara dalam penelitian ini. Adapun persamaan regresi dalam model *common effects* dapat ditulis sebagai berikut:

[3.2]

$$Prod_{it} = \beta_1 + \beta_2 \cdot Edu + \beta_3 \cdot hlt_{it} + \beta_4 \cdot Wage_{it} + \beta_4 \cdot Tech_{it} + \mu_{it}$$

Dimana i menunjukkan negara dan t menunjukkan periode waktunya. adalah Produktivitas tenaga kerja (Y), edu adalah pendidikan ($X1$), hlt adalah kesehatan ($X2$), $tech$ teknologi ($X3$), dan $wage$ upah ($X4$). Sedangkan μ_{it} adalah nilai residual, yaitu selisih antara nilai $Y_{observasi}$ dan $Y_{estimasi}$.

b. Model *Fixed Effects*

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

[3.3]

$$Prod_{it} = \beta_1 + \beta_2 \cdot Edu + \beta_3 \cdot hlt_{it} + \beta_4 \cdot Wage_{it} + \beta_4 \cdot Tech_{it} + \mu_{it}$$

Untuk *intercept* ditambahkan dengan notasi i untuk menggambarkan bahwa nilai *intercept* dari setiap negara berbeda-beda. Perbedaan tersebut dapat mengacu pada faktor-faktor lain yang mempengaruhi besarnya nilai dari variabel produktivitas tenaga kerja ketika variabel *explanatory* sama dengan nol. Dalam beberapa literatur model estimasi ini dikenal sebagai model *fixed effects*. Istilah *fixed effects* mengacu pada fakta bahwa, meskipun *intercept* berbeda pada setiap negara, namun

konstan dari setiap waktu. Selain itu, model *fixed effects* berasumsi bahwa *slope coefficient* tidaklah berbeda pada setiap negara dan waktu.

Membedakan nilai *intercept* dari setiap negara dalam suatu model, maka kita dapat menggunakan teknik variabel *dummy*. Dengan penggunaan variabel *dummy* dalam model estimasi regresi ini, kita dapat menuliskan persamaan regresi sebagai berikut:

[3.4]

$$Y_{it} = b_1 + a_1.INDO + a_2.THAI + a_3.PHIL + a_4.VIET + a_5.LAOS \\ + \beta_2.Edu + \beta_3.hlt_{it} + \beta_4.Wage_{it} + \beta_4.Tech_{it} + \mu_{it}$$

Dimana variabel *dummy* pada persamaan tersebut dinotasikan dengan D dan tambahan notasi i menggambarkan variasi nilai dari *intercept* dari setiap negara. Jumlah nilai *intercept* dari setiap negara yang di variabel *dummy*-kan adalah total negara yang diteliti (tujuh negara) dikurang satu, yaitu enam. Hal ini dilakukan agar kita dapat terhindar dari jebakan variabel *dummy*. Sehingga model estimasi yang digunakan terdapat multikolinearitas. Negara yang tidak di variabel *dummy*-kan menjadi komparasi dari nilai *intercept* negara-negara lain. Tentunya, peneliti bebas memilih negara mana yang dijadikan sebagai komparasi bagi negara-negara lain.

a. 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* dalam . model *random effects* disebut juga dengan model *error component*. karena di dalam model ini parameter yang berbeda antar negara maupun antar waktu dimasukkan ke dalam *error term* (residual).

Untuk persamaan regresi dari model *random effects* dapat dimulai dari persamaan berikut:

[3.5]

$$Prod_{it} = \beta_1 + \beta_2 \cdot Edu + \beta_3 \cdot hlt_{it} + \beta_4 \cdot Wage_{it} + \beta_4 \cdot Tech_{it} + \mu_{it}$$

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

[3.6]

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

Dimana ε_i adalah *random error term* dengan nilai rata-rata sama dengan nol dan merupakan nilai varians dari ε . Secara esensial, dapat dikatakan bahwa semua negara memiliki nilai rata-rata yang sama untuk *intercept*, yaitu sebesar β_1 . Sedangkan perbedaan nilai *intercept* dari setiap unit negara direfleksikan dalam *error term* ε_i . Apabila persamaan 3.5 dan 3.6 disubstitusikan, maka akan diperoleh persamaan regresi sebagai berikut:

[3.7]

$$Prod_{it} = \beta_1 + \beta_2 \cdot Edu + \beta_3 \cdot hlt_{it} + \beta_4 \cdot Wage_{it} + \beta_4 \cdot Tech_{it} + \mu_{it} + \varepsilon_i$$

$$= \beta_{1i} + \beta_2 \cdot Edu + \beta_3 \cdot hlt_{it} + \beta_4 \cdot Wage_{it} + \beta_4 \cdot Tech_{it} + \omega_{it}$$

$$\omega_{it} = \mu_{it} + \varepsilon_i$$

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 negara. Sedangkan komponen μ_{it} merupakan kombinasi antara *time series* dan *cross-section* dari residual, artinya terdapat perbedaan nilai residual 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* negara memiliki nilai *intercept* tersendiri yang fixed. Sedangkan pada model *random effects* setiap unit negara memiliki nilai *intercept* tersendiri yang masukkan ke dalam *error term* ε_i . Sedangkan nilai *intercept* rata-rata dari seluruh negara direfleksikan oleh β_1 .

2. Penentuan Model Estimasi Regresi Data Panel

Menurut Gujarati dalam (Gujarati 2014) 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:

H_0 : Model Common Effects

Ha: Model Fixed Effects

Statistik uji yang digunakan merupakan uji F, yaitu dengan membandingkan antara nilai F-statistik dan nilai F-tabel. Nilai F-tabel diperoleh dari $\{(\alpha, df_1(n-1), df_2(nt-n-k))\}$. Dimana α adalah taraf signifikansi (0.05), n adalah jumlah negara, t adalah jumlah waktu, nt adalah jumlah negara dikali waktu (jumlah observasi), dan k adalah 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 memodel estimasi regresi data panel. Sebaliknya, apabila F-statistik lebih kecil dari F-tabel, maka H_0 diterima, yang berarti model *common effects* yang paling baik untuk digunakan dalam memodel estimasi regresi data panel.

Selain dengan membandingkan 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 apakah

terdapat hubungan antara *error component* dengan satu atau lebih variabel independen dalam suatu model. Hipotesis awalnya adalah tidak terdapat hubungan antara *error component* dengan variabel independen.

Menurut Baltagi hipotesis dari uji Hausman adalah sebagai berikut:

H_0 : Model Random Effects

H_a : Model Fixed Effects

Statistik uji yang digunakan adalah uji *chi square*. Jika nilai *chi square*-statistik $>$ *chi square*-tabel ($\alpha, k-1$) atau nilai *p-value* kurang dari taraf signifikansi yang ditentukan, maka 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.

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 dapat digunakan model *fixed effects*.

b. 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*

H_a = Model *Common Effects*

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

[3.8]

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

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. Deteksi Gejala Asumsi Klasik

Model regresi data panel dapat dikatakan sebagai model yang baik, apabila memenuhi empat kriteria berikut: Best, Linear, Unbiased, dan Estimator. Keempat kriteria tersebut biasa disingkat dengan *BLUE*. Apabila model persamaan yang terbentuk tidak memenuhi kriteria *BLUE*, maka model persamaan tersebut diragukan dapat menghasilkan nilai-nilai prediksi yang akurat. Untuk itu perlu dilakukannya deteksi gejala asumsi klasik untuk mengetahui apakah model persamaan tersebut telah memenuhi kriteria *BLUE*. Hal ini dikarenakan model persamaan telah memenuhi kriteria *BLUE* apabila telah memenuhi asumsi klasik. Deteksi gejala asumsi klasik ini mencakup deteksi normalitas, deteksi linearitas, deteksi heterokedastisitas deteksi multikolinearitas, dan deteksi autokorelasi. Apabila model persamaan yang dideteksi telah bebas dari lima asumsi tersebut, maka dapat dikatakan model persamaan tersebut akan menjadi estimator yang baik.

a. Deteksi Normalitas

Deteksi normalitas juga dapat dilakukan dengan menggunakan uji *Jarque-Bera*, yaitu dengan mendeteksi normalitas pada residualnya yang dihasilkan dari model persamaan regresi linear yang digunakan. Uji *Jarque-Bera* ini menggunakan hipotesis sebagai berikut:

H_0 : Residual berdistribusi normal

H_a : Residual tidak berdistribusi normal

Kriteria uji: H_0 ditolak jika nilai $JB > \chi^2$ -tabel ($\alpha, k-1$) artinya residual tidak berdistribusi normal, dan jika sebaliknya maka residual berdistribusi normal. Selain melihat hasil dari nilai JB , dapat juga dengan melihat nilai probabilitas dari JB . Apabila P -value dari $JB < 0,05$, maka H_0 ditolak artinya residual tidak berdistribusi normal, jika sebaliknya maka H_0 diterima artinya residual berdistribusi normal.

b. Deteksi Heteroskedastisitas

Asumsi penting (Asumsi Gauss Markov) dalam Penggunaan OLS adalah varians residual yang konstan. Varian dari residual tidak berubah dengan berubahnya satu atau lebih variabel bebas. Jika asumsi ini terpenuhi, maka residual disebut heteroskedastisitas (Ariefianto 2012).

Deteksi heteroskedastisitas bertujuan untuk menguji apakah nilai residual yang ditentukan oleh variabel independen (*regressors*), memiliki nilai varians yang konstan atau sama dengan σ^2 (Gujarati 2014). Model regresi dikatakan baik apabila tidak terjadi heteroskedastisitas, artinya adanya ketetapan atau konstan antara varians dari nilai residual dari satu pengamatan ke pengamatan lainnya (Homoskedastisitas). Untuk mendeteksi heteroskedastisitas dilakukanlah uji *Glejser*, yaitu dengan meregressi nilai dari residual absolut dengan variabel X_1, X_2, X_3 , dan X_4 (Gujarati 2014). Hipotesis yang digunakan dari uji *Glejser* adalah sebagai berikut:

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

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

Berdasarkan hipotesis tersebut, maka kriteria pengambilan kesimpulan yakni jika nilai probabilitas (*p-value*) dari t-statistik $> 0,05$, maka H_0 diterima, artinya varians error bersifat homokedastik. Jika sebaliknya, maka H_0 ditolak, yang berarti varians error bersifat heterokedastik.

Adapun penyebab terjadinya heterokedasitas menurut Gujarati (2003) dan Pindyck dan Rubinfeld (1997) dalam (Ariefianto 2012) diantaranya ;

1. *Situasi Error Learning*. Misalnya kita ingin mengetahui hubungan tingkat kesalahan mengetik terhadap berbagai variable. Jika kita menggunakan sampel yang bersifat panel/time series akan sangat mungkin model yang dimiliki akan bersifat heterokedasitas. Hal ini disebabkan kesalahan pengetikan akan menurun dari waktu ke waktu dan terjadi konvergensi diantara elemen sampel (kesalahan anggota sampel yang paling tidak terampil akan menurun mendekati mereka yang awalnya sdah terampil)
2. *Kemampuan disreksi*. Hal ini tampak jelas pada penelitian dengan menggunakan variable pendapatan. Aktivitas oleh individu yang memiliki pendapatan tinggi akan jauh lebih variaetif dibandingkan mereka yang berpendapatan rendah. Dengan demikian suatu model regresi dengan menggunakan variable semacam ini akan mengalami peningkatan residual kuadrat dengan semakin besarnya pendapatan.

3. *Perbaiki Teknik pengambilan data.* Peneliti akan belajar untuk menarika informasi dengan benar, dengan demikian kesalahan akibat ekstraksi data akan semakin menurun.
4. *Keberadaan Outlier.* Outlier adlaah data yang memiliki karakteristik sangat berbeda dari kondisi yang umum. Misalnya kita memiliki suatu set data pendapatan dengan kisaran IDR 205 juta per bulan, keberadaan individu dengan pendapatan 100 juta dapat dikatakan outlier.
5. *Masalah spesifikasi.* Jika model pada populasi adalah nonlinear (misalnya eksponsial) namun kita memaksa menggunakan model linier. Disini, kuadrat residual akan menignkat cepat dengan meningkatnya nilai variable bebas.

c. Deteksi Multikolinieritas

Gujarati (2003) dalam (Ariefianto 2012) menyatakan bahwa multikolinieritas adalah fenomena sampling. Ia terjadi pada sampel dan bukan pada populasi. Adapun dalam definisi lain oleh Kmenta (1986) dalam (Ariefianto 2012) menyatakan permasalahan multikolinieritas adalah permasalahan derajat, bukan apakah ada atau tidak suatu kolinearitas pada data yang dimiliki. Deteksi multikolinieritas 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 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 dalam (Gujarati 2014) 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.

d. Deteksi Autokorelasi

Deteksi autokorelasi bertujuan untuk mendeteksi apakah dalam model regresi linier ada korelasi antara kesalahan pengganggu pada periode t dengan kesalahan pengganggu pada periode t dengan kesalahan pengganggu periode $t-1$ (tahun sebelumnya). Model regresi yang baik adalah tidak ada terjadi autokorelasi. Cara memprediksi dalam suatu model regresi terdapat autokorelasi atau tidak dapat dengan cara uji *Durbin-Watson* (DW test). Rumus statistik d *Durbin-Watson* sebagai berikut:

Uji *Durbin-Watson* akan menghasilkan nilai *Durbin-Watson* (DW) dan dari nilai *Durbin-Watson* tersebut dapat menentukan keputusan apakah terdapat autokorelasi atau tidak dengan melihat tabel berikut:

$$d = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e^2 t}$$

Dalam autokorelasi terhadap tabel yang menunjukkan keterangan menurut range durbin watson dalam (Firdaus 2004).

Tabel 3.1**Range Durbin-Watson untuk Autokorelasi**

Durbin-Watson	Kesimpulan
$DW < dl$	Ada autokorelasi Positif
$dl \leq DW \leq du$	Ragu-Ragu
$du \leq DW \leq 4-du$	Tidak ada autokorelasi
$4-du \leq DW \leq 4-dl$	Ragu-Ragu
$4-dl < DW$	Ada autokorelasi Negatif

Sumber: Muhammad Firdaus

Secara lebih spesifik beberapa penyebab autokorelasi atau disebut dengan korelasi serial menurut Wooldridge (2005), Vogelpang dan Gujarati (2003) dalam (Ariefianto 2012) menyebutkan diantaranya ;

1. *Inertia* . salah satu karakteristik umum dari data yang bersifat time series adalah inertia (*sluggishness*). Penyesuaian akibat suatu guncangan terhadap variable makro ekonomi adalah bersifat bertahap, dan berlangsung sepanjang waktu tertentu. Hal ini juga terjadi pada sekelompok variable.
2. *Specification bias*. Yakni kesalahan dalam menspesifikasi model. Terdapat dua tipe kesalahan, yakni (1) mengeluarkan variable yang seharusnya ada pada model (*omitted variable*) dan (2) bentuk fungsional yang tidak benar.
3. *Fenomena cobweb*. Sering terjadi pada pemodelan dimana terdapat suatu fenomena lagged response. Hal ini sering terjadi misalnya pada estimasi fungsi pasokan produk pertanian. Petani akan

mendasarkan keputusan jumlah produksi berdasarkan harga satu periode yang lalu.

4. *Rekayasa data*. Karena satu hal dan lain hal seseorang peneliti kadang harus “menukangi” data. Salah satu praktik “menukangi data” yang sering misalnya akibat perbedaan frekuensi
5. *Dampak musiman*. Misalnya variabel terikat kita gunakan memiliki karakter musiman (misalnya produksi beras), sedangkan variabel penjelas digunakan tidak. Apabila variabel terikat ini tidak disesuaikan terlebih dahulu (*deseasonalized*) maka residual dari regresi akan menunjukkan karakter musiman yang ada pada variabel terikat.

3. Uji Hipotesis

Uji hipotesis ini diperlukan untuk menguji apakah koefisien regresi yang didapat signifikan. Maksud dari signifikan di sini adalah suatu nilai koefisien regresi yang secara statistik tidak sama dengan nol. Jika *slope coefficient* sama dengan nol, berarti tidak dapat dikatakan bahwa terdapat cukup bukti untuk menyatakan variabel independen berpengaruh terhadap variabel dependen. Maka dari itu, semua koefisien yang terdapat pada persamaan regresi harus di uji. Terdapat dua jenis uji hipotesis yang dapat dilakukan untuk menguji koefisien regresi, yaitu uji t dan uji F. Uji t digunakan untuk mengetahui secara parsial apakah variabel-variabel independen berpengaruh signifikan terhadap variabel dependen. Dalam uji t pengujian hipotesis bersifat individual dengan melihat apakah suatu parameter regresi telah sesuai dengan hipotesis. Sedangkan uji F digunakan

untuk secara simultan/keseluruhan apakah variabel-variabel independen berpengaruh signifikan terhadap variabel dependen atau sering disebut dengan pengujian hipotesis berganda..

a. Uji t (Parsial)

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

Hipotesis statistik untuk variabel pendidikan :

$H_0: \beta_2 = 0$, artinya secara parsial tidak ada pengaruh Pendidikan terhadap produktivitas tenaga kerja

$H_a: \beta_2 \neq 0$, artinya secara parsial ada pengaruh Pendidikan terhadap produktivitas tenaga kerja

Hipotesis statistik untuk variabel kesehatan:

$H_0: \beta_3 = 0$, artinya secara parsial tidak ada pengaruh kesehatan terhadap produktivitas tenaga kerja

$H_a: \beta_3 \neq 0$, artinya secara parsial ada pengaruh kesehatan terhadap produktivitas tenaga kerja

Hipotesis statistik untuk variabel teknologi:

$H_0: \beta_4 = 0$, artinya secara parsial tidak ada pengaruh teknologi terhadap produktivitas tenaga kerja

$H_a: \beta_4 \neq 0$, artinya secara parsial ada pengaruh teknologi terhadap produktivitas tenaga kerja

Hipotesis statistik untuk variabel tingkat upah:

$H_0: \beta_5 = 0$, artinya secara parsial tidak ada pengaruh upah terhadap produktivitas tenaga kerja

$H_a: \beta_5 \neq 0$, artinya secara parsial ada pengaruh upah terhadap produktivitas tenaga kerja

Dasar pengambilan keputusan, apabila angka probabilitas signifikansi > 0.05 maka H_0 diterima dan H_a ditolak. Artinya variabel 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 dengan t-tabel. H_0 diterima jika t-tabel $>$ t-statistik dan ditolak jika t-tabel $<$ t-statistik. Nilai t-statistik dapat dicari dengan menggunakan rumus sebagai berikut

[3.9]

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

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

$0 =$ Hipotesis awal = nol

Keterangan:

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 F (Simultan)

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:

$$H_0 : \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$$

$$H_a : \beta_2 \neq \beta_3 \neq \beta_4 \neq \beta_5 \neq 0$$

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 regresi, yaitu apabila $F\text{-tabel} > F\text{-statistik}$ maka H_0 diterima dan apabila $F\text{-tabel} < F\text{-statistik}$ maka H_0 ditolak. Nilai dari $F\text{-statistik}$ datang dicari dengan menggunakan rumus sebagai berikut:

$$[3.10]$$

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

Keterangan:

R^2 = koefisien determinasi (residual)

k = jumlah variabel independen ditambah intercept dari suatu model persamaan

n = jumlah sampel

4. Analisis Koefisien Determinasi (R^2)

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

[3.11]

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

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.

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.