

## BAB 3 METODOLOGI PENELITIAN

Bab ini akan membahas mengenai tahapan-tahapan yang dilakukan dalam pengolahan data dalam upaya mendapatkan hasil estimasi yang mampu menjawab tujuan penelitian yaitu meneliti pengaruh likuiditas terhadap imbal hasil (*return*) saham seluruh saham selama periode tahun 2004-2006 di bursa efek Indonesia.

### 3.1 Data

Pemilihan sampel data dalam sebuah penelitian dapat menggunakan dua cara yaitu *probability sampling* atau *non-probability sampling*. *Probability sampling* yaitu data dipilih secara acak, sehingga setiap calon data sampel mempunyai kesempatan atau probabilitas yang sama untuk terpilih menjadi data sampel suatu penelitian. Sedangkan *Non-probability sampling* adalah jika data yang digunakan sebagai sampel harus memenuhi kriteria-kriteria khusus dalam pemilihannya.

Pada skripsi ini, penentuan sampel menggunakan cara *non-probability sampling* dan *probability sampling*. Data sampel yang digunakan pada penelitian ini merupakan data-data sekunder berupa data mingguan yang diolah dari data awal yang berbentuk data harian yang diperoleh dari pusat statistik Bursa Efek Indonesia, dan berbagai sumber lainnya. Data sampel tersebut harus memiliki kriteria khusus yang terdiri dari :

- (1) Perusahaan yang digunakan sebagai sampel data merupakan perusahaan *go publik* yang terdaftar di Bursa Efek Indonesia (BEI) dari tahun 2004-2006
- (2) Data sampel harus memiliki data harga saham dari tahun 2004-2006
- (3) Data sampel harus memiliki data *bid price* dari tahun 2004-2006
- (4) Data sampel harus memiliki data *ask price* dari tahun 2004-2006
- (5) Data sampel harus memiliki data *bid volume* dari tahun 2004-2006
- (6) Data sampel harus memiliki data *ask volume* dari tahun 2004-2006

Dari 401 saham yang listing di BEI sampai maret 2009, diperoleh sampel yang memenuhi kriteria khusus sebanyak 309 saham. Selanjutnya dari 309 saham ini, diambil sampel sebanyak 100 saham secara acak (*Probability sampling*), sehingga setiap data saham mempunyai kesempatan atau probabilitas yang sama untuk terpilih menjadi data dalam penelitian ini. Daftar perusahaan sebagai sampel data penelitian skripsi ini terangkum pada tabel dibawah ini:

**Tabel 3-1**  
**Daftar Perusahaan (Sampel Data)**

NO	Simbol	Nama Perusahaan	NO	Simbol	Nama Perusahaan
1	ADES	Ades Waters Indonesia Tbk	26	BRNA	Berlina Tbk
2	AIMS	Akbar Indo Makmur Stimec Tbk	27	CNTX	Centex (Saham Preferen) Tbk
3	AKSI	Asia Kapitalindo Securities Tbk	28	CTBN	Citra Tubindo Tbk
4	ALKA	Alakasa Industrindo Tbk	29	CTRA	Ciputra Development Tbk
5	ALMI	Alumindo Light Metal Inds. Tbk	30	CTTH	Citatah Industri Marmer Tbk
6	ANTM	Aneka Tambang (Persero) Tbk	31	DEFI	Danasupra Erapacific Tbk
7	APEX	Apexindo Pratama Duta Tbk	32	DLTA	Delta Djakarta Tbk
8	AQUA	Aqua Golden Mississi Tbk	33	DPNS	Duta Pertiwi Nusantara Tbk
9	ARGO	Argo Pantas Tbk	34	DSUC	Daya Sakti Unggul Corp. Tbk
10	ARNA	Arwana Citramulia Tbk	35	EKAD	Ekadharma International Tbk
11	ASBI	Asuransi Bintang Tbk	36	ELTY	Bakrieland Development Tbk
12	ASGR	Astra Graphia Tbk	37	ERTX	Eratex Djaja Tbk
13	ASII	Astra International Tbk	38	ESTI	Ever Shine Textile Inds. Tbk
14	ASJT	Asuransi Jasa Tania Tbk	39	ETWA	Eterindo Wahanatama Tbk
15	ATPK	ATPK Resources Tbk	40	FASW	Fajar Surya Wisesa Tbk
16	AUTO	Astra Otoparts Tbk	41	GDYR	Goodyear Indonesia Tbk
17	BABP	Bank Bumiputera Indonesia Tbk	42	GGRM	Gudang Garam Tbk
18	BAYU	Bayu Buana Tbk	43	GJTL	Gajah Tunggal Tbk
19	BBCA	Bank Central Asia Tbk	44	GMTD	Gowa Makassar Tourism Dev. Tbk
20	BCAP	Bhakti Capital Indonesia Tbk	45	HERO	Hero Supermarket Tbk
21	BCIC	Bank Century Tbk	46	INAF	Indofarma Tbk
22	BKSL	Sentul City Tbk	47	INAI	Indal Aluminium Industry Tbk
23	BMSR	Bintang Mitra Semestaraya Tbk	48	INCF	Indo Citra Finance Tbk
24	BNBR	Bakrie & Brothers Tbk	49	INCO	International Nickel Ind .Tbk
25	BNGA	Bank Niaga Tbk	50	INDR	Indorama Syntetics Tbk

**Tabel 3-1 (lanjutan)**  
**Daftar Perusahaan (Sampel Data)**

NO	Simbol	Nama Perusahaan	NO	Simbol	Nama Perusahaan
51	INKP	Indah Kiat Pulp & Paper Tbk	76	RIGS	Rig Tenders Tbk
52	ISAT	Indosat Tbk	77	RODA	Roda Panggon Harapan Tbk.
53	JIHD	Jakarta Int l Hotel & Dev. Tbk	78	SHID	Hotel Sahid Jaya Tbk
54	JPFA	JAPFA Tbk	79	SIIP	Suryainti Permata Tbk
55	KAEF	Kimia Farma Tbk	80	SMGR	Semen Gresik (Persero) Tbk
56	KBLI	GT Kabel Indonesia Tbk	81	SMMA	Sinar Mas Multiartha Tbk
57	KBLM	Kabelindo Murni Tbk	82	SMRA	Summarecon Agung Tbk
58	KLBF	Kalbe Farma Tbk	83	SOBI	Sorini Agro Asia Corporindo
59	KONI	Perdana Bangun Pusaka Tbk	84	SULI	Sumalindo Lestari Jaya Tbk
60	LION	Lion Metal Works Tbk	85	TCID	Mandom Indonesia Tbk
61	LPCK	Lippo Cikarang Tbk	86	TEJA	Texmaco Jaya Tbk
62	MBAI	Multibreeder Adirama Ind. Tbk	87	TGKA	Tigaraksa Satria Tbk
63	MEDC	Medco Energi International Tbk	88	TLKM	Telekomunikasi Indonesia Tbk
64	META	Nusantara Infrastructure Tbk	89	TMAS	Pelayaran Tempuran Emas Tbk
65	MITI	Mitra Investindo Tbk	90	TMPI	AGIS Tbk
66	MYOR	Mayora Indah Tbk	91	TMPO	Tempo Inti Media Tbk
67	MYTX	Apac Citra Centertex Tbk	92	TOTO	Surya Toto Indonesia Tbk
68	PLAS	Palm Asia Corpora Tbk	93	TRST	Trias Sentosa Tbk
69	PNBN	Bank Pan Indonesia Tbk	94	TRUS	Trust Finance Indonesia Tbk
70	PSDN	Prasidha Aneka Niaga Tbk	95	TSPC	Tempo Scan Pacific Tbk
71	PTRA	New Century Development Tbk	96	ULTJ	Ultra Jaya Milk Tbk
72	PTSP	Pioneerindo Gourmet Int l Tbk	97	UNTR	United Tractors Tbk
73	PUDP	Pudjiadi Prestige Limited Tbk	98	UNVR	Unilever Indonesia Tbk
74	PWON	Pakuwon Jati Tbk	99	VOKS	Voksel Electric Tbk
75	RBMS	Ristia Bintang Mahkotasejati T	100	WAPO	Wahana Phonix Mandiri Tbk

Sumber: Olahan Penulis

### 3.1.1 Sumber Data

Data yang dibutuhkan dalam penelitian ini adalah data mengenai harga saham, *bid price*, *ask price*, *bid volume*, *ask volume* dari seluruh sampel. Data-data tersebut diperoleh dari data-data sekunder yang diperoleh dari [www.yahooofinance.com](http://www.yahooofinance.com), pusat statistik Bursa Efek Indonesia, dan [www.idx.co.id](http://www.idx.co.id).

## 3.2 Metode Penelitian

Pada penelitian ini menggunakan program E-Views untuk melakukan pengolahan data. Metode yang digunakan dalam pengolahan data yaitu dengan metode panel data. Karena itu, pembahasan berikutnya dalam bab ini akan dibahas mengenai metode data panel.

### 3.2.1 Variabel Penelitian

Variabel yang akan digunakan pada penelitian skripsi ini yaitu variabel terikat (*dependent variable*) dan variabel bebas (*independent variable*).

#### 3.2.1.1 Variabel Terikat

Berdasarkan tujuan dari penelitian ini maka variabel terikat (*dependent variable*) yang digunakan yaitu imbal hasil (*return*) saham. Rumus yang dipakai untuk mengukur tingkat return saham berdasarkan rumus Sharpe (1999) sebagai berikut :

$$\text{Return} = \frac{\text{end of period wealth} - \text{beginning of period wealth}}{\text{beginning of period wealth}} \quad (3.1)$$

#### 3.2.1.1 Variabel Bebas

Variabel bebas dari penelitian ini adalah *Depth to relative spread* (DTRS) dari masing-masing sampel dan *risk premium*.

- (1) *Depth to relative spread* (DTRS) adalah rata-rata *ask-depth* dan *bid-depth* dibagi *relative spread*, atau secara matematis dapat ditulis (Ekaputra, 2006) :

$$\text{Depth - to - Relative Spread} = \frac{(\text{ask depth} + \text{bid depth})/2}{\text{relative spread}} \quad (3.2)$$

Dimana :

$$\text{relative spread} = \frac{\text{Ask price} - \text{bid price}}{(\text{ask price} + \text{bid price})/2} \quad (3.3)$$

(2) Variabel bebas lainnya adalah *risk premium*, yang didapat dengan mengurangi *return market* dengan *risk free* (SBI).

$$\text{risk premium} = R_m - r_f \quad (3.4)$$

### 3.2.2 Model Penelitian

Model penelitian ini merujuk kepada penelitian Amihud dan Mendelson (1986) , dengan melakukan modifikasi yaitu dengan hanya menggunakan dua variabel yaitu *risk premium* ( $R_m - r_f$ ) dan mengganti variabel *bid-ask spread* dengan variabel *depth to relative spread* (DTRS) karena pada DTRS telah pengukuran dilakukan dengan menggunakan tiga dimensi likuiditas, dibandingkan dengan *bid-ask spread*, yang hanya menggunakan dua dimensi pengukuran. Maka model untuk penelitian ini dengan menggabungkan variabel-variabel tersebut menjadi:

$$R_{it} = \alpha + \beta \text{riskprem}_t + \beta \text{DTRS}_{it} + e_{it} \quad (3.5)$$

Dimana:

$$\text{riskprem}_t = R_m - r_f$$

$$\text{DTRS}_{it} = \text{depth to relative spread}$$

$$e_{it} = \text{simbol error dalam model}$$

$$\alpha, \beta = \text{koefisien parameter model penelitian}$$

Imbal hasil (*return*) saham sebagai variabel terikat (*dependent variable*), sedangkan *risk premium* dan *depth to relative spread* sebagai variabel bebas (*independent variable*). Pada persamaan di atas koefisien  $\alpha$  merupakan *intercept* dan  $\beta$  merupakan *slope* yang menunjukkan besar pengaruh dari suatu variabel bebas yang digunakan dalam persamaan terhadap variabel terikat, dengan mengasumsikan variabel lain bernilai nol (0). Tingkat signifikansi dari setiap variabel dapat dilihat dari probabilitinya. Apabila probabilitinya bernilai kurang dari 0.05 (dengan tingkat signifikansi 5%) atau bernilai kurang dari 0.1 (dengan tingkat signifikansi 10%) maka dapat dikatakan bahwa variabel bebas tersebut mempunyai pengaruh yang signifikan terhadap variabel terikat. Dan berlaku sebaliknya, apabila nilai probabilitinya lebih dari 0.05 atau 0.1 maka variabel bebas tersebut tidak memiliki pengaruh yang signifikan terhadap variabel terikat tersebut. Umumnya, untuk ilmu sosial, termasuk ekonomi dan keuangan, besarnya probabiliti adalah 5% (Nachrowi,2006).

### 3.2.3 Hipotesis

Untuk membuktikan apakah likuiditas saham mempengaruhi imbal hasil saham (*stock return*) dengan menggunakan *depth to relative spread* sebagai pengukur likuiditas. Hipotesis yang digunakan untuk membuktikan pengaruh likuiditas terhadap likuiditas imbal hasil saham (*stock return*) adalah:

**H<sub>0</sub>**: Likuiditas saham tidak mempengaruhi imbal hasil saham.

**H<sub>1</sub>**: Likuiditas saham mempengaruhi imbal hasil saham.

Selain itu, Untuk membuktikan apakah *risk premium* mempengaruhi imbal hasil saham (*stock return*). Hipotesis yang digunakan untuk membuktikan pengaruh *risk premium* terhadap likuiditas imbal hasil saham (*stock return*) adalah:

**H<sub>0</sub>**: *Risk premium* tidak mempengaruhi imbal hasil saham.

**H<sub>1</sub>**: *Risk premium* mempengaruhi imbal hasil saham.

### 3.2.4 Teknik Pengolahan Data

Dalam suatu persamaan regresi yang hanya terdapat lebih dari satu variabel bebas seperti dalam penelitian ini, maka model tersebut dikenal dengan sebutan regresi linier majemuk (Nachrowi, 2006). Metode penelitian ini akan menggunakan permodelan regresi karena tujuan dari penelitian ini ialah untuk mengetahui pengaruh yang terjadi antara *given variable* dengan satu atau lebih variabel lainnya (Brooks, 2005). Teknik pengolahan data yang digunakan dalam penelitian ini yaitu menggunakan regresi panel data.

#### 3.2.4.1 Data Panel

Teknik pemodelan terhadap data beberapa individu yang dikumpulkan dari waktu ke waktu atau dapat dikatakan sebagai data *cross section* dan diikuti pada periode waktu tertentu dikenal dengan nama data panel (Nachrowi, 2006). Panel data merupakan gabungan antara jenis data *time series* dan *cross section* sehingga panel data merupakan data yang memiliki dimensi waktu dan ruang.

Dalam menggunakan permodelan dengan panel data terdapat kesulitan yang mungkin bisa ditemukan yaitu dalam mengidentifikasi *t-ratios* atau *f-stat* dari model regresinya. Hal ini dapat terjadi jika jumlah observasi *cross section* hanya sedikit dan jumlah data *time series* banyak. Maka terdapat beberapa pendekatan yang dapat dilakukan untuk mengefisiensikan perhitungan model regresi panel data. Pendekatan-pendekatan tersebut yaitu:

1. Metode *Common-Constant (The Pooled OLS Method)*
2. Metode *Fixed Effect (FEM)*
3. Metode *Random Effect (REM)*

##### 3.2.4.1.1 Metode *Common-Constant (The Pooled OLS Method)*

Pada metode OLS diasumsikan baik *intercept* maupun *slope* tidak berubah baik antara individu maupun antar waktu (Nachrowi, 2006). Karena itu, metode OLS merupakan metode yang paling sederhana sehingga dapat dikatakan bahwa nilai  $\alpha$  dan  $\beta$  akan sama (konstan) untuk setiap data *time series* dan *cross section*. Hal ini

menimbulkan pertanyaan apakah realistis jika dibuat suatu model, di mana sebuah perusahaan yang bergerak di industri makanan mempunyai *intercept* yang sama dengan perusahaan yang bergerak di sektor jasa?, untuk mengatasi masalah tersebut, ada dua teknik lainnya yang bisa digunakan untuk membuat model dari data panel yaitu Metode Efek Tetap (*The Fixed Effect*) dan Metode Efek Random (*The Random Effect*) (Nachrowi, 2006).

### 3.2.4.1.2 Metode *Fixed Effect* (*Fixed Effect Model*)

Telah dinyatakan diatas bahwa asumsi pembuatan model yang menghasilkan  $\alpha$  konstan untuk setiap individu (i) dan waktu (t) kurang realistis. Dalam Metode Efek Tetap (MET) kita dapat mengatasi hal tersebut, karena metode ini memungkinkan adanya perubahan  $\alpha$  pada setiap i dan t. Secara matematis model Metode Efek Tetap dinyatakan sebagai berikut (Nachrowi, 2006) :

$$Y_{it} = \alpha + \beta X_{it} + \gamma_2 W_{2t} + \gamma_3 W_{3t} + \dots + \gamma_N W_{Nt} + \delta_2 Z_{i2} + \delta_3 Z_{i3} + \dots + \delta_T Z_{iT} + s_{it} \quad (3.6)$$

Di mana :  $Y_{it}$  = variabel terikat untuk individu ke-i dan waktu ke-t

$X_{it}$  = variabel bebas untuk individu ke-i dan waktu ke-t

$W_{it}$  dan  $Z_{it}$  merupakan variabel dummy yang didefinisikan sebagai berikut :

$W_{it} = 1$  ; untuk individu i;  $i = 1, 2, \dots, N$   
 $= 0$  ; lainnya.

$Z_{it} = 1$  ; untuk periode t;  $t = 1, 2, \dots, T$   
 $= 0$  ; lainnya.

Variabel *dummy* digunakan untuk membedakan *intercept*nya sehingga metode ini juga dikenal dengan *Least Square Dummy Variable* (LSDV) Model. Metode ini tetap mempunyai kelemahan yaitu semakin banyak jumlah data *cross section* maka akan membuat *degrees of freedom* semakin besar dengan memperkenalkan N *dummies*.



### 3.2.4.1.3 Metode *Random Effect (Random Effect Model)*

Sebagaimana telah dibahas diatas bahwa pada Metode Efek Tetap, perbedaan karakteristik individu dan waktu diakomodasikan pada *intercept* sehingga *intercept*nya berubah antar individu dan antar waktu. Sementara Model Efek Random perbedaan karakteristik individu dan waktu diakomodasikan pada *error* dari model. Mengingat ada dua komponen yang mempunyai kontribusi pada pembentukan *error*, yaitu individu dan waktu, maka *random error* pada Metode Efek Random juga perlu diurai menjadi *error* untuk komponen individu, *error* komponen waktu dan *error* gabungan. Dengan demikian, persamaan Metode Efek Random diformulasikan sebagai berikut (Nachrowi, 2006) :

$$Y_{it} = \alpha + \beta X_{it} + \varepsilon_{it} ; \varepsilon_{it} = u_i + v_t + w_{it} \quad (3.7)$$

di mana :

$u_i$  : Komponen *error cross section*

$v_t$  : Komponen *error time series*

$w_{it}$  : Komponen *error* gabungan

### 3.2.4.2 Pemilihan Metode Estimasi

Dalam memodelkan panel data, telah dijelaskan terdapat tiga metode. Pemilihan metode yang akan digunakan dalam melakukan suatu penelitian harus dilakukan secara valid. Maka untuk melakukan pemilihan metode yang tepat terdapat tiga tahap uji yang dapat dilakukan. Pemilihan metode yang tepat yang akan digunakan untuk memodelkan panel data bertujuan agar pendekatan yang dipilih cocok dengan tujuan penelitian dan cocok dengan karakteristik data sampel yang digunakan sehingga proses estimasi memberikan hasil yang lebih tepat (Pinagara,2009) :

(1) Pemilihan secara teoritis.

Menurut Nachrowi (2006) metode *common constant* dikatakan terlalu sederhana untuk mendeskripsikan fenomena yang ada. Sebagai pilihan selanjutnya dapat memilih antara metode *fixed effect* dan metode *random effect*. Penentuan pertama dapat dilakukan secara teoritis dengan melihat

hubungan korelasi antara individual *cross section*, komponen error  $\varepsilon_i$  dan X sebagai *regressor* (variabel bebas) (Gujarati, 2004). Jika diasumsikan  $\varepsilon_i$  dan X tidak berkorelasi maka metode *random effect* yang digunakan. Sebaliknya apabila  $\varepsilon_i$  dan X berkorelasi maka metode yang paling tepat digunakan ialah metode *fixed effect*. Namun, kelebihan pada metode *fixed effect* tidak perlu mengasumsikan bahwa komponen error tidak berkorelasi dengan variabel bebas yang mungkin sulit dipenuhi (Nachrowi, 2006)

(2) Pemilihan atas dasar sampel data penelitian

Jika ternyata secara teoritis penentuan model tidak dapat memberikan jawaban secara tepat, maka dasar pemilihan model selanjutnya dapat didasarkan pada sampel penelitian. Apabila data sampel diambil atas suatu populasi secara acak maka lebih tepat menggunakan metode *random effect*.

Apabila pemilihan sampel data telah ditentukan berdasarkan populasi yang ada maka pemilihan metode *fixed effect* lebih tepat digunakan. Sebagai tambahan, jumlah data *cross section* dengan data *time series* juga dapat menentukan permodelan mana yang lebih tepat digunakan. Jika jumlah T (data *time series*) lebih besar daripada jumlah N (data *cross section*), maka metode *fixed effect* lebih dipilih. Saat jumlah N lebih besar daripada jumlah T, maka digunakan metode *random effect* dalam pengolahannya (Gujarati, 2004).

(3) Pemilihan dengan Uji Formal Statistik

Pengujian secara formal dapat pula dilakukan. Metode *fixed effects* dengan metode *common constant* dapat diuji dengan (*incremental*) *F-Test* atau menggunakan LM-Test, sedangkan metode *random effects* dengan metode *common constant* diuji dengan *Lagrange Multiplier (LM) test* (Breusch dan Pagan, 1980). Uji tersebut mempunyai hipotesis  $H_0$  : Metode *common constant*, sehingga apabila  $H_0$  tidak ditolak, metode *common constant (pooled*

*OLS regression*) yang akan dipilih. Sedangkan *The Hausman specification test* (Hausman, 1978) membandingkan antara metode *fixed effect* dan metode *random effect*. Pemilihan uji formal statistik, yaitu :

- a) Untuk pemilihan model yang baik, dapat diketahui dengan menggunakan *F-test*. Persamaan yang diestimasi dengan OLS adalah persamaan *common constant* sedangkan yang diestimasi dengan LSDV adalah *fixed effect* (Asteriou, 2007:347).

$$F = \frac{(R_{FE}^2 - R_{CC}^2) / N - 1}{(1 - R_{FE}^2) / NT - N - K} \approx F(N - 1, NT - N - K) \quad (3.8)$$

dimana,

$R_{FE}^2$  = koefisien determinasi untuk model *fixed effect* (LSDV model)

$R_{CC}^2$  = koefisien determinasi untuk model *common constant* (OLS model)

N = jumlah sampel *cross section*

T = jumlah sampel *time series*

K = total jumlah koefisien regresi (termasuk konstanta)

Jika F-hitung > F-tabel (n-1, nt-n-k) maka tolak  $H_0$ , OLS model *invalid* sehingga LSDV atau FEM adalah *valid*. Dengan hipotesis sebagai berikut

$H_0$  : metode *pooled least square*

$H_1$  : metode *fixed effects*

- b) Untuk memilih antara *random effects* dan *pooled least square* digunakan *The Breusch-Pagan LM Test* dengan menggunakan hipotesis sebagai berikut:

$H_0$  : metode *pooled least square*

$H_1$  : metode *random effects*

Dasar penolakan  $H_0$  menggunakan statistik LM-Test yang berdasarkan distribusi *Chi-Square*.

- c) Untuk memilih antara penggunaan *random effects* dan *fixed effects* maka digunakan pertimbangan statistik *Chi-Square* (Gujarati, 2004:651).

*Haussman test* menggunakan hipotesis sebagai berikut:

$H_0$ : metode *random effects*

$H_1$ : metode *fixed effects*

Tolak  $H_0$  jika *Chi-Square* hitung > *Chi-Square* tabel

### 3.2.4.3 Pengujian Asumsi

Disaat membuat persamaan regresi dari sekumpulan data dalam suatu penelitian, dianggap bahwa model yang diduga dengan OLS tersebut tidak melanggar asumsi-asumsi yang membuat penduga menjadi BLUE (*Best Linier Unbiased Estimated*), atau mempunyai sifat yang linier, tidak bias, dan varian minimum.

Metode OLS dapat memberikan penduga koefisien regresi yang baik atau bersifat BLUE. Akan tetapi, sifat tersebut didasarkan pada berbagai asumsi yang tidak boleh dilanggar agar penduga tetap BLUE. Teori tersebut dikenal dengan sebutan Teorema Gauss-Markov. Asumsi-asumsi yang melandasi estimasi koefisien regresi dengan metode OLS tersebut adalah :

- (1)  $E(u_t) = 0$  Nilai rata-rata dari error nol
- (2)  $Var(u_t) = \sigma^2$  Varians dari error konstan dan finite untuk setiap nilai  $x_t$  (tidak terdapat heterokedastisitas)
- (3)  $Cov(u_i, u_j) = 0$  Error secara statistik terbukti independen satu sama lain (tidak terdapat autokorelasi)
- (4)  $Cov(u_t, x_t) = 0$  Tidak ada hubungan antara error dengan  $x$  (tidak terdapat multikolinieritas)
- (5)  $u_t \sim N(0, \sigma^2)$   $u_t$  terdistribusi secara normal

Jika asumsi 1 hingga 4 terpenuhi maka estimator pada *ordinary least square* dikatakan BLUE (*Best Linear Unbiased Estimators*). *Best*; estimator memiliki varians minimum, *Linear*; parameter yang diestimasi memiliki hubungan linear, *Unbiased*; nilai sesungguhnya dari parameter akan sama dengan nilai estimator, *Estimators*;

bahwa adalah  $\hat{\beta}$  (sampel) adalah estimator  $\beta$  (populasi) . Konsisten: kemungkinan nilai estimasi akan berbeda jauh dengan nilai parameter populasi akan mendekati nol seiring dengan penambahan jumlah sample. Tidak Bias: jika dirata-rata, nilai estimasi akan sama dengan nilai yang sebenarnya. Efisien: estimator dari parameter disebut efisien jika telah terbukti tidak bias dan tidak ada estimator lain yang memiliki varians yang lebih kecil.

#### 3.2.4.3.1 Heterokedastisitas

Dalam salah satu asumsi yang harus dipenuhi agar taksiran parameter dalam model regresi bersifat BLUE maka  $\text{Var}(u_i) = \sigma^2$ , atau dengan kata lain, semua residual atau *error* mempunyai varian yang sama. Kondisi seperti itu disebut dengan homoskedastis. Sedangkan jika varian tidak konstan atau berubah-ubah disebut dengan heteroskedastis. Pengujian heterokedastisitas dapat dilakukan dengan uji grafik (*Residual Graph*) maupun uji *white (Residual Test)*. Hipotesis pada *White Heterokedasticity - Test* yaitu  $H_0$ : tidak terdapat heterokedastisitas, yang berarti tidak ada hubungan antara error dengan variabel bebasnya.  $H_1$ : terdapat heterokedastisitas, yang berarti ada hubungan antara error dengan variabel bebasnya

#### 3.2.4.3.2 Autokorelasi

Dalam menduga parameter dalam regresi, OLS mengasumsikan bahwa error merupakan variabel random yang independen (tidak berkorelasi) agar penduga bersifat BLUE. Atau secara matematis dituliskan  $\text{Cov}(u_i, u_j) = 0$ , Artinya tidak ada korelasi antara  $u_i$  dan  $u_j$ . Otokorelasi dapat terjadi, jika observasi yang berturut-turut sepanjang waktu mempunyai korelasi satu dengan yang lainnya. Cara untuk medeteksinya dapat digunakan dua cara (pinagara,2009), yaitu:

- (1) Uji *Durbin Watson* (Uji-DW). Dengan hipotesis yaitu  $H_0$  : Tidak ada autokorelasi dan  $H_1$ : Ada autokorelasi. Uji-DW ini akan tidak tolak  $H_0$  saat angka DW-Stat pada tabel statistik pengujian berada disekitar angka 2 ( $1.5 < \text{DW-Stat} < 2.5$ ), yang berarti dapat disimpulkan bahwa error tidak berautokorelasi.

- (2) Uji Residual pada program E-Views. Uji residual ini menguji korelogram pada spesifikasi lag. Akan dikatakan bahwa tidak ada autokorelasi saat *pvalue* dari *Q-Stat* signifikan (lebih kecil dari 0,025, *two-tailed* pada tingkat signifikansi 5%).

#### 3.2.4.3.3 Multikolinearitas

Multikolinearitas dapat diartikan sebagai adanya keberadaan dari satu atau lebih hubungan linear diantara beberapa variabel bebas pada permodelan regresi (Gujarati, 2004:342). Hal ini melanggar asumsi regresi dimana disyaratkan sebaliknya untuk menghasilkan estimator BLUE. Uji multikolinearitas dapat dilakukan dengan menghitung koefisien korelasi antar variabel independen. Koefisien korelasi yang tinggi antar variabel bebas mengindikasikan makin tingginya masalah multikolinearitas pada permodelan regresi. Pendeteksian masalah ini dapat dilihat dari:

1. Nilai  $R^2$  yang terlampau tinggi (lebih dari 0,8) tetapi tidak ada atau sedikit T-stat yang signifikan
2. Nilai F-Stat yang signifikan namun T-Stat dari masing-masing variabel bebas tidak signifikan

Pengujian atas multikolinearitas masih banyak diperdebatkan, sehingga cara untuk mengatasinya biasanya dengan mentransformasi data, men-*drop* salah satu variabel yang tidak signifikan atau dengan membiarkannya (Gujarati, 2003:348). Bagaimanapun dengan tingkat multikolinearitas yang tinggi, dikatakan bahwa permodelan regresi tetap memegang asumsi dari BLUE *property*.