

# Model Data Panel: Sebuah Survei

Oleh: Prof. Dr. Adler Haymans Manurung<sup>1</sup>

## Pendahuluan

Koutsoyiannis (1977) menyatakan bahwa ada empat jenis yang dipergunakan untuk estimasi model yaitu time series (runtun waktu), Cross-section; panel data dan engineering data. Wooldridge (2006) membagi data menjadi empat kelompok dimulai dengan cross-section; time series (runtun waktu); pooled cross-section dan Panel atau Longitudinal Data. Umumnya data yang tersedia dalam bentuk runtun waktu, karena data tersebut bisa memperlihatkan hasil yang diperoleh/dicapai atas kebijakan yang diputuskan dan dilaksanakan beberapa waktu lalu. Sehingga data ini sering dipublikasikan oleh lembaga yang berwenang mempublikasikannya. Wooldridge (2006) menyatakan data runtun waktu sangat sulit menganalisisnya. Belakangan, mulai muncul pemakaian data yang berisikan waktu tertentu dan objek penelitian. Adapun objek penelitian seperti perusahaan, rumahtangga, kabupaten, Propinsi, Negara dan perorangan atau kelembagaan lainnya. Data seperti ini disebut dengan data cross-section. Wooldridge (2006) menyatakan ciri terpenting data cross-section bahwa diasumsikan data dipilih secara acak (*random sampling*). Data ini banyak digunakan dalam menganalisis ekonomi mikro dan evaluasi kebijakan ekonomi.

Selanjutnya, data cross-section digabungkan dengan data runtun waktu dikenal dengan pooled cross-section. Salah satu contoh data pooled cross-section yaitu menggabungkan dua data cross-section. Misalkan, Total asset perusahaan terdaftar di BEI pada tahun 2000 dan 2010. Kedua data digabungkan sehingga sampel penelitian menjadi lebih besar. Data seperti ini biasanya dilakukan peneliti menggabungkan dua hasil sensus atau survey yang berbeda dalam rangka melihat kebijakan yang dilakukan sebelumnya. Kemudian, kumpulan data runtun waktu dimana setiap waktu merupakan data cross-section disebut dengan data panel (Longitudinal) atau sering juga disebut dengan bahasa inggrisnya panel data. Pada paper ini data panel dan panel data dianggap satu konteks dan bisa ditulis saling bergantian tetapi maksud sama. Kumpulan runtun waktu yang dimaksudkan harus berurutan dengan unit seperti pada data cross-section.

Penelitian empiris yang menggunakan data panel sudah sangat bervariasi ke berbagai bidang seperti dilakukan Goeltom (1995) melakukan penelitian liberalisasi keuangan Indonesia periode 1981 – 1988. Heshmati (1995) untuk industri pork. Lopez-Pueyo dkk (2008) melakukan penelitian produktifitas manufaktur dan spillovers dari international Riset dan Pengembangan. Alessie dan Lusardi (1997) membahas tabungan dan penghalusan pendapatan (*income smoothing*). Mainardi (2005) membahas pendapatan dan risiko kecelakaan berkerja pada industry pertambangan. Niu dkk (2011) melakukan penelitian mengenai pertumbuhan ekonomi, konservasi energy dan reduksi emisi di delapan Negara. Marin (2001) melakukan penelitian mengenai evolusi dari pH selama periode 1986 – 1997 di Eropa. Manurung (2011, 2012) menganalisis determinan dan optimal struktur kapital untuk perusahaan manufaktur di Indonesia. Egger (2004) melakukan estimasi pengaruh regional blok perdagangan. Egger dkk (2005) melakukan penelitian kompetisi pajak negara bagian di Amerika Serikat. Tsay and Ando (2012) melakukan penelitian pengaruh krisis sub-prime terhadap bursa saham. Gonzalez dan Marrero (2012) melakukan penelitian dalam bidang transportasi.

---

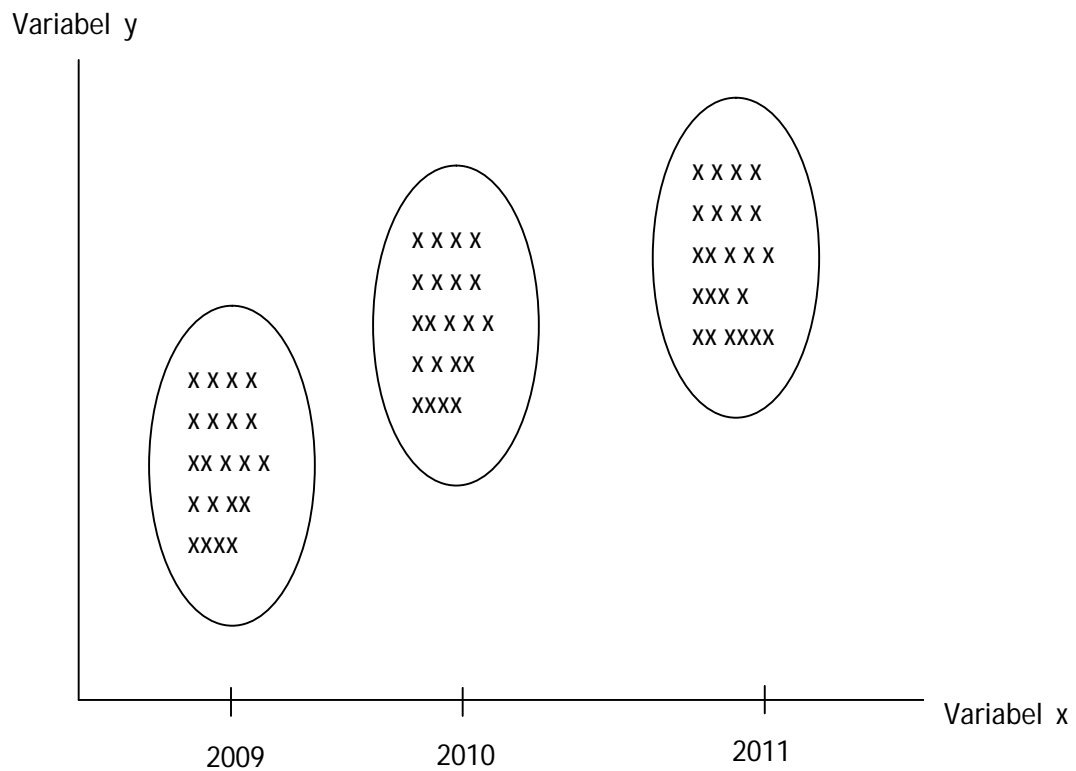
<sup>1</sup> Penulis adalah Guru Besar Pasar Modal dan Perbankan.

Panel data dikenal sudah sejak lama, tetapi penelitian panel terus berkembang sejak Mudlak (1961) dan Balestra dan Nerlove (1966) menerbitkan tulisannya dalam jurnal<sup>2</sup>. Adapun buku teks ekonometrika yang membahas data panel seperti Chamberlain (1984); Maddala (1993); Baltagi (1995); Johnston and Dinardo (1997); Hsiao (1999); Hayashi (2000); Nerlove (2002); Arellano (2003); Stock dan Watson (2003); Wooldridge (2002, 2006); Verbeek (2008); Cameron dan Trivedi (2008); Gujarati dan Porter (2009); Gujarati (2011); Greene (2012) dan Hill dkk (2012). Sangat bervariasinya penelitian yang menggunakan data panel, sehingga tulisan ini mencoba membahas data panel tersebut sebagai bahan bacaan terutama belum terlihat dalam bahasa Indonesia.

## Data Panel

Sesuai uraian sebelumnya bahwa data panel adalah sekumpulan data runtun waktu yang berisikan sekumpulan cross-section. Cross-section yang dimaksud merupakan objek penelitian seperti perorangan, lembaga bank, perusahaan, KUD, kabupaten, provinsi, Negara dan lembaga lainnya. Waktu penelitian harus dalam berurutan seperti terlihat pada Gambar 1 berikut ini.

Gambar 1: Grafik sampel data untuk Panel Data



Panel data ingin melakukan estimasi model atas bentuk data pada Gambar 1. Baltagi (1995) menyatakan keuntungan penggunaan data panel sebagai berikut:

<sup>2</sup> Nerlove, M. (2002); Essays in Panel Data Econometrics; Cambridge University Press, p.2.

- Pengendalian untuk heterogenitas individu
- Memberikan informasi yang lebih besar tentang data, lebih bervariasi, kurang berhubungan antar variabel, lebih besar derajat kebebasan dan lebih efisien
- Lebih mampu mempelajari penyesuaian dinamis (dynamic of adjustment)
- Lebih mampu mengidentifikasi dan mengukur, yang secara sederhana tidak dapat terdeteksi pada cross-section murni atau data runtun waktu.
- Modelnya mengizinkan peneliti untuk membentuk dan menguji model perilaku yang lebih komplik daripada data cross-section dan runtun waktu murni.
- Data yang dikumpulkan umumnya pada level mikro.

Batalgi (1995) juga menyebutkan kelemahan dari data panel ini yaitu:

- Rancangan dan pengumpulan datanya
- Distorsi dari pengukuran kesalahan (measurement errors)
- Persoalan Selektifitas (selectivity problem)  
Persoalan ini termasuk adanya pemilihan sendiri, non-respon dan atrisi.
- Dimensi waktu yang sangat pendek

Berdasarkan Gambar 1, model dapat diestimasi hanya untuk periode 2009 dan juga demikian untuk 2010 dan 2011, sehingga ada tiga model yang dapat dihasilkan. Tiga model tersebut dikenal dengan cross-section. Jika data 2009, 2010 dan 2011 digabung dan dibuat satu model yang dikenal dengan data pooled. Tetapi, bila data ketiga tahun diperlihatkan dan juga cross-section maka dikenal dengan data panel. Gujarati dan Porter (2009) menyebutkan data panel ini bisa juga disebut sebagai micropanel data atau longitudinal data dan cohort analysis.

Selanjutnya, bila data unit observasi seperti perorangan, kabupaten, perusahaan, Negara sama di setiap waktu selama runtun waktu maka datanya disebut data panel sama (balanced panel). Bila jumlah unit observasi berbeda setiap waktu selama periode runtun waktu disebut data panel tidak sama (unbalanced panel). Bila runtun waktu data yang dipergunakan juga pendek dikenal panel pendek (short panel) dan bila data runtun waktunya panjang dikenal dengan panel panjang (long panel). Pada panel data, jumlah unit observasi atau cross-sectionnya disimbol N dan runtun waktu disimbol T. Panel pendek terjadi  $N > T$  sementara pada panel panjang terjadi sebaliknya  $T > N$ . Estimasi atas data panel ini tergantung kepada pendek panjangnya panel tersebut (Gujarati dan Porter, 2009). Model umum matematis dari data panel (Gujarati, 2011) sebagai berikut:

$$Y_{it} = \beta_1 + \beta_2 X_{2it} + \beta_3 X_{3it} + \mu_{it} \quad (1)$$

$i = 1, 2, \dots, k$   
 $t = 1, 2, \dots, n$

Pada model (1) diatas terlihat adanya subscript pada variabel tak bebas, variabel bebas dan suku kesalahan (error terms) dengan symbol  $it$ .  $i$  menyatakan cross-section data yang menyatakan observasi dimana jumlahnya sebanyak  $k$  dan  $t$  menyatakan waktu yang banyaknya  $n$ . Model (1) dikenal dengan Pooled Data Model dan bila diestimasi dengan menggunakan metode kwadrat terkecil (ordinary least square) dikenal Pooled OLS Regression. Manurung (2002) membangun sebuah model logit pilihan untuk masuknya saham dalam portofolio pada periode 1996 – 2000. Penelitian tersebut memberikan hasil perhitungan model logit setiap tahun selama periode dan

juga pooled data untuk periode 1996 sampai 2000. Adapun model hasil penelitiannya sebagai berikut:

Tabel 1: Hasil Logit Untuk masing-masing variabel dalam Pemilihan saham

Tahun	Konstan	Aset	DER	Dividen Yield	Likuiditas	Kapitalisasi Pasar	Jumlah Saham	PBV	PER	Return	ROE	ROIC	Volatilitas
1995	-44.98095 (-3.3810*)	2.37499 (2.9957*)	-0.49571 (-2.56356*)	-5.18921 (-0.59223)	-0.51725 (-0.341705)	-1.36977 (-2.0975**)	N A	0.91112 (1.5965****)	-0.02433 (-0.61389)	1.13498 (1.2457)	-0.005977 (-0.07136)	5.759389 (1.91834***)	-3.14761 (-1.35983)
1996	-96.23042 (-3.4247*)	4.8942 (3.1706*)	-2.08828 (-2.5981*)	-41.13014 (-1.59278)	-1.44994 (-0.89029)	-2.80447 (-2.69354*)	N A	2.34442 (2.6305*)	-0.04325 (-1.16826)	-3.14107 (-2.6046*)	0.055574 (0.62573)	6.5449 (2.1759**)	1.49134 (0.51865)
1997	-24.31514 (-2.74223*)	0.07346 (0.25007)	-0.16804 (-1.26747)	-3.75884 (-0.71702)	-3.071254 (-2.95478*)	0.80504 (2.59106*)	N A	0.24497 (1.00043)	-0.0097 (-0.6914)	-1.11286 (-1.27263)	0.00025 (0.05391)	0.020793 (0.07116)	1.17384 (0.91006)
1998	-21.9656 (-3.687*)	0.77187 (2.359**)	0.04206 (1.4389)	-42.6767 (-1.930***)	0.05567 (0.18836)	-0.012031 (-0.04234)	-0.02507 (-0.09767)	0.73263 (1.8994**)	0.00534 (1.0319)	0.10763 (0.8023)	0.00355 (1.80768***)	0.03862 (0.7023)	-0.87577 (-1.5372****)
1999	-28.3595 (-3.7125*)	-0.4778 (-1.06241)	0.03441 (0.59779)	-1.91906 (-0.23328)	0.44512 (1.5971****)	1.51878 (3.38999*)	0.0874 (0.22994)	0.00535 (0.0709)	0.00226 (1.9289***)	0.38677 (2.8368*)	0.00189 (0.35484)	-0.01619 (-0.27287)	-3.47453 (-3.2885*)
2000	-43.5946 (-4.0934*)	-0.67502 (-1.3882)	-0.18959 (-1.54103****)	0.18889 (0.07664)	0.88161 (1.8791***)	2.39943 (3.6331*)	-0.007227 (-0.027428)	-0.61833 (-1.2323)	-0.06212 (-2,5182**)	-0.57272 (-0.56941)	-0.012752 (-1.6289****)	-0.35401 (-0.51664)	-4.3633 (-2.2416**)
Seluruhnya	-1.76931 (-6.39587*)	-0.15073 (-4.30421*)	-0.03756 (-1.79810***)	-3.21618 (-1.46716)	-0.49173 (-2.76923*)	0.25638 (5.39248*)	N A	0.155198 (2.74221*)	0.001568 (1.93749***)	0.079404 (1.48975****)	-0.000331 (-0.66802)	0.007456 (0.227622)	-1.6861 (-5.8448*)

Catatan: Angka dalam kurung merupakan nilai Z test  
 \*) Signifikan pada tingkat 1%  
 \*\*) Signifikan pada tingkat 5%  
 \*\*\*) Signifikan pada tingkat 10%  
 \*\*\*\*) Signifikan pada tingkat 15%

Sumber: Adler H. Manurung (2002)

Pada Tabel 1 terlihat secara jelas adanya model pemilihan saham dipengaruhi oleh beberapa peubah bebas untuk setiap tahunnya dan datanya dijadikan satu tanpa memandang tahun aka diperoleh model yang ditunjukkan oleh baris terakhir. Tetapi, data panel tidak memperlihatkan seperti pada Tabel 1 tersebut dan estimasi koefisien akan diuraikan selanjutnya.

## Estimasi Parameter Model Data Panel

Gujarati dan Porter (2009) menyatakan pilihan yang mungkin dilakukan untuk teknik mengestimasi model data panel sebagai berikut:

- Model Pooled OLS – semua sampel observasi misalkan, perusahaan, dikumpulkan menjadi satu dimana tindakan ini menghilangkan cross-section dan runtun waktu. Model ini sering juga disebut Model koefisien konstan (Constant Coefficient Model).
- Model Efek Tetap Kwadrat Terkecil Variabel Boneka (Fixed Effects Least Squares Dummy Variable Model – LSDV) - semua data digabungkan tetapi diijinkannya cross-section memiliki peubah boneka.
- Model Efek Tetap dalam kelompok (Fixed Effects Within-Group Model) – semua data digabungkan tetapi unit observasi (objek) mengekspresikan setiap peubah sebagai selisih dari rata-ratanya dan diestimasi menggunakan metoda kwadrat terkecil.
- Model Efek Acak (Random Effects Model) – seperti model LSDV dimana setiap observasi memiliki nilai intersep dan diasumsikan diambil secara acak dari populasi yang lebih besar.

Selanjutnya, pembahasan akan dilakukan satu-persatu sesuai pilihan yang diberikan sebelumnya.

### ***Model Koefisien Konstan***

Hill dkk (2012) menyatakan model koefisien konstan atau pooled sebagai berikut:

$$Y_{it} = \beta_1 + \beta_2 X_{2it} + \beta_3 X_{3it} + \mu_{it} \quad (2)$$

$i = 1, 2, \dots, k$   
 $t = 1, 2, \dots, n$

Pada persamaan (2) terdapat ada 2 subscript yaitu  $i$  dan  $t$ , dimana  $i$  menyatakan observasi (cross-section) dan  $t$  menyatakan waktu.  $Y_{it}$  menyatakan variabel  $Y$  pada periode ke  $t$  untuk observasi  $i$ . Koefisien  $\beta$  tidak mempunyai subscript sehingga nilainya tetap untuk semua periode dan observasi, sehingga tidak terjadi heterogenitas untuk setiap individu.  $\mu_{it}$  mempunyai nilai rata-rata nol dan konstan varians, dimana tidak berkorelasi dengan waktu  $t$  dan observasi  $i$  dan juga tidak berkorelasi dengan  $X_2$  dan  $X_3$ . Hasil estimasi model dapat diperhatikan pada Tabel 1 yang ditunjukkan sebelumnya.

### ***Model Efek Tetap (MET)***

Gujarati (2011) menyatakan model efek tetap yang diestimasi dengan LSDV memberikan argumentasi bahwa terjadi heterogenitas diantara observasi sehingga mengijinkan intersep berbeda setiap observasi dimana modelnya ditulis sebagai berikut:

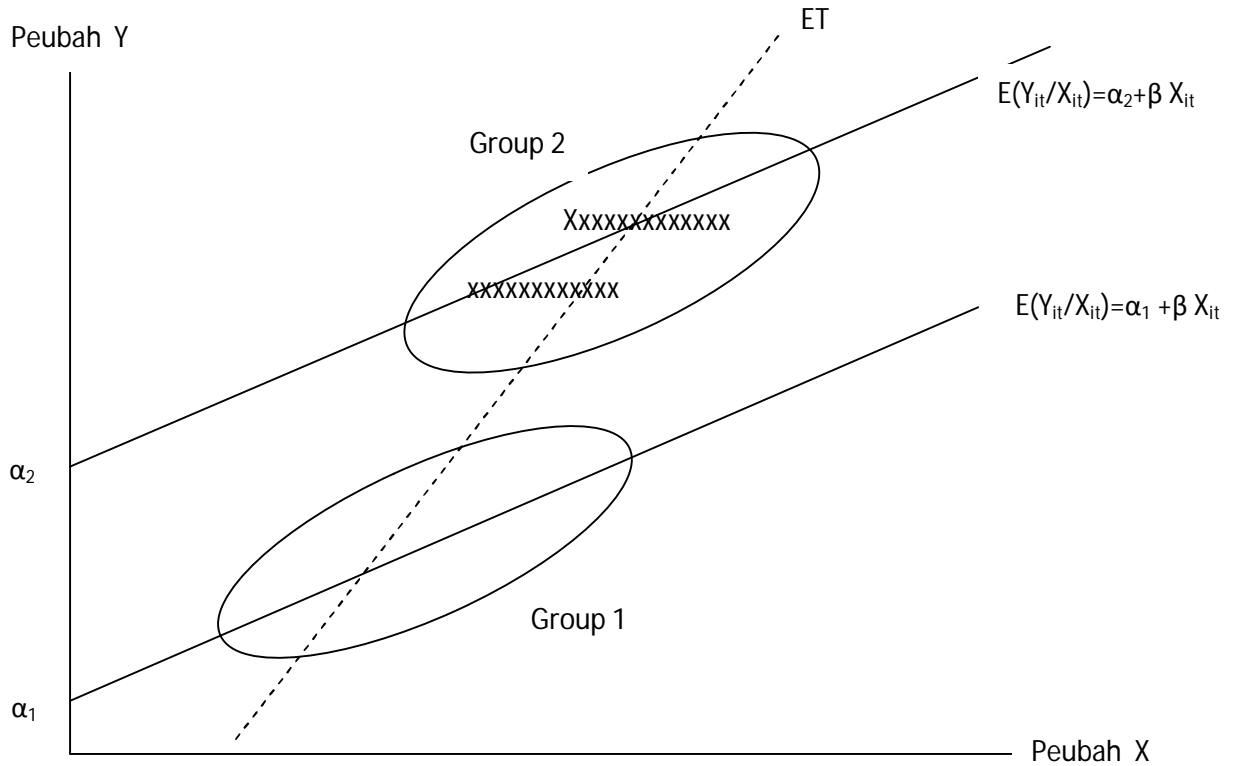
$$Y_{it} = \beta_{1i} + \beta_2 X_{2it} + \beta_3 X_{3it} + \mu_{it} \quad (3)$$

$i = 1, 2, \dots, k$   
 $t = 1, 2, \dots, n$

Persamaan (3) berbeda dengan persamaan (2) ditunjukkannya adanya subscript pada intersep yang menyatakan bahwa terjadi perbedaan intersep untuk setiap observasi penelitian. Perbedaan intersep pada observasi dikarenakan karakteristik observasi berbeda dari satu ke yang lainnya, misalnya karakteristik industri dan gaya manajemen. Bila sampel data observasi yang dipergunakan di Bursa Efek Indonesia, maka terjadi karakteristik perusahaan yang berbeda, gaya manajemen yang berbeda, pasar yang berbeda, dan sebagainya sehingga perbedaan ini dinyatakan oleh intersep tersebut. Gujarati dan Porter (2009) menyatakan bahwa pernyataan "Efek Tetap" menyatakan adanya perbedaan intersep tetapi intersep setiap observasi tidak berbeda selama periode runtun waktu penelitian yang dikenal dengan *Time-Invariant*. Bila intersep ditulis dengan menambahkan subscript kedua yaitu  $t$ ,  $\beta_{it}$ , maka intersep berbeda untuk setiap observasi dan juga periode runtun waktu penelitian dikenal dengan *Time-Variant*. Pada persamaan (3) juga dinyatakan bahwa slope model tetap untuk setiap observasi dan juga periode runtun waktu penelitian dimana besarnya  $\beta_2$  dan  $\beta_3$ .

Dalam kerangka memahami Efek Tetap dengan metoda LSDV ini perlu diperlihatkan perbedaannya dengan metode pooled yang diuraikan sebelumnya. Penjelasannya dibuat dalam bentuk gambar yang menyatakan ada dua kelompok data dengan berbeda waktu ditunjukkan oleh Gambar 2 berikut dibawah ini.

Gambar 2: Bias akibat Penghilang Efek Tetap



Pada Gambar 2 diperlihatkan adanya dua kelompok dimana masing-masing kelompok dibuat model estimasinya, peubah X mempengaruhi peubah Y. Group 1 mempunyai intersep sebesar  $\alpha_1$  dengan persamaan model estimasi  $E(Y_{it} / X_{it}) = \alpha_1 + \beta * X_{it}$  dan group 2 mempunyai intersep sebesar  $\alpha_2$  dengan persamaan model estimasi  $E(Y_{it} / X_{it}) = \alpha_2 + \beta * X_{it}$ . Kedua model kelihatan dibuat mempunyai slope yang sama. Selanjutnya, bila dibuat model persamaan dengan menggabungkan kedua kelompok data maka diperoleh garis ET. Adapun garis ET akan mempunyai intersep dan slope yang sama untuk setiap observasi dan waktu. Adapun group pada gambar menyatakan waktu. Garis ET memperlihatkan slope yang bias dikarenakan melupakan adanya Efek Tetap. Untuk menyatakan adanya perbedaan intersep untuk setiap observasi maka dalam persamaan 3 dibuat tambahan peubah boneka sebanyak observasi atau sampel, misalkan dua observasi setiap tahun sehingga  $D_1 = 1$  untuk observasi satu dan nol yang lain;  $D_2 = 1$  untuk observasi dua dan nol yang lain, sehingga persamaannya sebagai berikut:

$$Y_{it} = \alpha_1 + \alpha_2 D_{1i} + \alpha_3 D_{3i} + \beta_2 X_{2it} + \beta_3 X_{3it} + \mu_{it} \quad (4)$$

Tabel berikut dibawah ini memperlihatkan hasil Eviews dengan model efek tetap untuk penelitian determinan struktur kapital di Indonesia (Manurung, 2011). Pada Tabel terlihat banyak observasi 28 perusahaan sehingga muncul 28 nilai setelah model untuk menjadi tambahan intersep.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.613037	0.063419	9.66638	0
CRTO?	0.002554	0.006227	0.410121	0.6822
WCTA?	-0.374613	0.171164	-2.188619	0.0299
GRW?	-0.003585	0.017966	-0.199529	0.8421
MBV?	-0.004604	0.002438	-1.888384	0.0605
RETA?	-0.192189	0.10839	-1.773124	0.0778
EBITTA?	-0.090442	0.042637	-2.121203	0.0352
ROA?	0.05822	0.086698	0.671527	0.5027
STA?	0.038591	0.03779	1.021179	0.3085
Fixed Effects (Cross)				
_1--C	-0.036559			
_2--C	0.178169			
_3--C	-0.328891			
_4--C	0.307517			
_5--C	0.26492			
_6--C	-0.045454			
_7--C	-0.022786			
_8--C	0.193655			
_9--C	-0.101319			
_10--C	-0.147433			
_11--C	-0.054519			
_12--C	0.041227			
_13--C	0.081581			
_14--C	0.018352			
_15--C	-0.019524			
_16--C	0.133517			
_17--C	-0.140502			
_18--C	0.15576			
_19--C	-0.102498			
_20--C	-0.077592			
_21--C	-0.135338			
_22--C	-0.144258			
_23--C	-0.023106			
_24--C	0.014026			
_25--C	-0.017009			
_26--C	0.039776			
_27--C	-0.17596			
_28--C	0.144249			
Effects Specification				
Cross-section fixed (dummy variables)				
R-squared	0.71155	Mean dependent var	0.504004	
Adjusted R-squared	0.657849	S.D. dependent var	0.226371	
S.E. of regression	0.132413	Akaike info criterion	-1.059562	
Sum squared resid	3.296228	Schwarz criterion	-0.511261	
Log likelihood	154.6709	Hannan-Quinn criter.	-0.838241	
F-statistic	13.25024	Durbin-Watson stat	1.318988	
Prob(F-statistic)	0			

Sumber: Adler H. Manurung (2011).

Adapun nilai intersep untuk observasi pertama yaitu nilai  $c$  pada table ditambah dengan nilai fixed effect  $1\_C$  yang besarnya  $0,613037 + (-0,036559) = 0,576478$ . Sedangkan nilai intersep pada observasi ke-28 yaitu  $0,613037 + 0,144249 = 0,757286$ , demikian juga perhitungan untuk intersep pada observasi yang lain. Artinya, nilai intersep tidak sama setiap observasi dan bila terjadi hanya merupakan kebetulan.

### **Model Efek Random (MER)**

Greene (2012) menyatakan model Efek Random dinyatakan sebagai berikut:

$$Y_{it} = (\beta_1 + \varepsilon_i) + \beta_2 X_{2it} + \beta_3 X_{3it} + \mu_{it} \quad (5)$$

$i = 1, 2, \dots, k$   
 $t = 1, 2, \dots, n$

Persamaan (5) memperlihatkan bahwa intersep dari model terdiri dari koefisien tetap sebesar  $\beta_1$  ditambah  $\varepsilon_i$  yang menyatakan adanya keacakan untuk pada setiap observasi mempresentasikan karakteristik observasi, diman  $\varepsilon_i$  memiliki rata-rata dan varians  $\sigma_\varepsilon^2$ . Akibatnya, keacakan dari persamaan (5) menjadi  $\varepsilon_i + \mu_{it}$ . Akibatnya persamaan (5) dapat ditulis kembali sebagai berikut:

$$Y_{it} = \beta_1 + \beta_2 X_{2it} + \beta_3 X_{3it} + \omega_{it} \quad (6)$$

dimana

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

Komponen kesalahan  $\omega_{it}$  mempunyai dua komponen kesalahan yaitu kesalahan cross-section  $\varepsilon_i$  atau kesalahan spesifik individu dan  $\mu_{it}$  kesalahan gabungan observasi dan periode runtun waktu. Sehingga, Model Efek Randon ini disebut juga Model Komponen Kesalahan (Error Compponnets Model - ERM). Adapun asumsi pada ERM sebagai berikut:

$$\left. \begin{aligned} \varepsilon &\sim N(0, \sigma_\varepsilon^2) \\ \mu_{it} &\sim N(0, \sigma_\mu^2) \\ E(\varepsilon_i \mu_{it}) &= 0; E(\varepsilon_i \varepsilon_j) = 0 \quad (i \neq j) \\ E(\mu_{it} \mu_{is}) &= E(\mu_{it} \mu_{ij}) = E(\mu_{it} \mu_{js}) = 0 \quad (i \neq j; t \neq s) \end{aligned} \right\} \quad (7)$$

Sebagai catatan bahwa  $\omega_{it}$  tidak berkorelasi dengan setiap perubah penjelas (explanatory variables). Akibat  $\varepsilon_i$  merupakan komponen  $\omega_{it}$  ada kemungkinan akan berkorelasi dengan satu atau lebih peubah bebas. Oleh karenanya, REM bukanlah estimator yang consistent untuk koefisien model. Adanya persamaan (7) maka diperoleh persamaan sebagai berikut:

$$E(\omega_{it}) = 0 \quad (8)$$



$$\text{Var}(\omega_{it}) = \sigma_{\varepsilon}^2 + \sigma_{\mu}^2 \quad (9)$$

Bila nilai  $\sigma_{\varepsilon}^2 = 0$  maka persamaan (9) sama persis dengan persamaan (1) yang menyatakan bahwa modelnya adalah model Pooled.

Persamaan (9) juga menyatakan adanya varians homoskedastis, dapat ditunjukkan dengan  $\omega_{it}$  dan  $\omega_{is}$  ( $t \neq s$ ) berkorelasi, maka kesalahan dari sebuah unit cross-section tertentu pada dua waktu yang berbeda berkorelasi. Korelasi antara keduanya ditunjukkan sebagai berikut:

$$\rho = \text{corr}(\omega_{it}, \omega_{is}) = \frac{\sigma_{\varepsilon}^2}{\sigma_{\varepsilon}^2 + \sigma_{\mu}^2}; t \neq s \quad (10).$$

Gujarati (2011) menyatakan bahwa pertama, untuk setiap unit cross-section  $p$  tetap sama tidak memandang jarak dua waktu tersebut, dan kedua  $\rho$  tetap sama untuk semua unit cross-section. Estimator dari REM yang efisien harus menggunakan Metoda Kwadrat Terkecil Umum (Generalized Least Square). Adapun hasil empiris model REM ini (Manurung, 12) sebagai berikut:

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	6.887726	0.398706	17.27521	0.0000
DER?	-0.013629	0.015247	-0.893876	0.3722
DM	-0.518765	0.158784	-3.267118	0.0012
Random Effects (Cross)				
_1--C	0.043860			
_2--C	-0.973641			
_3--C	1.353276			
_4--C	0.191824			
_5--C	2.130803			
_6--C	-2.814941			
_7--C	0.460287			
_8--C	-0.148238			
_9--C	-0.861065			
_10--C	-0.469848			
_11--C	-1.312075			
_12--C	1.096430			
_13--C	2.195355			
_14--C	-0.892028			
Weighted Statistics				
R-squared	0.044377	Mean dependent var		1.151709
Adjusted R-squared	0.037477	S.D. dependent var		1.154009
S.E. of regression	1.132178	Sum squared resid		355.0659
F-statistic	6.431660	Durbin-Watson stat		0.343815
Prob(F-statistic)	0.001861			

Pada Tabel diatas terlihat secara jelas bahwa ada pengaruh keacakan sehingga intersep dari perusahaan pertama sebesar 6,931586 (6,887726 + 0,043860) dan intersep pada perusahaan ke-14 yaitu 5,9955698 (6,887726 – 0,892028).

## Pengujian Pemilihan Model

Uraian sebelumnya telah memberikan penjelasan mengenai Model Pooled, Model Efek Tetap (MET) dan Metode Efek Random (MER). Tetapi, timbul pertanyaan apakah ketiga model harus diestimasi atau salah satu dari ketiga model tersebut. Dalam membuat penelitian, harus satu model yang diperlihatkan dan diuraikan sehingga diperlukan pengujian. Adapun pengujian yang selalu dipergunakan untuk menggunakan Pooled, MET atau MER dikenal dengan Pengujian Model. Pengujian Model Pooled dibandingkan dengan model MET mempunyai hipotesis sebagai berikut:

$$\begin{aligned} H_0 &: \text{Model Pooled lebih baik dari MET} \\ H_1 &: \text{Model MET lebih baik dari Model Pooled} \end{aligned}$$

Pengujian ini dikenal dengan pengujian Chow dimana Statistik ujinya mengikuti distribusi F sebagai berikut:

$$F = \frac{(R_{MET}^2 - R_{Pol}^2)/(m-1)}{(1 - R_{FEM}^2)/(n-k-1)} \quad (11)$$

dimana

- m adalah jumlah cross-section
- n adalah jumlah seluruh observasi
- k adalah jumlah parameter yang diestimasi

Bila F statistik lebih besar dari F tabel dimana numeratornya (m-1) dan denominatornya (n-k-1) maka H<sub>0</sub> ditolak yang menyatakan bahwa nilai intersep tidak sama sehingga model yang dipergunakan model MET.

Kemudian, pemilihan pemilihan model Pooled dengan MER dilakukan pengujian yang dikenal dengan pengujian hipotesis sebagai berikut:

$$\begin{aligned} H_0 &: \text{Model Pooled lebih cocok} \\ H_1 &: \text{MER lebih cocok} \end{aligned}$$

Adapun uji statistiknya sebagai berikut:

$$LM = \sqrt{\frac{NT}{2(T-1)}} \left\{ \frac{\sum_{i=1}^N \left( \sum_{t=1}^T \hat{e}_{it} \right)^2}{\sum_{i=1}^N \sum_{t=1}^T \hat{e}_{it}} - 1 \right\} \quad (12)$$

Uji statistik tersebut mengikuti Distribusi Normal (0,1), sehingga diperbandingkan dengan  $Z_\alpha$ , bila nilai LM lebih besar dari nilai  $Z_\alpha$  maka tolak H<sub>0</sub> sehingga penolakan H<sub>0</sub> menyatakan menerima MER yang cocok.

Selanjutnya, pemilihan MET dan MER perlu dilakukan pengujian dimana pengujiannya dikenal dengan pengujian Hausman yang mempunyai hipotesa sebagai berikut:

- $H_0$  : MER yang paling cocok  
 $H_1$  : MET yang paling cocok

Adapun uji statistiknya sebagai berikut:

$$\chi^2_{df} = \frac{(\beta_{MER} - \beta_{MET})^2}{\sigma_{MER-MET}} \quad (12)$$

Bila  $\chi^2$  observasi ini lebih tinggi dari  $\chi^2$  tabel dengan derajat kebebasannya jumlah variabel bebas, maka hipotesis nol ditolak yang menyatakan bahwa MET yang cocok dipergunakan dalam model tersebut. Bentuk pengujian Hausman tersebut dapat diperhatikan pada Tabel berikut:

Correlated Random Effects - Hausman Test  
 Pool: PML01  
 Test cross-section random effects

Test Summary	Chi-Sq. Statistic	Chi-Sq. d.f.	Prob.
Cross-section random	14.726586	8	0.0647

Bila dilihat Tabel  $\chi^2$  dengan df 8 diperoleh nilai sebesar 13,362 dimana level signifikansi 10% sehingga  $H_0$  ditolak dan model yang tepat yaitu Model Efek Tetap. Salah satu cara lain melihat diterima atau ditolaknya hipotesis dengan memperbandingkan level signifikansi dengan probabilitas yang tercantum pada Tabel. Adapun bentuk keputusannya sebagai berikut:

- $\alpha > \text{Prob}$  -----  $> H_0$  ditolak  
 $\alpha < \text{Prob}$  -----  $> H_0$  diterima

Pada kasus diatas  $\alpha = 10\%$  lebih kecil dari  $\text{Prob} = 6,47\%$  maka  $H_0$  ditolak.

Pemilihan MET dan MER ini juga dapat menggunakan kaidah sebagai berikut:

- Jika T (jumlah periode waktu) sangat besar dan N (jumlah cross-section) sangat kecil maka sangat kecil perbedaan nilai parameter yang diestimasi dengan MET dan MER tetapi lebih disukai menggunakan MET
- Pada Panel Pendek (N besar dan T kecil) dan terjadi perbedaan yang cukup besar pada estimasi, lebih disukai (tepat) menggunakan MET.
- N sangat besar dan T sangat kecil, dimana asumsi REM tetap ada maka MER lebih tepat sebagai estimator dibandingkan MET.
- Jika komponen kesalahan individu  $\epsilon_i$  dan salah satu peubah penjelas (regressor) berkorelasi maka estimasi dengan MER akan bias sehingga lebih tepat menggunakan MET.

## Daftar Pustaka

Agung, I Gusti N. (2011); Cross Section and Experimental Data Analysis; John Wiley & Sons

Agung, I Gusti N. (2009); Time Series Data Analysis: Using Eviews; John Wiley & Sons

Alessie, Rob and Annamaria Lusardi (1997); Saving and income smoothing: Evidence from panel data; European Economic Review, Vol. 41; pp. 1251-1279.

Apergis, N.; Filippidis, I. and C. Economidou (2007); Financial Deepening and Economic Growth Linkages: A Panel Data Analysis; Review of World Economics, Vol. 143, No. 1; pp. 179 – 198.

Arellano, Manuel (2003); Panel Data Econometrics: Advanced Texts in Econometrics; Oxford University Press

Asteriou, D. and S. G. Hall (2007); Applied Econometrics: A Modern Approach; Palgrave Macmillan

Batalgi, Badi H. (1999); Econometrics; 2<sup>nd</sup> eds.; Springer

Batalgi, Badi H. (1995); Econometric Analysis of Panel Data; John Wiley & Sons

Cameron, A. C. and P. K. Trivedi (2008); Microeconometrics: Methods and Applications; Cambridge University Press.

Chamberlain, G. (1997); Panel Data; in Zvi Griliches and M. D. Intriligator; Handbook of Econometrics, Volume 2; North-Holland.

Davidson, Russell and J. G. Mackinnon (2004); Econometric Theory and Methods; Oxford University Press

Egger, Peter (2004); Estimating Regional Trading Bloc Effects with Panel Data; Review of World Economics, Vol. 140, No. 1; pp. 151 – 166.

Egger, P.; Pfaffermayr, M.; and Hannes Winner (2005); An unbalanced spatial panel data approach to US state tax competition; Economics Letters, Vol. 88; pp. 329 – 335.

Gaulier, G.; Hurlin, C. and P. Jean-Pierre (1999); Testing Convergence: A Panel Data Approach; Annals of Economics and Statistics, No. 55/56; pp. 411 – 427.

Goeltom, Miranda S. (1995); Indonesia's Financial Liberalization: An Empirical Analysis of 1981 – 1988 Panel Data; Institute of Southeast Asian Studies.

Gonzalez, Rosa M. and Gustavo A. Marrero (2012); Induced road traffic in Spanish Regions: A dynamic panel data model; Transportation Research Part A, Vol. 46; pp. 435 – 445.

Greene, William H. (2012); *Econometric Analysis*, 7<sup>th</sup> Eds., Pearson

Gujarati, Damodar (2011); *Econometrics by Example*; Palgrave Macmillan.

Gujarati, Damodar and Dawn C. Porter (2009); *Basic Econometrics*; 5<sup>th</sup> eds.; Palgrave Macmillan.

Hayashi, Fumio (2000); *Econometrics*; Princeton University Press

Heshmati, A.; Kumbhakar, S. C. and Lennart Hjalmarsson (1995); Efficiency of the Swedish pork industry: A farm level study using rotating panel data 1976-1988; *European Journal of Operational Research*, Vol. 80; pp. 519-533

Hill, R. C.; Griffith, W. E and G. C. Lim (2012); *Principles of Econometrics*; 4<sup>th</sup> eds.; John Wiley & Sons.

Honore, B. E and E. Kyriazidou (2000); Panel Data Discrete Choice Models with Lagged Dependent Variables; *Econometrica*, Vol. 68, No. 4; pp. 839 – 874.

Hsiao, Cheng (1999); *Analysis of Panel Data*; Cambridge University Press.

Johnston, J. and J. Dinardo (1997); *Econometric Methods*; 4<sup>th</sup> eds.; McGraw Hill.

Joreskog, K. G. (1978); An Econometric Model for Multivariate Panel Data; *Econometrics of Panel Data*, Apr – Sep; pp. 355 – 366.

Judge, G. G.; Hill, R. C.; Griffiths, W. E.; Lutkepohl, H. and T. C. Lee (1988); *Introduction to the Theory and Practice of Econometrics*; 2<sup>nd</sup> eds.; John Wiley & Sons.

\_\_\_\_\_ (1985); *The Theory and Practice of Econometrics*; 2<sup>nd</sup> eds.; John Wiley & Sons.

Judson, R. A. and A. L. Owen (1999); Estimating Dynamic Panel Data Models: A Guide for Macroeconomics; *Economics Letters*, Vol. 65; pp. 9 -15.

Kiviet, Jan F. (1995); On bias, inconsistency, and efficiency of various estimators in dynamic panel data models; *Journal of Econometrics*, Vol. 68; pp. 53-78

Koutsoyiannis, A. (1977); *Theory of Econometrics*; Barnes and Nobles Book.

Lancaster, Tony (2002); Orthogonal Parameters and Panel Data; *Review of Economic Studies*, Vol. 69, No. 3; pp. 647 – 666.

Lee, E. W. and M. Y. Kim (1998); The Analysis of Correlated Panel Data Using a Continuous-Time Markov Model; *Biometrics*, Vol. 54, No. 4; pp. 1638 – 1644.

Li, Qi and T. Stengos (1996); Semiparametric Estimation of Partially Linear Panel Data Models; *Journal of Econometrics*, Vol. 71; pp. 389 – 397.

Lopez-Pueyo, C.; Barcenilla-Visus, S. and J. Sanau (2008); International R&D spillovers and manufacturing productivity: A panel data analysis; *Structural Change and Economic Dynamics*, Vol. 19; pp. 152–172

Lott, W. F. and S. C. Ray (1992); *Applied Econometrics: Problems with Data Sets*; The Dryden Press.

Maddala, G. S. (1993); *The Econometrics of Panel Data; Volume 1*; Edward Elgar Publishing.

Mainardi, Stefano (2005); Earnings and work accident risk: a panel data analysis on mining; *Resources Policy*, Vol. 30; pp. 156–167.

Manurung, Adler H. (2011); Determinan Struktur Pasar Kapital Perusahaan di Indonesia; *Jurnal Akuntansi FE UNTAR*, Vol. 15, No. 3; pp. 250 – 261.

Manurung, Adler H. (2002); Konsistensi Pemilihan Saham dalam Pembentukan Portofolio Optimal di BEJ oleh Manajer Investasi dikaitkan dengan Variabel Rasio Empiris Kinerja Perusahaan; Disertasi Tidak Dipublikasikan, Pascasarjana FEUI.

Marin, E.; Perez-Amaral, T.; Rua, A. and E. Hernandez (2001); The Evolution of the pH in Europe (1986 – 1997) using Panel Data; *Chemosphere*, Vol. 45; pp. 329 – 337.

Nerlove, Marc (2002); *Essay in Panel Data Econometrics*; Cambridge University Press.

Niu, S.; Ding, Y.; Niu, Y.; Li, Y.; and GuanghuaLuo (2011); Economic growth, energy conservation and emissions reduction: A comparative analysis based on panel data for 8 Asian-Pacific countries; *Energy Policy*, Vol. 39; pp. 2121–2131.

Phillips, P. C. B. and H. R. Moon (1999); Linear Regression Limit Theory for Nonstationary Panel Data; *Econometrica*, Vol. 67, No. 5; pp. 1057 – 1111.

Ramanathan, Ramu (1995); *Introductory Econometrics*; 3<sup>rd</sup> eds.; The Dryden Press.

Stock, J. H. and M. W. Watson (2003); *Introduction to Econometrics*; Addison Wesley.

Tsay, Ruey S.; Tomohiro Ando (2012); Bayesian panel data analysis for exploring the impact of subprime financial crisis on the US stock market; *Computational Statistics and Data Analysis*, Vol. 56; pp. 3345–3365

Verbeek, Marno (2008); *A Guide to Modern Econometrics*; 3<sup>rd</sup> eds.; John Wiley & Sons.

Vogelvang, Ben (2005); *Econometrics: Theory and Applications with Eviews*; Prentice Hall.

Wansbeek, Tom (2001); GMM estimation in panel data models with measurement error; *Journal of Econometrics*, Vol. 104; pp. 259 – 268.

Wansbeek, T. and Paul Bekker (1996); On IV, GMM and ML in a dynamic panel data model; *Economics Letters* 51 (1996) 145-152.

Wooldrige, J. M (2006); *Introductory Econometric: A Modern Approach*; 2<sup>nd</sup> eds.; Thomson South Western.

Wooldrige, J. M (2002); *Econometric Analysis of Cross Section and Panel Data*; The MIT Press

Wooldrige, J. M (1995); Selection Corrections for Panel Data Models under conditional Mean Independence Assumptions; *Journal of Econometrics*, Vol. 68; pp. 115 – 132.

Zhang, Ying (2002); A Semiparametric Pseudolikelihood Estimation Method for Panel Count Data; *Biometrika*, Vol. 89, No. 1; pp.. 39 – 48.