TESIS

ESTIMASI PARAMETER MODEL PERSAMAAN SIMULTAN MENGGUNAKAN METODE *TWO STAGE RIDGE REGRESSION* DAN *TWO STAGE RIDGE LIU*

SIMULTANEOUS EQUATION MODEL PARAMETER ESTIMATION USING TWO STAGE RIDGE REGRESSION AND *TWO STAGE RIDGE LIU METHOD*

SARINA

H062192008

PROGRAM STUDI MAGISTER STATISTIKA

DEPARTEMEN STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS HASANUDDIN

MAKASSAR

2023

ESTIMASI PARAMETER MODEL PERSAMAAN SIMULTAN MENGGUNAKAN METODE *TWO STAGE RIDGE REGRESSION* DAN *TWO STAGE RIDGE LIU*

SIMULTANEOUS EQUATION MODEL PARAMETER ESTIMATION USING TWO STAGE RIDGE REGRESSION AND *TWO STAGE RIDGE LIU METHOD*

SARINA

PROGRAM STUDI MAGISTER STATISTIKA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR
2023

ESTIMASI PARAMETER MODEL PERSAMAAN SIMULTAN MENGGUNAKAN METODE *TWO STAGE RIDGE REGRESSION* DAN *TWO STAGE RIDGE LIU*

Tesis

sebagai salah satu syarat untuk mencapai gelar magister

Program Studi Magister Statistika

Disusun dan diajukan oleh

SARINA

H062192008

kepada

PROGRAM STUDI MAGISTER STATISTIKA

DEPARTEMEN STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS HASANUDDIN

MAKASSAR

2023

TESIS

ESTIMASI PARAMETER MODEL PERSAMAAN SIMULTAN MENGGUNAKAN METODE *TWO STAGE RIDGE REGRESSION* DAN *TWO STAGE RIDGE LIU*

SARINA

H062192008

Telah dipertahankan di hadapan Panitia Ujian yang dibentuk dalam rangka Penyelesaian Program Studi Magister Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin
pada tanggal 01 Maret 2023
dan dinyatakan telah memenuhi syarat kelulusan

Menyetujui,

Pembimbing Utama,

Dr. Nurtiti Sunusi, S.Si M.Si.
NIP. 1972017 199703 2 002

Pembimbing Pendamping

Dr. Erna Tri Herdiani, S.Si M.Si.
NIP. 19750429 200003 2 001

Ketua Program Studi
Magister Statistika,

Dr. Dr. Georgina M. Tinungki, M.Si
NIP. 19620926 198702 2 001

Dekan Fakultas Matematika dan
Ilmu Pengetahuan Alam
Universitas Hasanuddin

KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET DAN TEKNOLOGI
UNIVERSITAS HASANUDDIN

Dr. Eng. Amiruddin, M.Si
NIP. 19720515 199702 1 002

PERNYATAAN KEASLIAN TESIS DAN PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa, tesis berjudul Estimasi Parameter Model Persamaan Simultan Menggunakan Metode *Estimasi Parameter Model Persamaan Simultan Menggunakan Metode Two Stage Ridge Regression Dan Two Stage Ridge Liu* adalah benar karya saya dengan arahan dari komisi pembimbing (Dr. Nurtiti Sunusi, S.Si., M.Si dan Dr. Erna Tri Herdiani,S.Si., M.Si). Karya ilmiah ini belum diajukan dan tidak sedang diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka tesis ini. Sebagian dari isi tesis ini telah dipublikasikan di Jurnal (International Journal of Research Publications (IJRP) ISSN: 2708-3578 Vol. 105 Issue 1, Juli - 2022, Pages:106-116) sebagai artikel dengan judul "An Estimator For Simultaneous Equation Model Using Two Stage Ridge Method (Case Study On Farmer Exchange Rate Data In Indonesia)".

Dengan ini saya melimpahkan hak cipta dari karya tulis saya berupa tesis ini kepada Universitas Hasanuddin.

Makassar, 01 Maret 2023

Yang Menyatakan,

Sarina

NIM. H062192008

UCAPAN TERIMA KASIH

Segala puji bagi Allah *Subhanahu Wa ta'ala Rabb* semesta alam, shalawat serta salam semoga senantiasa tercurahkan kepada Nabi yang paling dimuliakan, pemimpin orang-orang bertakwa yakni Rasulullah *Shallallahu Alaihi Wasallam* dan kepada para keluarga serta shahabat beliau yang senantiasa kita rindukan perjumpaan dengannya. Amma ba'du.

Puji syukur penulis panjatkan ke hadirat Allah *Subhanahu Wa ta'ala* atas limpahan rahmat dan karunia-Nya sehingga penulis dapat menyusun dan menyelesaikan tesis ini. Penulis juga percaya tesis ini dapat selesai bukan hanya dengan kekuatan pikiran penulis semata akan tetapi karena bantuan dari berbagai pihak juga, baik selama proses perkuliahan bahkan sampai proses pengerjaan tesis di Program Magister Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Hasanuddin.

Penulisan tesis ini dapat terlaksana karena bantuan dan dukungan dari berbagai pihak baik secara langsung maupun tidak langsung, sehingga pada kesempatan ini menulis ingin mengucapkan terima kasih yang tak terhingga kepada kedua Orang Tua tercinta dan saudara atas doa yang tak pernah putus, dukungan serta segala kebaikan mereka yang sampai kapan pun takkan pernah bisa terbalaskan atas kasih sayang yang tiada henti dalam penyelesaian tesis ini.

Selanjutnya, Penghargaan yang tulus dan ucapan terima kasih dengan penuh keikhlasan juga penulis ucapkan kepada:

1. **Prof. Dr. Ir. Jamaluddin Jompa, M.Sc.** selaku Rektor Universitas Hasanuddin.
2. **Dr. Eng. Amiruddin, M.Si**. selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam beserta seluruh jajarannya.
3. **Dr. Nurtiti Sanusi, M.Si.** selaku Ketua Departemen Statistika sekaligus yang menjadi Pembimbing utama yang telah bersabar dan bersedia meluangkan banyak waktunya untuk membimbing penulis dan memberikan masukan dalam penyelesaian tesis ini.
4. **Dr. Dr. Georgina Maria Tinungki, M.Si**. selaku Ketua Program Studi Magister Statistika yang menjadi salah satu tim penguji tesis untuk memberikan masukan dan saran dalam penulisan tesis ini.

5. **Dr. Erna Tri Herdiani, M.Si.** selaku pembimbing akademik sekaligus sebagai pembimbing pertama yang telah bersabar dan bersedia meluangkan banyak waktunya untuk membimbing penulis dan memberikan masukan dalam penyelesaian tesis ini.
6. **Dr. Nirwan Ilyas, M.Si.** selaku salah satu tim penguji tesis sekaligus memberikan ilmu, dukungan, dan motivasi serta kemudahan kepada penulis dalam berbagai hal selama menjalani Pendidikan di Departemen Statistika.
7. **Dr. Anna Islamiyati, S.Si., M.Si.** selaku penguji penulis yang telah bersedia memberikan masukan-masukan dan arahan dalam penyusunan tesis.
8. Bapak dan Ibu Dosen di Departemen Statistika Fakultas MIPA Universitas Hasanuddin, yang dengan tulus ikhlas memberikan ilmu pengetahuan dan pengalaman yang dimilikinya selama perkuliahan berlangsung sehingga memberikan banyak manfaat bagi penulis untuk saat ini maupun di masa mendatang.
9. Teman-teman Mahasiswa Program Magister Statistika khususnya Angkatan 1, terima kasih atas nasehat dan dukungan luar biasa kepada penulis.
10. Serta semua pihak yang telah membantu dalam penyelesaian tesis ini yang tidak dapat penulis sebutkan satu per satu.

Akhirnya dengan segala kerendahan hati, penulis berharap tesis ini dapat bermanfaat bagi pihak-pihak yang membutuhkan dan dapat dijadikan referensi bagi penelitian-penelitian selanjutnya khususnya dalam dunia statistika dan data sains. Penulis juga menyadari bahwa penulisan tesis ini masih jauh dari kesempurnaan dan banyak kelemahan, sehingga penulis tak lupa mengharapkan saran dan kritik atas tesis ini.

Makassar, 01 Maret 2023

Sarina

ABSTRAK

SARINA. **Metode Estimasi Parameter Model Persamaan Simultan Menggunakan Metode *Two Stage Ridge Regression* dan *Two Stage Ridge Liu*** (dibimbing oleh Dr. Nurtiti Sunusi, S.Si., M.Si dan Dr. Erna Tri Herdiani, S.Si., M.Si)

Model persamaan simultan merupakan suatu sistem persamaan yang terdiri dari dua atau lebih persamaan yang saling berkaitan satu dengan yang lain, variabel respon pada satu persamaan dapat berperan sebagai variabel bebas pada persamaan lain, dengan kata lain terdapat hubungan yang simultan dalam sistem persamaan tersebut. Dalam model persamaan simultan ini terdapat dua jenis variabel yaitu variabel endogen yang nilainya ditentukan di dalam model dan variabel eksogen yang nilainya ditetapkan terlebih dahulu di luar model. Adanya simultanitas antar variabel endogen menyebabkan terjadinya korelasi antara variabel respon dengan faktor galat. Salah satu metode alternatif yang dapat digunakan untuk mengestimasi parameter model persamaan simultan yaitu *two stage least square* (2SLS). Metode 2SLS akan dapat mengestimasi parameter dengan baik apabila antara variabel bebasnya tidak terjadi masalah multikolinearitas. Untuk mengatasi multikolinearitas digunakan analisis regresi *ridge* atau dan *two stage ridge liu* regresi liu. Pada penelitian ini, digunakan metode *two stage ridge regression* dan liu pada persamaan simultan. Metode ini kemudian diaplikasikan pada data simultan nilai tukar petani. Dari hasil penelitian ini diperoleh bahwa metode *two stage ridge liu* lebih baik dibandingkan dengan metode *two stage ridge*. Hal ini didasarkan pada nilai MSE untuk setiap metode.

Kata kunci: Model Persamaan Simultan, *Two Stage Least Square*, Multikolinearitas, Regresi Ridge, Regresi Liu

ABSTRACT

SARINA. **Simultaneous Equation Model Parameter Estimation Using Two Stage Ridge Regression And Two Stage Ridge Liu Method** (supervised by Dr. Nurtiti Sunusi, S.Si., M.Si dan Dr. Erna Tri Herdiani, S.Si., M.Si).

Simultaneous equation model is a system of equations consisting of two or more equations that are related to one another, the response variable in one equation can act as an independent variable in other equations, in other words there is a simultaneous relationship between the two equations. in the system of equations. In this simultaneous equation model, there are two types of variables, namely endogenous variables whose values are determined in the model and exogenous variables whose values are determined in advance outside the model. The existence of simultaneity between endogenous variables causes a correlation between the response variables and the residual factors. One alternative method that can be used to estimate the parameters of the simultaneous equation model is two stage least square (2SLS). The 2SLS method will be able to estimate parameters well if there is no multicholinearity problem between the independent variables. To overcome multicollinearity used ridge regression analysis or Liu regression. In this research, *two stage ridge regression* and liu method is used in simultaneous equations. This method is then applied to simultaneous data on farmer exchange rates. From the results of this study, it was found that the two stage ridge liu method based on two bias estimators was better than the *two stage ridge regression* method. It is based on the MSE value for each method.

Keywords: Simultaneous Equation Model, Two Stage Least Square, Multicollinearity, Ridge Regression, Liu Regression.

DAFTAR ISI

Halaman

DAFTAR TABEL

DAFTAR GAMBAR

Nomor Urut **Halaman**

BAB I
PENDAHULUAN

1.1 Latar Belakang

Analisis regresi merupakan suatu model statistik yang digunakan dalam penentuan hubungan sebab akibat antara variabel terikat dengan variabel prediktor. Berdasarkan jumlah persamaan, suatu model dapat dibedakan atas model persamaan tunggal (*single equation model*) dan model persamaan simultan (*simultaneous equation model*). Persamaan tunggal adalah persamaan dimana variabel terikat dinyatakan sebagai sebuah fungsi dari satu atau lebih variabel bebas, sehingga hubungan sebab akibat antara variabel terikat dan variabel bebas merupakan hubungan satu arah (Shina, 2016). Sedangkan model persamaan simultan adalah suatu model yang mengandung lebih dari satu persamaan yang saling mempengaruhi dimana ada beberapa variabel dalam suatu persamaan mempunyai keterkaitan dengan variabel yang sama yang terdapat di dalam persamaan lainnya. Persamaan simultan terdiri atas dua jenis persamaan yaitu persamaan struktural, merupakan persamaan yang berupa suatu fungsi, terdiri dari beberapa variabel dan persamaan bentuk reduksi, yaitu persamaan yang bukan merupakan fungsi, namun hanya persamaan yang terdiri dari penjumlahan beberapa variabel. Variabel-variabel dalam persamaan bentuk reduksi dapat berasal dari variabel dependen pada persamaan struktural, maupun variabel yang berasal dari luar persamaan struktural. Seperti dengan klasifikasi variabel sebagai variabel terikat dan variabel bebas dalam model regresi linier, variabel dalam model persamaan simultan diklasifikasikan sebagai variabel endogen dan variabel eksogen (Shalabh, 2010).

Ordinary least square (OLS) tidak dapat digunakan untuk mengestimasi parameter dalam model persamaan simultan (Gujarati, 2010). Hal ini disebabkan karena salah satu asumsi klasik yaitu non-autokorelasi dilanggar (Falah, Mustafid, & Sudarno, 2016). Estimasi parameter model persamaan simultan dapat diperoleh dengan menggunakan dua metode yaitu metode persamaan tunggal dan metode sistem. Akan tetapi, sebelum melakukan estimasi parameter model persamaaan simultan diperlukan adanya suatu identifikasi model. Identifikasi ini bertujuan untuk menentukan apakah suatu model persamaan simultan dapat dilakukan penaksiran dan mengetahui metode penaksiran apa yang sebaiknya digunakan pada persamaan simultan tersebut (Qingli, 2019).

Hausman telah mengemukakan bahwa dalam mengidentifikasi model persamaan simultan terdapat dua aturan yaitu aturan kondisi orde dan aturan kondisi rank. Dari identifikasi tersebut akan diperoleh tiga kondisi yang mungkin terjadi terhadap model persamaan simultan yaitu: tidak teridentifikasi, teridentifikasi dan teridentifikasi berlebih.

Metode alternatif yang dapat digunakan untuk mengestimasi parameter model persamaan simultan yaitu *two stage least square* (2SLS). *Two stage least square* (2SLS) adalah salah satu metode persamaan tunggal yang termasuk ke dalam kelompok analisis persamaan struktural. Metode ini digunakan dengan kondisi dimana suatu variabel endogen akan menjadi variabel eksogen pada persamaan lain sehingga kemungkinan berkorelasi dengan galat cukup besar (terjadi endogenitas). Beberapa penelitian yang telah menggunakan metode 2SLS diantaranya Soemartini (2016) menggunakan 2SLS pada model persamaan simultan untuk memprediksi PDRB dan pertumbuhan ekonomi, Shina (2018) melakukan estimasi parameter sistem model persamaan simultan pada data panel dengan metode 2SLS, dan Yusuf et al. (2021) menggunakan 2SLS untuk memodelkan persamaan simultan pada kasus inflasi dan nilai tukar rupiah. Dari beberapa hasil penelitian tersebut diperoleh bahwa penggunaan metode 2SLS ialah metode yang tepat, karena metode estimasi ini dapat digunakan pada kondisi *identified* maupun *overidentified*. Selain itu juga disebabkan karena 2SLS lebih sensitif terhadap kesalahan spesifikasi pada persamaan (Toker, Kacıranlar, & Güler, 2018).

Metode 2SLS akan dapat mengestimasi parameter dengan baik apabila antara variabel bebasnya tidak terjadi masalah multikolinearitas. Multikolinearitas adalah suatu keadaan yang mana variabel-variabel bebas dalam model regresi linier berganda saling bergantung linier. Ketika variabel bebas saling bergantung linier, estimasi parameter dengan menggunakan metode 2SLS menjadi kurang tepat karena akan menghasilkan hasil estimasi dengan variansi yang besar, interval kepercayaan yang tidak tepat sehingga estimasi dengan metode 2SLS menjadi tidak stabil dan menunjukkan kerugian *mean squared errors* (MSE) yang signifikan (Lukman, Ayinde, Siok Kun, & Adewuyi, 2019). Salah satu estimator bias sebagai alternatif OLS untuk menghindari masalah multikolinearitas yaitu regresi ridge (RR) yang diusulkan Hoerlddan Kennard (1970). Estimator RR adalah salah satu yang paling banyak digunakan. Estimator ini membantu mengatasi masalah multikolinearitas dengan menambahkan nilai positif k, ke elemen diagonal matriks $\boldsymbol{Z'Z}$. Konstanta $\boldsymbol{k}$ ini dikenal sebagai parameter bias atau parameter penyusutan (Aslam & Ahmad, 2020).

Beberapa peneliti menggabungkan antara RR dan 2SLS untuk masalah multikolinearitas dan autokorelasi diantaranya adalah Eledum dan Mustafa (2013) dalam jurnalnya menggunakan metode *two stage ridge regression* yang merupakan metode gabungan antara 2SLS dan RR. Metode ini digunakan untuk mengatasi multikolinearitas dan autokorelasi galat pada regresi berganda. Selain itu, Eledum dan Alkhaklifa (2012) menggunakan *generalized two stages ridge regression* yang dapat mengatasi masalah multikolinearitas dan autokorelasi galat. Dari hasil penelitiannya, diperoleh metode ini lebih baik dibanding metode *two stage ridge regression*. Hal ini berdasarkan nilai bias pada masing-masing parameter yang diperoleh. Namun, estimator RR adalah fungsi nonlinear dari parameter bias *k*, sehingga menyebabkan kesulitan dalam pemilihan parameter bias k berdasarkan fungsi parameternya. Untuk mengatasi kelemahan ini, Liu (1993) memperkenalkan estimator lain yang kemudian dikenal estimator Liu (LE) yang menggabungkan estimator ridge dan Stein. LE menambahkan d sebagai parameter bias. Keuntungan utama LE atas RR adalah bahwa LE merupakan fungsi linier dari parameter bias d. Dengan demikian, LE menjadi pilihan yang terbaik untuk menentukan parameter bias d yang sesuai. Sehingga estimator ini lebih mudah digunakan untuk memilih parameter bias d daripada memilih k pada estimator ridge. Selanjutnya (Akdeniz & Kaçıranlar, 1995) menamakan penduga baru ini sebagai penduga Liu. Perbandingan nilai MSE antara OLS dan Liu terdapat didalam penelitian (Sakallıoğlu, Kaçıranlar, & Akdeniz, 1996) yang membandingkan dengan kriteria yang sama. Beberapa peneliti yang menggunakan metode Liu adalah (Li & Yang, 2012) menggunakan metode modified Liu estimator yang merupakan gabungan antra OLS dan RR untuk masalah multikolinearitas pada regresi berganda. Toker dan Guller (2018) kemudian menggunakan estimator *two stage Liu* pada persamaan simultan yang terindikasi mengalami masalah multikolinearitas. Selanjutnya Omara (2019) menggunakan estimator ridge liu berdasarkan estimasi ridge pada regresi berganda. Dari hasil penelitiannya diperoleh bahwa estimator ridge liu lebih baik dibandingkan estimator ridge dan estimator liu berdasarkan nilai MSE.

Dalam menyusun sebuah model ekonomi, diperlukan ketelitian yang kuat dalam memilih variabel-variabel yang saling berhubungan. Sesuai dengan karakteristik dari setiap variabel ekonomi yang bersifat sebagai variabel independen, akan tetapi di satu sisi variabel tersebut dapat menjadi variabel dependen (Basmann, 1957). Gujarati dan Porter (2010) menyebutkan salah satu contoh simultanitas dalam regresi yaitu fungsi upah dan harga. Dalam bidang pertanian, fungsi upah dirumuskan ke dalam indeks harga yang diterima petani, sedangkan fungsi harga dirumuskan ke dalam indeks harga yang dibayar petani. Berdasarkan studi kasus tersebut, maka diduga terdapat simultanitas antara variabel indeks harga yang diterima petani dan indeks harga yang dibayar petani. Indeks harga yang diterima petani merupakan indeks harga yang menunjukkan perkembangan harga produsen atas hasil produksi petani, sedangkan indeks harga yang dibayar petani merupakan indeks harga yang menunjukkan perkembangan harga kebutuhan rumah tangga petani, baik itu kebutuhan untuk konsumsi rumah tangga maupun kebutuhan untuk proses produksi pertanian.

Berdasarkan beberapa penelitian tersebut, penulis mengkaji penggunaan metode *two stage ridge regression* dan *two stage ridge liu* untuk memperoleh estimasi parameter model persamaan simultan. Metode ini diaplikasikan pada data bulanan indeks harga yang diterima dan dibayar petani di Indonesia.

1.2 Rumusan Masalah

Berdasarkan uraian yang telah diberikan sebelumnya, maka diperoleh rumusan masalah sebagai berikut:

1. Bagaimana mengestimasi parameter model persamaan simultan menggunakan metode *two stage ridge regression dan two stage ridge liu*?
2. Bagaimana model persamaan simultan menggunakan metode *two stage ridge liu* data bulanan indeks harga yang diterima dan dibayar petani di Indonesia?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah diuraikan, maka tujuan penelitian ini yaitu:

1. Mendapatkan estimator model persamaan simultan menggunakan metode *two stage ridge regression dan two stage ridge liu*
2. Memperoleh estimasi model persamaan simultan menggunakan metode menggunakan metode *two stage ridge liu* data bulanan indeks harga yang diterima dan dibayar petani di Indonesia

1.4 Batasan Masalah

Batasan masalah pada penelitian ini adalah sebagai berikut:

1. Model persamaan yang digunakan dibatasi pada model persamaan simultan yang teridentifikasi atau *overidentified*, yang mengandung masalah multikolinearitas.
2. Data yang digunakan adalah data nilai tukar petani Indonesia tahun 2016-2021.
3. Estimator bias yang digunakan yaitu dari estimator ridge dan estimator liu

1.5 Manfaat Penelitian

Hasil dari penelitian ini diharapkan dapat menambah wawasan dan pengetahuan tentang penanganan masalah multikolinearitas pada model persamaan simultan menggunakan metode *two stage ridge liu* pada indeks harga yang diterima dan dibayar petani.

BAB II
TINJAUAN PUSTAKA

Di bagian ini, dibahas tentang konsep teoritis yang akan digunakan dalam hasil dan pembahasan

2.1 Model Regresi Linear

Regresi linear adalah metode statistika yang digunakan untuk membentuk model hubungan antara variabel respon (terikat; dependen; Y) dengan satu atau lebih variabel prediktor (bebas; independen; X). Apabila terdapat hubungan linear antara variabel respon (Y) dengan satu variabel prediktor (X) disebut regresi linear sederhana. Sedangkan hubungan linear antara dua atau lebih variabel prediktor ($X_1, X_2, \ldots, X_n$) dengan variabel respon (Y) disebut sebagai regresi linear berganda (Kurniawan, 2008). Analisis regresi dapat digunakan untuk menganalisis data dan mengambil kesimpulan yang bermakna tentang hubungan ketergantungan dari satu variabel respon kepada satu atau lebih variabel prediktor. Hubungan antara dua atau lebih variabel dapat berpola linear ataupun nonlinear (Draper & Smith, 1992). Model regresi linier paling sederhana terdiri atas satu variabel respon dan satu variabel prediktor, seperti pada persamaan (1):

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \tag{1}$$

keterangan:

i : $1, 2, \ldots, n$

n : banyaknya pengamatan

Y : variabel respon

β_0 : rata-rata sebaran nilai Y ketika X bernilai nol

β_1 : parameter regresi

X : variabel prediktor

ε : besaran yang membuat nilai Y menyimpang dari garis regresinya (galat)

Apabila persamaan memiliki lebih dari satu variabel prediktor, maka model disebut sebagai model regresi linear berganda. Bentuk umum persamaan regresi linier berganda dengan p variabel bebas adalah seperti pada persamaan (2) (Kutner, Nachtsheim, & Neter, 2004):

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_p X_{pi} + \varepsilon_i, i = 1,2,3, \dots, n \tag{2}$$

keterangan:

p : banyaknya variabel prediktor

β_j : parameter regresi variabel prediktor ke-$j, j = 1,2, \dots p$

Y_i : variabel respon $(i = 1,2,3, \dots, n)$

X_i : Pengamatan ke$-i$ variabel prediktor X

ε_i :nilai galat regresi yang saling bebas dan berdistribusi normal $\varepsilon_i \sim N(0, \sigma^2)$

Jika nilai-nilai variabel disusun sebagai suatu elemen matriks, maka model regresi persamaan (2) dalam matrik dinyatakan dengan

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \tag{3}$$

dengan: $\boldsymbol{y}' = (\boldsymbol{Y}_1, \boldsymbol{Y}_2, \dots, \boldsymbol{Y}_n)$

$$\boldsymbol{X} = \begin{bmatrix} 1 & X_{11} & X_{21} & \cdots & X_{p1} \\ 1 & X_{12} & X_{22} & \cdots & X_{p2} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & X_{1n} & X_{2n} & \cdots & X_{pn} \end{bmatrix}, \boldsymbol{\beta}' = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}, \boldsymbol{\varepsilon}' = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

dimana diasumsikan bahwa $\boldsymbol{Y}$ adalah vektor $(n \times 1)$ variabel respon, $\mathbf{X}$ adalah matriks variabel prediktor berukuran $n \times (p+1)$, $\boldsymbol{\beta}$ adalah vektor $(p+1) \times 1$ dari parameter regresi yang tidak diketahui dan ε adalah vektor galat dengan elemen diasumsikan independen dan identik secara normal sehingga $E(\varepsilon) = 0$ dan matriks dispersi $E(\varepsilon'\varepsilon) = \sigma^2 I$ (Supranto, 2001).

Dalam regresi linear berganda yang akan diduga adalah parameter regresinya. Persamaan linear untuk pendugaan garis regresi linear ditulis dalam bentuk

$$\widehat{\boldsymbol{y}} = \boldsymbol{X}\widehat{\boldsymbol{\beta}} \tag{4}$$

Nilai sampel yang sesungguhnya dari $\boldsymbol{y}$ tidak akan serupa dengan nilai estimasi dari $\widehat{\boldsymbol{y}}$ dan perbedaan antara nilai $\boldsymbol{y}$ yang sebenarnya dengan nilai estimasi $\widehat{\boldsymbol{y}}$ disebut dengan galat dan dinotasikan dengan $\boldsymbol{\varepsilon}$. Sehingga $\boldsymbol{\varepsilon} = \boldsymbol{y} - \widehat{\boldsymbol{y}}$. Dengan menggunakan metode OLS, $\boldsymbol{\beta}$ dapat diduga dengan meminimumkan jumlah kuadrat dari galat $\boldsymbol{\varepsilon} = \boldsymbol{y} - \hat{\boldsymbol{y}}$. Berdasarkan definisi tersebut maka $\hat{\boldsymbol{\beta}}$ dapat dituliskan sebagai:

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} \tag{5}$$

Hoerl dan Kennard menunjukkan bahwa solusi untuk OLS tidak selalu ada dan tidak ada solusi unik ketika matriks $\boldsymbol{X}'\boldsymbol{X}$ tidak terkondisi (tidak dapat dibalik) karena masalah multikolinearitas. Dalam situasi ini estimator OLS cenderung menjadi sangat besar dan mungkin memiliki varian yang besar. Teorema Gauss-Markov menyatakan bahwa OLS memiliki beberapa kriteria ideal bagi penduganya yang dikenal sebagai BLUE (*best linear and unbiased estimator*). Penduga $\boldsymbol{\beta}$ dapat dinyatakan penduga BLUE jika memenuhi kriteria berikut (Gujarati, 2003):

a. Konsisten

Sifat kebaikan penduga β_j sehubungan dengan ukuran sampel n. Idealnya, semakin besar ukuran sampel maka nilai penduga mendekati nilai parameternya dan keragaman dari penduga mendekati nol.

$$\lim_{n\to\infty} var(\hat{\boldsymbol{\beta}}_j) = 0$$

b. Tak Bias

Suatu penduga β_j dikatakan penduga yang tidak bias bagi parameternya apabila nilai harapan penduga memiliki selisih nol dengan nilai parameter yang sebenarnya.

$$E(\hat{\boldsymbol{\beta}}_j) = \boldsymbol{\beta}_j$$

c. Ragam Minimum

Ketersebaran nilai penduga β_j dari nilai parameter yang sebenarnya perlu diperhatikan. Nilai duga yang sangat menyebar dari nilai parameter menghasilkan penduga dengan ragam besar dan tidak akurat. Oleh karena itu, diharapkan suatu penduga memiliki $var(\hat{\boldsymbol{\beta}}_j)$minimum. Penduga tak bias yang memiliki ragam terkecil disebut sebagai penduga yang efisien. Dalam upaya mencapai hasil pendugaan parameter yang baik tersebut, OLS memiliki empat asumsi klasik yang harus dipenuhi. Asumsi-asumsi tersebut adalah normalitas sisaan, homoskedastisitas, non autokorelasi, dan non multikolinieritas.

2.2 Model Persamaan Simultan

Dalam bidang bisnis dan ekonomi banyak hubungan antara variabel dinyatakan dalam persamaan tunggal. Dalam model persamaan tunggal variabel tak bebas (*dependent variable*) dinyatakan sebagai fungsi linear dari satu atau lebih variabel bebas (*independent variable*). Model persamaan tuggal dapat dituliskan:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \qquad (i = 1,2,\dots,n) \tag{6}$$

dimana:

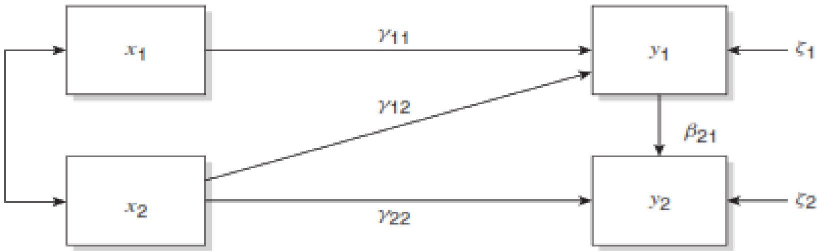
X_i : pengamatan ke$-i$ variabel bebas X

Y : variabel tak bebas

β : parameter

ε : galat

Dalam model persamaan tunggal hubungan sebab akibat antara variabel tak bebas dengan variabel bebas merupakan hubungan satu arah yaitu variabel bebas merupakan penyebab dan variabel tak bebas merupakan akibat. Model persamaan tunggal dapat diselesaikan dengan menggunakan model regresi persamaan tunggal. Akan tetapi model regresi persamaan tunggal tidak dapat digunakan pada model persamaan simultan. Hal ini disebabkan karena model persamaan simultan tidak mungkin diestimasi dengan hanya menggunakan satu persamaan dengan mengabaikan informasi yang ada pada persamaan lainnya. Model persamaan simultan ini kemudian dapat digambarkan sebagai berikut:



Gambar 1. Diagram Hubungan Model Pesamaan Simultan

Berbeda dengan model-model ekonometrika, model persamaan simultan memiliki lebih dari satu persamaan dengan masing-masing variabel independent dan variabel dependent. Karena terdapat banyak persamaan maka terdapat lebih dari satu variabel dependent. Persamaan-persamaan tersebut menggambarkan hubungan antar fenomena ekonomi. Selain itu, variabel dependent pada satu persamaan dapat menjadi variabel independent pada persamaan lain. Sehingga antar persamaan saling berhubungan. Pada model ini terdapat dua jenis variabel yaitu variabel endogen dan variabel eksogen.

2.2.1 Variabel Pada Model Persamaan Simultan

Penyebutan variabel independen (penjelas) dan variabel dependen tidak tepat lagi jika digunakan pada model persamaan simultan, karena variabel dependen bisa juga menjadi variabel independen. Menurut Gujarati (2012: 360), dalam konteks model persamaan simultan, terdapat 2 jenis variabel yaitu:

1. Variabel endogen

Variabel endogen ialah variabel respon di dalam sistem persamaan simultan yang nilainya ditentukan di dalam sistem persamaan walaupun variabel-variabel tersebut juga muncul sebagai variabel bebas di dalam sistem persamaan.

2. Variabel eksogen

Variabel eksogen ialah variabel yang nilainya ditentukan di luar model sebab nilainya sudah diketahui sebelumnya. Jumlah persamaan dalam model persamaan simultan sama dengan jumlah seluruh variabel endogennya.

2.2.2 Notasi Persamaan Simultan

Ada dua persamaan dalam model simultan, yaitu persamaan struktural dan *reduced form*. Persamaan struktural yaitu persamaan asli yang menggambarkan perilaku hubungan antar variabel dalam persamaan. Sedangkan persamaan *reduced form* merupakan suatu persamaan yang diperoleh dari persamaan-persamaan struktural yang telah dikaitkan. Dengan menyelesaikan persamaan *reduced form*, selanjutnya menghitung koefisien-koefisien dalam persamaan struktural. Oleh karena itu, penaksiran terhadap persamaan structural akan tergantung dari hasil penaksiran pada persamaan *reduced form*. Sebelum menerapkan suatu metode untuk menduga parameter pada setiap persamaan dalam model tersebut, perlu dilakukan identifikasi model terlebih dahulu. Gujarati (2003) mengemukakan bahwa untuk dapat diduga parameternya, suatu model persamaan simultan harus teridentifikasi.

Model persamaan simultan dapat direpresentasikan dengan T observasi dalam m variabel endogen dinotasikan dengan $\boldsymbol{y}_{t1},\ \boldsymbol{y}_{t2,\dots,}\ \boldsymbol{y}_{tm}$ dan k variabel eksogen dinotasikan dengan $\boldsymbol{x}_{t1,}\ \boldsymbol{x}_{t2}, \dots\ \boldsymbol{x}_{tk}$ serta m variabel galat acak dinotasikan dengan $\boldsymbol{\varepsilon}_{t1}, \boldsymbol{\varepsilon}_{t2,\dots,}\boldsymbol{\varepsilon}_{tm}$.

Notasi umum dalam m persamaan untuk merepresentasikan persamaan simultan dapat dituliskan sebagai berikut:

$$\begin{aligned} y_{t1}\gamma_{11} + \cdots + y_{tm}\gamma_{1m} + x_{t1}\beta_{11} + \cdots + x_{tk}\beta_{1k} &= \varepsilon_{t1} \\ y_{t1}\gamma_{21} + \cdots + y_{tm}\gamma_{2m} + x_{t1}\beta_{21} + \cdots + x_{tk}\beta_{2k} &= \varepsilon_{t2} \\ &\vdots \\ y_{t1}\gamma_{m1} + \cdots + y_{tm}\gamma_{mm} + x_{t1}\beta_{k1} + \cdots + x_{tk}\beta_{km} &= \varepsilon_{tm} \end{aligned} \tag{7}$$

dimana γ dan β adalah parameter struktural dari sistem persamaan yang tidak diketahui dan akan ditaksir dari data. Dalam notasi matriks dapat dinyatakan sebagai berikut:

$$\begin{bmatrix} \boldsymbol{y}_{t1} \\ \boldsymbol{y}_{t2} \\ \vdots \\ \boldsymbol{y}_{tm} \end{bmatrix} \begin{bmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1m} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{m1} & \gamma_{m2} & \cdots & \gamma_{mm} \end{bmatrix} + \begin{bmatrix} \boldsymbol{X}_{t1} \\ \boldsymbol{X}_{t2} \\ \vdots \\ \boldsymbol{X}_{tk} \end{bmatrix} \begin{bmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1m} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{p1} & \beta_{p2} & \cdots & \beta_{pm} \end{bmatrix} = \begin{bmatrix} \varepsilon_{t1} \\ \varepsilon_{t2} \\ \vdots \\ \varepsilon_{tm} \end{bmatrix}$$

Dalam bentuk matriks, persamaan (7) dapat dibentuk menjadi:

$$\mathbf{Y\Gamma} + \mathbf{XB} = \mathbf{U} \tag{8}$$

dimana:

$\boldsymbol{Y}$: Vektor variabel endogen berukuran $m \times 1$

$\boldsymbol{X}$: Vektor variabel eksogen yang berukuran $p \times 1$

$\boldsymbol{\Gamma}$: Matriks koefisien variabel endogen yang tidak diketahui berukuran $m \times m$

$\boldsymbol{B}$: Matriks koefisien variabel eksogen yang tidak diketahui berukuran $m \times p$

$\mathbf{U}$ adalah matriks struktural galat. Elemen untuk setiap baris $\boldsymbol{X}$ berdistribusi bebas dari galat struktural untuk setiap elemen $\mathbf{U}$. Galat struktural $E(\varepsilon) = 0$ dan bersifat homoskedastik (Toker, Kacıranlar, & Güler, 2018). Tabel 1 menyajikan rangkuman notasi pesamaan simultan yang digunakan.

Tabel 1. Notasi untuk Model Persamaan Simultan

Vektor/Matriks	**Definisi**	**Dimensi**
Variabel:		
$\boldsymbol{y}$	Variabel Endogen	$m \times 1$
$\boldsymbol{X}$	Variabel Eksogen	$p \times 1$

ε	Galat atau eror dalam persamaan	$m \times 1$
Koefisien:		
$\boldsymbol{\Gamma}$	Matriks koefisien untuk variabel eksogen; pengaruh variabel eksogen terhadap endogen; pengaruh langsung $\boldsymbol{y}$ terhadap $\boldsymbol{y}$	$m \times m$
$\boldsymbol{B}$	Matriks koefisien untuk variabel endogen; pengaruh variabel endogen teerhadap endogen; pengaruh langsung $\boldsymbol{x}$ terhadap $\boldsymbol{y}$	$m \times p$

Persamaan (8) merupakan model lengkap yang secara umum dapat diselesaikan dengan menggunakan model *reduced form* atau model bentuk reduksi.

2.2.3 Bentuk Persamaan Yang Direduksi (*Reduced Form*)

Menurut Gujarati (2012), dari persamaan struktural dapat diperoleh bentuk persamaan reduksi (*reduced form equation*) dan koefisien bentuk reduksis yang berhubungan. Bentuk reduksi adalah suatu bentuk persamaan dimana semua variabel endogen merupakan fungsi dari semua variabel eksogen dan galat. Bentuk persamaan reduksi adalah jika model dari persamaan struktural jika $\boldsymbol{\Gamma}$ adalah nonsingular dapat mengekspresikan variabel endogen sebagai fungsi dari variabel eksogen dan galat stokastik. Persamaan ini didapat dengan memecahkan bentuk persamaan struktural sehingga variabel endogen pada setiap persamaan sebagai fungsi dari variabel predetermine dan galat stokastik (Jugde, 1980).

Model *reduced form* dapat dituliskan sebagai berikut:

$$\begin{aligned} Y_{1i} &= \pi_{11}X_{1i} + \pi_{12}X_{2i} + \cdots + \pi_{1k}X_{pi} + v_{1i} \\ Y_{2i} &= \pi_{21}X_{1i} + \pi_{22}X_{2i} + \cdots + \pi_{2k}X_{pi} + v_{2i} \\ \vdots & \qquad \vdots \qquad\quad \cdots \quad \vdots \qquad \vdots \\ Y_{mi} &= \pi_{m1}X_{1i} + \pi_{m2}X_{mi} + \cdots + \pi_{mp}X_{pi} + v_{mi} \end{aligned} \tag{9}$$

dimana π adalah koefisien *reduced form* dan v adalah galat *reduced form*. Persamaan (9) merupakan reduksi dari persamaan (7). Persamaan (9) dalam bentuk matriks dituliskan sebagai berikut:

$$\begin{bmatrix} Y_{1i} \\ Y_{2i} \\ \vdots \\ Y_{mi} \end{bmatrix} = \begin{bmatrix} \pi_{11} & \pi_{12} & \cdots & \pi_{1m} \\ \pi_{21} & \pi_{22} & \cdots & \pi_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \pi_{m1} & \pi_{m2} & \cdots & \pi_{mp} \end{bmatrix} \begin{bmatrix} X_{1i} \\ X_{2i} \\ \vdots \\ X_{pi} \end{bmatrix} + \begin{bmatrix} v_{1i} \\ v_{2i} \\ \vdots \\ v_{mi} \end{bmatrix} \tag{10}$$

Matriks diatas dapat ditulis menjadi:

$$\boldsymbol{Y}_i = \boldsymbol{\Pi}\boldsymbol{X}_i + \boldsymbol{v}_i \tag{11}$$

dimana $\boldsymbol{\Pi}$ adalah matriks koefisien dengan ukuran $m \times p$.

2.2.4 Identifikasi Model Persamaan Simultan

Dalam model persamaan simultan, identifikasi dilakukan pada awal sebelum melakukan penaksiran untuk menentukan apakah suatu model persamaan simultan dapat dilakukan penaksiran atau tidak, dan mengetahui metode penaksiran apa yang sebaiknya digunakan pada persamaan simultan. Menurut Koutsoyiannis (1977), tujuan dari identifikasi adalah mendeteksi apakah persamaan teridentifikasi (*exactly identified)* atau tidak teridentifikasi (*not identified*). Identifikasi persamaan simultan ada tiga bentuk yaitu:

1. Persamaan tidak teridentifikasi (*not identified)*

Keadaan ini terjadi jika banyaknya variabel eksogen yang tidak tercakup dalam suatu persamaan lebih kecil dari pada banyaknya variabel endogen dalam persamaan tersebut dikurangi satu. Dalam model seperti ini, parameter struktural tidak dapat diestimasi. Notasinya yaitu: $(p - p^*) < (m^* - 1)$. Jika persamaan dalam kondisi tidak teridentifikasi, maka tidak ada teknik yang dapat digunakan untuk menaksir parameternya (Koutsoyiannis, A., 1973).

2. Persamaan tidak teridentifikasi (*exactly identified)*

Keadaan ini terjadi jika banyaknya variabel eksogen yang tidak tercakup dalam suatu persamaan sama dengan banyaknya variabel endogen dalam persamaan tersebut dikurangi satu. Dalam model seperti ini, parameter struktural dapat diestimasi. Notasinya yaitu: $(p - p^{*)} = (m^* - 1)$ dengan rank $A = m - 1$. Dimana A adalah matriks koefisien dari variabel yang tidak dimasukkan dalam persamaan tertentu, tetapi terkandung dalam persamaan lain dalam model. Persamaan dalam kondisi teridentifikasi tepat, dapat ditaksir parameternya menggunakan *indirect least square* (ILS) dan *two stage least square* (2SLS) (Ananta, 1987; Gurajati, ekonometrika dasar, 1978).

3. Persamaan Identifikasi Berlebih (*overidentified*)

Persamaan simultan dikatakan teridentifikasi berlebih apabila banyaknya variabel eksogen yang tercakup dalam persmaan lebih banyak dari banyaknya variabel endogen dalam persamaan dikurangi satu. Dalam model seperti ini, parameter struktural dapat diestimasi. Notasinya yaitu: $(p - p^*) > (m^* - 1)$

dengan rank $A = m - 1$. Pada persamaan dengan kondisi teridentifikasi berlebihan, metode ILS tidak dapat digunakan untuk mengestimasi persamaan tersebut, biasanya ditaksir menggunakan 2SLS atau 3SLS (Koutsoyiannis, A., 1973).

2.2.5 Aturan Identifikasi

Dalam melakukan identifikasi persamaan simultan ada dua cara yaitu kondisi order (*order condition*) dan kondisi peringkat (*rank condition*). Untuk keperluan pengujian identifikasi persamaan perlu dibuat notasi berikut:

m : jumlah variabel endogen dalam model/sistem

m^* : jumlah variabel endogen dalam suatu persamaan

p : jumlah variabel eksogen dalam model/sistem

p^* : jumlah variabel eksogen dalam suatu persamaan

1. Kondisi Order (*Order Condition*)

kondisi orde merupakan suatu kondisi yang diperlukan (*necessary*) namun belum menjadi kondisi cukup (*sufficient*) untuk identifikasi (Koutsoyiannis, 1977). Ada dua cara untuk mengidentifikasi kondisi orde, yang masing-masing sebenarnya menghasilkan hasil yang setara. Gujarati (2012), menyatakan:

a. Pada model persamaan simultan agar dapat diidentifikasi, setidaknya harus mengeluarkan $m - 1$ variabel (endogen dan predetermine) yang terdapat dalam model. Koutsoyiannis (1977) juga menyatakan untuk persamaan yang teridentifikasi, jumlah variabel yang dikeluarkan (endogen dan predetermine) dari model harus sama dengan atau lebih besar dari jumlah variabel endogen dikurangi satu. Dinotasikan dengan,

$$(m - m^*) + (p - p^* \geq m - 1 \quad (12)$$

Jika variabel yang dikeluarkan tepat sejumlah variabel, maka persamaan tersebut tepat teridentifikasi (*exactly identified*). Jika variabel yang dikeluarkan lebih dari $m - 1$ variabel, maka persamaan tersebut terlalu teridentifikasi (*overidentified*).

b. Dalam model m persamaan simultan agar dapat diidentifikasi, jumlah dari variabel predetermine yang dikeluarkan dari persamaan tidak boleh kurang dari jumlah variabel endogen yang dimasukan dalam persamaan dikurangi dengan satu. Dinotasikan dengan,

$$(p - p^*) \geq m^* - 1 \quad (13)$$

Jika $(p - p^*) = m - 1$, maka persamaan tersebut tepat teridentifikasi (exactly identified). Jika $(p - p^*) > m - 1$, maka persamaan tersebut terlalu teridentifikasi (*overidentified*).

Apabila suatu persamaan dalam model tidak memenuhi syarat order, maka tidak diperlukanmenyelidiki syarat rank, sebaliknya bila syarat order dipenuhi, maka dilanjutkan menyelidiki syarat rank.

2. Kondisi Rank (*Rank Condition*)

Dengan kondisi rank ini, dapat memenuhi dua aturan formal dalam pengidentifikasian. Rank berkenaan dengan konsep matriks dengan orde yang mempuyai determinan sama dengan nol atau jumlah maksimum baris-baris atau kolom-kolom yang bebas linear (independen). Kondisi rank diperlukan karena walaupun melalui pengujian kondisi orde suatu persamaan teridentifikasi namun bisa saja melalui pengujian kondisi rank tidak terpenuhi sehingga penaksiran parameter untuk persamaan simultan tidak dapat dilakukan. Hal ini mungkin terjadi jika kolom-kolom atau baris-baris matriks dari suatu persamaan tidak bebas linear atau terdapat hubungan antar variabelnya.

Gujarati (2012) menyatakan bahwa dalam model persamaan simultan dapat diidentifikasi, jika dan hanya jika setidaknya terdapat satu determinan yang tidak nol dari matrik yang orde $(m - 1) \times (m - 1)$. Matriks itu dapat dibentuk dari koefisien variabel (endogen dan predetermine) yang dikeluarkan dari persamaan tetapi ada dalam persamaan lainnya dari model.

Untuk melakukan pengujian dalam kondisi rank, langkah-langkah yang dapat dilakukan adalah:

1. Manipulasi persamaan dengan memindahkan semua variabel sisi kanan ke sebelah kiri kecuali variabel galat stokastik.
2. Tuliskan hasil manipulasi tersebut dalam bentuk tabel.
3. Mencoret koefisien-koefisien dari baris yang didalamnya ada persamaan yang sedang diperhatikan untuk pengidentifikasian.
4. Mencoret koefisien-koefisien dari kolom yang berhubungan dengan langkah 2 yaitu yang tidak sama dengan nol.
5. Data yang tersisa dalam tabel yaitu data dari koefisien-koefisien yang tidak berhubungan dengan langkah 2 dan 3 (data dari seluruh variabel dalam model tapi tidak termasuk persamaan yang sedang diperhatikan untuk pengidentifikasian). Hitung determinan berorde $(m - 1)$ dari data yang tersisa.

Jika seluruh kemungkinan dari determinan yang berorde $(m-1)$adalah nol, maka persamaan yang sedang diperhatikan tidak dapat diidentifikasi.

Kondisi rank menyatakan apakah persamaan dapat diidentifikasi atau tidak, sedangkan kondisi orde menyatakan jika hal itu dapat secara tepat teridentifikasi atau terlalu terindentifikasi (Gujarati, 2012). Prinsip umum dari identifikasi persamaan struktural pada model persamaan simultan, yaitu:

1. Jika $(p-p^*) > m-1$ dan rank dari matriks $\boldsymbol{R}$ adalah $(m-1)$, maka persamaan tersebut terlalu teridentifikasi (*overidentified*).
2. Jika $(p-p^*) = m-1$ dan rank dari matriks $\boldsymbol{R}$ adalah $(m-1)$, maka persamaan tersebut tepat teridentifikasi (*exactly identified*).
3. Jika$(p-p^*) \geq m-1$ dan rank dari matriks $\boldsymbol{R}$ adalah kurang dari $(m-1)$, maka persamaan tersebut kurang teridentifikasi (*underidentified*).
4. Jika $(p-p^*) < m-1$ dan rank dari matriks $\boldsymbol{R}$ adalah kurang dari $(m-1)$, maka persamaan tersebut tidak teridentifikasi.

2.2.6 Uji Simultan

Pengujian simultanitas yang dikemukakan oleh Hausman (1978) bertujuan membuktikan secara empiris bahwa suatu sistem model persamaan benar-benar memiliki hubungan simultan antar persamaan strukturalnya. Misalkan kita memiliki model sebagai berikut

$$Y_1 = \gamma_{10} + \gamma_{11}Y_2 + \beta_{11}X_1 + \beta_{12}X_2 + \varepsilon_1 \qquad (14)$$

dimana variabel Y_2 diduga endogen dan membutuhkan instrumen variabel. Misalkan telah diasumsikan variabel Y_2 dipengaruhi oleh $X_3\,, X_4$ dan X_5 merupakan instrumen variabel bagi Y_2

$$Y_2 = \gamma_{20} + \gamma_{21}Y_1 + \beta_{21}X_3 + \beta_{22}X_4 + \beta_{23}X_5 + \varepsilon_2 \qquad (15)$$

Akan digunakan ε_2 yang diperoleh dari Persamaan (13) sebagai variabel baru pada Persamaan **(**14**)** untuk mengetahui apakah Y_2 merupakan variabel endogen yang membutuhkan instrumen variabel. Selanjutnya dilakukan uji signifikansi, jika ε_2 signifikan maka dapat disimpulkan bahwa ε_2 adalah variabel endogen dan membutuhkan instrumen variabel.

a. Hipotesis

$\mathrm{H}_0 : \mathrm{Y}_1 = 0$ (variabel Y_1 tidak signifikan berpengaruh ε_2)

$\mathrm{H}_1 : \mathrm{Y}_1 \neq 0$ (variabel Y_1 signifikan berpengaruh ε_2)

b. Taraf nyata $\alpha = 0{,}05$

c. Statistik Uji

$$P_{value} = 0{,}000$$

d. Kriteria Keputusan

$$\mathrm{H_0} \text{ ditolak jika } \mathrm{P_{value}} < \alpha$$

dengan memasukkan unsur galat pada tiap persamaan maka dapat diketahui pengaruh galat tersebut. Pengaruh galat pada persamaan ini yang menjadi objek pengamatan keberadaan simultanitas pada persamaan.

2.3 Metode Two Stage Least Square

Setelah model persamaan simultan dapat diidentifikasi, maka langkah selanjutnya dilakukan penaksiran terhadap model. Untuk menguji hipotesis yang telah dikembangkan maka digunakan kuadrat terkecil dua tahap atau *two stage least square* (2SLS). Metode 2SLS digunakan ketika terjadi korelasi antara variabel bebas dengan galat. Ide dasar metode 2SLS adalah menghilangkan korelasi yang terjadi antara variabel bebas dengan galat. Bentuk umum persamaan struktural ke-i sebagai berikut:

$$Y_i = \gamma_{i1}Y_1 + \gamma_{i2}Y_2 + \cdots + \gamma_{im}Y_m + \beta_{i1}X_1 + \cdots + \beta_{ip}X_p + \varepsilon_i$$

Langkah-langkah penyelesaiaan metode 2SLS dibagi dalam dua langkah sebagai berikut:

i. Menggunakan metode OLS pada persamaan reduksi. Persamaan reduksi adalah persamaan setiap variabel endogen dengan semua variabel eksogen

$$Y_1 = \pi_{11}X_1 + \pi_{12}X_2 + \cdots + \pi_{1p}X_p + \varepsilon_1$$

$$Y_2 = \pi_{21}X_1 + \pi_{22}X_2 + \cdots + \pi_{2p}X_p + \varepsilon_2$$

$$\vdots \quad \vdots \quad \vdots \quad \vdots$$

$$Y_i = \pi_{i1}X_1 + \pi_{i2}X_2 + \cdots + \pi_{ip}X_p + \varepsilon_i$$

diperoleh

$$\hat{Y}_1 = \pi_{11}X_1 + \pi_{12}X_2 + \cdots + \pi_{1p}X_p$$

$$\hat{Y}_2 = \pi_{21}X_1 + \pi_{22}X_2 + \cdots + \pi_{2p}X_p$$

$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots$$

$$\hat{Y}_i = \pi_{i1}X_1 + \pi_{i2}X_2 + \cdots + \pi_{ip}X_p$$

Sehingga persamaan simultan dapat dinyatakan sebagai

$$Y_1 = \hat{Y}_1 + \varepsilon_1$$

$$Y_2 = \hat{Y}_2 + \varepsilon_3$$

$$\vdots$$

$$Y_i = \hat{Y}_i + \varepsilon_i$$

ii. Variabel endogen yang muncul disisi kiri pada persamaan struktural diganti dengan persamaan $Y_i = \hat{Y}_i + \varepsilon_i$ dan memperoleh persamaan sebagai berikut:

$$Y_i = \gamma_{i1}(\hat{Y}_1 + \varepsilon_1) + \gamma_{i2}(\hat{Y}_2 + \varepsilon_3) + \cdots + \gamma_{im}(\hat{Y}_i + \varepsilon_i) + \beta_{i1}X_1 + \cdots + \beta_{ip}X_p + \varepsilon_i$$

$$Y_i = \gamma_{i1}\hat{Y}_1 + \gamma_{i1}\varepsilon_1 + \gamma_{i2}\hat{Y}_2 + \gamma_{i2}\varepsilon_2 \ldots + \gamma_{im}\hat{Y}_i + \gamma_{im}\varepsilon_i + \beta_{i1}X_1 + \cdots + \beta_{ip}X_p + \varepsilon_i$$

$$Y_i = \gamma_{i1}\hat{Y}_1 + \gamma_{i2}\hat{Y}_2 + \cdots + \gamma_{im}\hat{Y}_i + \gamma_{im}\varepsilon_i + \beta_{i1}X_1 + \cdots + \beta_{ip}X_p + \varepsilon_i^*$$

dimana $\varepsilon_i^* = \varepsilon_i + \gamma_{i1}\varepsilon_1 + \gamma_{i2}\varepsilon_2 + \gamma_{im}\varepsilon_i$

Metode 2SLS ini dapat digunakan pada persamaan simultan yang kondisi identifikasinya *overidentified*. Selain itu, dapat digunakan untuk menaksir persamaan yang *exactly identified.* Meskipun 2SLS metode estimasi persamaan yang disukai, karena memiliki beberapa sifat yang diinginkan, namun sangat sensitif terhadap multikolinearitas.

2.4 Regresi Ridge

Regresi ridge (*Ridge* Regression) ditujukan untuk mengatasi kondis buruk (*ill* conditioned) yang diakibatkan oleh korelasi yang tinggi antara beberapa variable bebas dalam model. Hal ini menyebabkan matriks $\boldsymbol{X'X}$ hamper singular, sehingga menghasilkan nilai estimasi parameter model yang tidak stabil (Draper & Smith, 1992). Regresi ini merupakan modifikasi dari metode OLS dengan cara menambah tetapan bias (k) yang kecil pada diagonal matriks $\boldsymbol{X'X}$ (Hoerl & Kennard, 1970).

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^{*\prime}\boldsymbol{X}^* + k\boldsymbol{I})^{-1}\boldsymbol{X}^{*\prime}\boldsymbol{Y}^* \qquad (16)$$

Dimana k adalah sebuah bilangan yang positif atau $k \geq 0$, umumnya k terletak antara interval $0 < k < 1$.

dengan:

$\boldsymbol{X}^*$: Matriks $\mathbf{X}$ yang telah ditransformasi dengan pemusatan dan penskalaan.

$\mathbf{Y}^*$: Matriks Y yang telah ditransformasi dengan pemusatan dan penskalaan.

I : Matriks identitas

k : Nilai tetapan bias dimana nilai adalah tetap $0 < k \leq 1$.

2.5 Metode Liu

Regresi ridge menggunakan konstanta k yang sulit ditentukan dan tidak mengatasi pengaruh multikolinearitas secara signifikan. Pada tahun 1993, Liu Kejian mengusulkan metode penaksiran baru yang dapat memperbaiki kelemahan pada regresi ridge yang dikenal dengan metode Liu. Metode Liu menghasilkan estimasi parameter yang disebut estimator Liu. Estimator Liu ini bersifat bias dan bergantung pada suatu konstanta yang disebut konstanta estimator Liu. Metode Liu pada regresi linear diperoleh dengan menambahkan baris pada persamaan (3) dengan fungsi kendala $d\widehat{\boldsymbol{\beta}} = \boldsymbol{\beta} + \tilde{\boldsymbol{\varepsilon}}$. Secara umum, estimator Liu dituliskan sebagai berikut:

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X} + \boldsymbol{I})^{-1}(\boldsymbol{X}'\boldsymbol{Y} + d\widehat{\boldsymbol{\beta}}) \tag{17}$$

dimana d: nilai tetapan bias dengan $0 < d < 1$, $\widehat{\boldsymbol{\beta}}$ adalah estimator OLS.

Perlu dicatat bahwa penaksir regresi ridge dan Liu dipengaruhi oleh multikolinearitas karena bergantung pada penaksir OLS. Oleh karena itu, penaksir ridge dan Liu dapat memberikan informasi yang bias dengan adanya multikolinearitas. Metode Ridge Liu pada regresi linear diperoleh dengan menambahkan baris pada persamaan (3) dengan fungsi kendala $dk(\widehat{\boldsymbol{\beta}}) = k\boldsymbol{\beta} + \tilde{\boldsymbol{\varepsilon}}$. Sehingga metode ridge liu dirumuskan sebagai berikut:

$$\begin{aligned}\widehat{\boldsymbol{\beta}} &= (\boldsymbol{X}'\boldsymbol{X} + k\boldsymbol{I})^{-1}\left(\boldsymbol{X}'\boldsymbol{Y} + kd\widehat{\boldsymbol{\beta}}_{ridge}\right), k > 0, 0 < d < 1 \\ &= (\boldsymbol{X}'\boldsymbol{X} + k\boldsymbol{I})^{-1}(\boldsymbol{X}'\boldsymbol{Y} + kd)\widehat{\boldsymbol{\beta}}_{ridge} \qquad (18)\end{aligned}$$

dimana $\widehat{\boldsymbol{\beta}}_{ridge}$ estimator estimator ridge.

2.6 Pemilihan Nilai Bias

Menentukan nilai k merupakan hal yang sulit untuk mencapai koefisien regresi yang stabil. Terdapat beberapa metode untuk menentukan kontanta bias k, yaitu :Pemilihan parameter yang sesuai sangat penting dalam penelitian ini. Penggunaan estimator ridge sangat bergantung pada parameter ridge k. Beberapa metode untuk memperkirakan parameter ridge telah diusulkan. Tujuan dari aplikasi penduga baru ini, secara optimal nilai k dan d diperoleh. Untuk mendapatkan hasil yang optimal nilai k, kita asumsikan nilai d tetap. Menurut Montgomery dan Peck (1992), cara untuk memilih parameter bias (k) yaitu dengan menggunakan rumus HKB yang diperkenalkan oleh Hoerdl, Kennard dan Balwin:

$$\hat{k}_{HKB} = \frac{p\hat{\sigma}^2}{\sum_i^p \hat{\alpha}_i^2} \tag{19}$$

dimana $\hat{\alpha}\,\mathrm{dan}\,\hat{\sigma}$ diperoleh dari OLS, dengan p adalah banyaknya parameter regresi yang diestimasi.

Berikut ini akan dijelaskan pemilihan konstanta d yang optimal pada skala estimator Liu. Skalar mean square galat estimator Liu dapar dipandang sebagai fungsi dari konstanta estimator Liu yaitu:

$$f(d) = \sigma^2 \sum_{j=1}^{k} \frac{(\lambda_j + d)^2}{\lambda_j(\lambda_j + 1)^2} + (d-1)^2 \sum_{j=1}^{k} \frac{\beta_j^2}{(\lambda_j + 1)^2}$$

Untuk mendapatkan d optimal yang meminimumkan fungsi $f(d)$, maka fungsi $f(d)$ diturunkan satu kali terhadap d dan hasil turunannya disamakan dengan 0. Turunan fungsi $f(d)$ terhadap d adalah:

$$f'(d) = 2\sigma^2 \sum_{j=1}^{k} \frac{(\lambda_j + d)}{\lambda_j(\lambda_j + 1)^2} + 2(d-1) \sum_{i=1}^{k} \frac{\beta_j^2}{(\lambda_j + 1)^2} = 0$$

$$\sigma^2 \sum_{i=1}^{k} \frac{(\lambda_j + d)}{\lambda_j(\lambda_j + 1)^2} + 2(d-1) \sum_{i=1}^{k} \frac{\delta_i^2}{(\lambda_j + 1)^2} = 0$$

Nilai tetapan d berada diantara 0 dan 1. Berdasarkan Lukman et.al (2019), parameter bias d dapat diperoleh dengan:

$$d_{opt} = \sum_{j=1}^{k} \frac{(\beta_j^2 - \sigma^2)}{\lambda_j(\lambda_j + 1)^2} \Big/ \sum_{j=1}^{k} \frac{\sigma^2 + \beta_j^2}{\lambda_i(\lambda_i + 1)^2}$$

Dengan λ_i adalah nilai eigen dari matriks $\boldsymbol{X'X}$, $\beta_j, j = 1,2, \dots, p$ adalah parameter model regresi dan σ^2 adalah variansi galat pada model regresi. Karena β_j dan σ^2 adalah parameter yang nilainya tidak diketahui maka β_j dan σ^2 akan diestimasi menggunakan estimator OLS. Sehingga nilai d yang optimal dengan menggunakan $\widehat{\beta_j}$ dan $\hat{\sigma}^2$ adalah:

$$d_{opt} = \sum_{j=1}^{k} \frac{\left(\widehat{\beta_j}^2 - \hat{\sigma}^2\right)}{\lambda_j\left(\lambda_j+1\right)^2} \Big/ \sum_{j=1}^{k} \frac{\hat{\sigma}^2 + \lambda_i \widehat{\beta_j}^2}{\lambda_i(\lambda_i+1)^2} \tag{20}$$

2.7 Multikolinieritas

Multikolinieritas pertama kali diperkenalkan oleh Ragnar Frisch yang berarti adanya hubungan linier diantara beberapa atau semua variabel bebas pada model regresi linier berganda. Dalam regresi linier berganda seharusnya tidak terjadi multikolinieritas. Jika multikolinieritas terjadi, maka dapat memberikan dampak variansi yang besar sehingga sulit mendapatkan estimasi yang tepat (Widarjono, 2005). Dalam bentuk matriks, multikolinearitas adalah kondisi buruk (*ill condition*) dari matriks $\boldsymbol{X'X}$ yaitu kondisi yang tidak memenuhi asumsi klasik. Matriks $\boldsymbol{X'X}$ akan singular, sehingga nilai perkiraan koefisien regresi menjadi tidak berhingga dan tak mungkin menduganya (Gujarati, 2003).

Terdapat beberapa metode yang dapat digunakan untuk mendeteksi multikolinieritas, salah satunya yaitu dengan menggunakan nilai VIF (*Variance Inflation Factor*) dari masing-masing variabel bebas terhadap variabel respon. Nilai VIF dapat diperoleh dengan:

$$VIF = \frac{1}{1-R_j^2} \tag{21}$$

dimana R_j^2 adalah koefisien determinasi dari X_j dan X_j^* variabel bebas lainnya. Nilai R_j^2 memiliki pengaruh terhadap nilai *VIF* yang dihasilkan, dimana semakin besar nilai R_j^2 maka semakin besar pula nilai *VIF* yang dihasilkan. Jika nilai $VIF > 10$ maka dapat disimpulkan bahwa terdapat masalah multikolinieritas (Gujarati, 2003). Salah satu prosedur paling populer untuk menangani masalah multikolinearitas adalah menggunakan salah estimator yang bias, dengan varians minimum. Salah satunya dengan Regresi Ridge.

2.8 Metode Pemusatan dan Penskalaan

Pemusatan dan penskalaan data merupakan bagian dari membakukan variabel. Modifikasi sederhana dari pembakuan variabel ini adalah transformasi korelasi (*Correlation Transformation*). Salah satu metode pemusatan dan penskalaan data adalah *unit length* scalling. Berikut ini merupakan pembakuan variabel dependen Y dan variabel independen X:

$$Y_i^* = \frac{Y_i - \bar{Y}}{S_y}\ , S_y = \sqrt{\frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{n-1}}$$

$$X_{ij}^* = \frac{X_{ij} - \bar{X}_j}{S_{xj}}, S_{xj} = \sqrt{\frac{\sum_{i=1}^{n}\left(X_{ij} - \bar{X}_j\right)^2}{n-1}}\ , j = 1,2, \dots, p$$

dimana:

$\bar{Y}$: rata-rata dari Y

$\bar{X}_j$: rata-rata dari X_j

S_{x_j} : standar deviasi dari X_j

S_y : standar deviasi dari Y

Model regresi dengan transformasi variabel Y_i^* dan X_{ij}^* yang didefiniskan dengan transformasi korelasi diatas disebut model regresi baku. Model tersebut adalah sebagai berikut:

$$Y_i^* = \beta_1^* X_{1i}^* + \beta_2^* X_{2i}^* + \cdots + \beta_p^* X_{pi}^* + \varepsilon_i^*.$$

2.9 Variabel Penelitian

Definisi variabel yang digunakan dalam penelitian ini adalah definisi yang dikeluarkan BPS dan beberapa literatur lainnya. Adapun definisi tersebut adalah:

2.9.1 Indeks Harga yang Diterima Petani

Indeks Harga yang Diterima Petani (It) merupakan indikator tingkat pendapatan produsen petani. It dapat menggambarkan fluktuasi harga komoditas pertanian yang dihasilkan petani. Dari Indeks Harga yang Diterima Petani (It), dapat dilihat perkembangan harga komoditas/produk pertanian yang dijual oleh petani secara perodik. Indeks ini digunakan juga sebagai data penunjang dalam penghitungan Pendapatan Domestik Bruto/Pendapatan Domestik Regional Bruto sektor pertanian. Sebagai data pokok untuk penghitungan It ini diperlukan tiga macam data yaitu:

a. Kuantitas produksi Tiap Jenis Produk Pertanian

Data kuantitas produksi untuk subsektor tanaman pangan, tanaman hortikultura, tanaman perkebunan rakyat, peternakan, dan perikanan.

b. Harga Produsen

Harga produsen adalah harga produksi dari petani sebelum memasukkan biaya pengepakan dan transportasi ke dalam harga penjualannya atau dengan kata lain harga di ladang atau sawah setelah pemetikan (farm gate). Harga yang dicakup adalah harga transaksi dengan sistim penjualan umum atau tebasan, sedangkan penjualan dengan sistim ijon tidak dicatat karena tidak mewakili harga yang sebenarnya

c. Persentase Marketed Surplus

Persentase Marketed Surplus adalah perbandingan antara nilai produksi yang dijual petani dan nilai produksi yang dihasilkan untuk setiap jenis tanaman pertanian.

2.9.2 Indeks Harga yang Dibayar Petani

Indeks harga yang dibayar petani (Ib) merupakan indkator dari sisi kebutuhan petani baik untuk konsumsi maupun biaya produksi. Pada Ib, indeks konsumsi rumah tangga menunjukan fluktuasi harga barang/jasa yang dikonsumsi oleh petani yang merupakan bagian terbesar dari masyarakat di perdesaan dan dapat juga digunakan sebagai proxy inflasi perdesaan. Dari Indeks Harga yang Dibayar Petani (Ib), dapat dilihat perkembangan harga barang dan jasa yang dikonsumsi oleh petani di pedesaan maupun perkembangan harga barang dan jasa yang yang diperlukan untuk memproduksi komoditas/produk pertanian. Di sisi lain, indeks biaya produksi dan penambahan barang modal menunjukan fluktuasi harga barang/jasa yang digunakan untuk memproduksi komoditas pertanian. Ib adalah setiap jenis barang/jasa yang tercakup dalam pengeluaran konsumsi rumah tangga, biaya produksi dan penambahan barang modal adalah nilai setiap jenis barang yang dibeli petani, tidak termasuk nilai barang yang diproduksi sendiri.

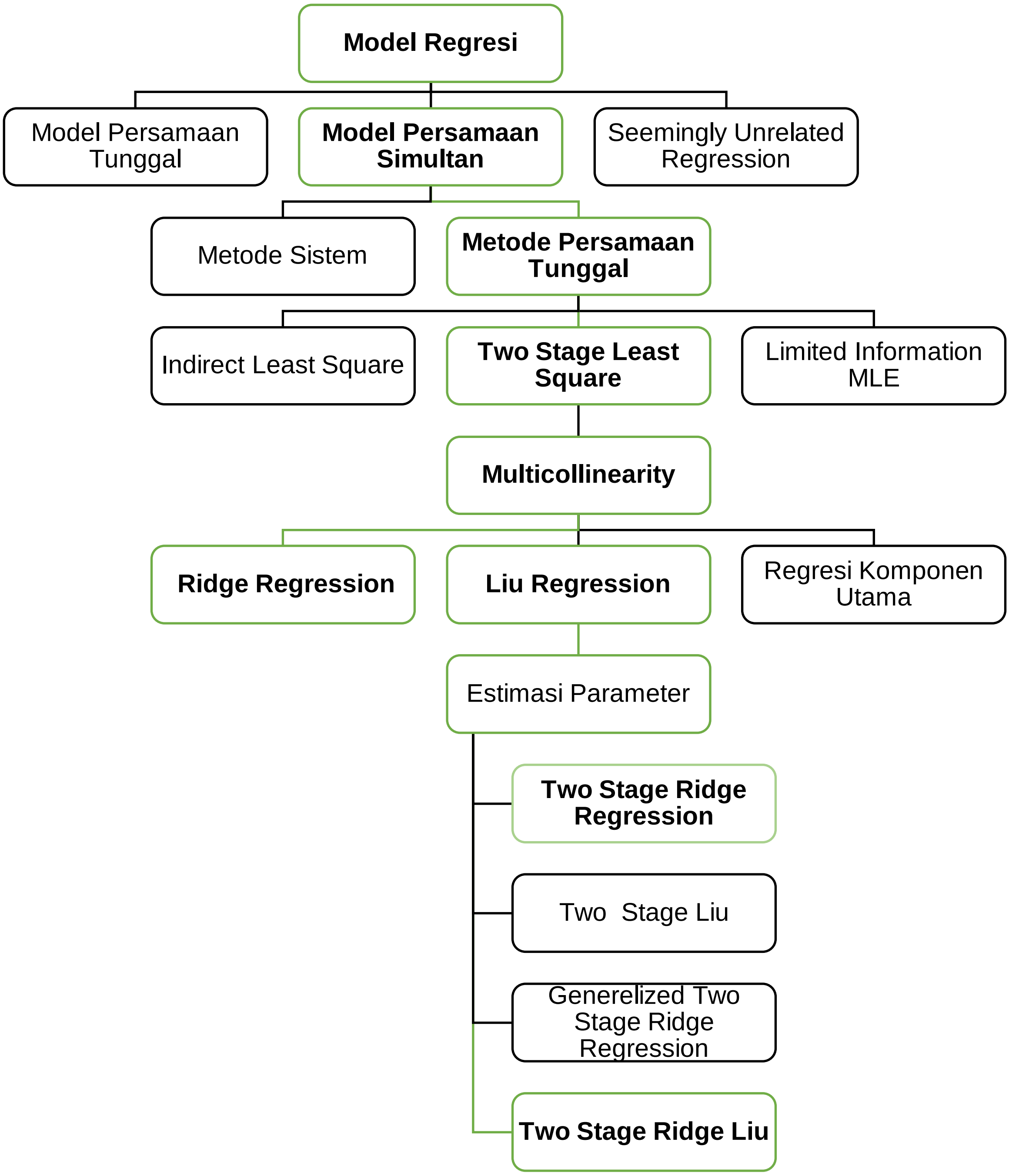
1. Konsumsi Rumah Tangga

Konsumsi rumah tangga adalah nilai konsumsi total seluruh rumah tangga petani selama setahun, maka nilai konsumsi yang didapat dari hasil SPDT NTP 2012 ini harus dikalikan dengan jumlah rumah tangga tani di perdesaan dalam periode waktu selama setahun

2. Biaya Produksi dan Penambahan Barang Modal

Biaya produksi dan penambahan barang modal (BPPBM) terdiri dari biaya bibit; pupuk, obat-obatan, dan pakan; biaya sewa dan pengeluaran lainnya; transportasi; barang modal dan upah buruh tani.

2.10 Kerangka Konseptual



Gambar 2. Kerangka Konseptual