



TESIS - SS14 2501

**MODEL TOBIT KUANTIL BAYESIAN (TKB)
*ENDOGENOUS***
**(Aplikasi pada Pengeluaran Rumah Tangga
untuk Konsumsi Susu di Kabupaten Sidoarjo)**

SARTIKA AYU WULANDARI
NRP. 06211650017014

DOSEN PEMBIMBING :
Dr. Dra. Ismaini Zain, M. Si
Dr. Santi Puteri Rahayu, M. Si

**PROGRAM MAGISTER
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2018**



TESIS - SS14 2501

**MODEL TOBIT KUANTIL BAYESIAN (TKB)
ENDOGENOUS
(Aplikasi pada Pengeluaran Rumah Tangga
untuk Konsumsi Susu di Kabupaten Sidoarjo)**

**SARTIKA AYU WULANDARI
NRP. 06211650017014**

**DOSEN PEMBIMBING :
Dr. Dra. Ismaini Zain, M.Si
Dr. Santi Puteri Rahayu, M.Si**

**PROGRAM MAGISTER
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2018**



THESIS - SS14 2501

**ENDOGENOUS BAYESIAN TOBIT QUANTILE
(BTQ) MODEL
(Application on Household Expenditure for
Milk Consumption in Sidoarjo)**

**SARTIKA AYU WULANDARI
NRP. 06211650017014**

**SUPERVISORS :
Dr. Dra. Ismaini Zain, M.Si
Dr. Santi Puteri Rahayu, M.Si**

**PROGRAM OF MAGISTER
DEPARTEMENT OF STATISTICS
FACULTY OF MATHEMATICS, COMPUTATION, AND DATA SCIENCE
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2018**

MODEL TOBIT KUANTIL BAYESIAN (TKB) ENDOGENOUS
(Aplikasi pada Pengeluaran Rumah Tangga untuk Konsumsi
Susu di Kabupaten Sidoarjo)

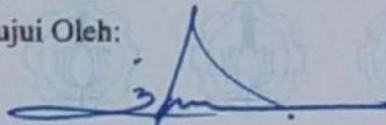
Tesis disusun untuk memenuhi salah satu syarat memperoleh gelar
Magister Sains (M.Si)

di
Institut Teknologi Sepuluh Nopember
Oleh:

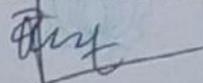
SARTIKA AYU WULANDARI
NRP. 06211650017014

Tanggal Ujian : 23 Januari 2018
Periode Wisuda : Maret 2018

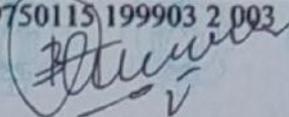
Disetujui Oleh:



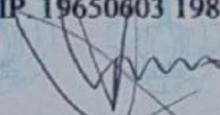
1. Dr. Ismaini Zain, M. Si. (Pembimbing I)
NIP. 19600525 198803 2 001



2. Dr. Santi Puteri Rahayu, M. Si (Pembimbing II)
NIP. 19750115 199903 2 003



3. Prof. Dr. Drs. I Nyoman Budiantara, M.Si (Penguji)
NIP. 19650603 198903 1 003

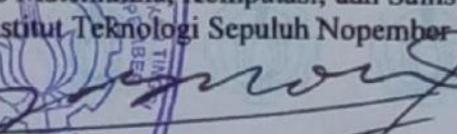


4. Dr. Kadarmanto, M. A. (Penguji)
NIP. 19710707 199312 1 001



Dekan

Fakultas Matematika, Komputasi, dan Sains Data
Institut Teknologi Sepuluh Nopember


Prof. Dr. Basuki Widodo, M.Sc.
NIP. 19650605 198903 1 002

MODEL TOBIT KUANTIL BAYESIAN *ENDOGENOUS* (Aplikasi pada Pengeluaran Rumah Tangga untuk Konsumsi Susu di Kabupaten Sidoarjo)

Nama : Sartika Ayu Wulandari
NRP : 06211650017014
Dosen Pembimbing : Dr. Dra. Ismaini Zain, M.Si
Dr. Santi Puteri Rahayu, M.Si

ABSTRAK

Pada survei-survei yang diselenggarakan oleh Badan Pusat Statistik (BPS) dengan unit observasi rumah tangga, banyak tercatat rumah tangga yang tidak mengalokasikan pengeluaran untuk beberapa jenis barang konsumsi. Hal ini menyebabkan banyak ditemukan data bersifat *mixture*, yaitu campuran antara data diskret dan kontinyu, yang dikategorikan sebagai data tersensor. Model yang dapat digunakan untuk menganalisis masalah dengan variabel respon yang merupakan data tersensor adalah model regresi tobit. Namun, estimator tobit memiliki kelemahan, yaitu menghasilkan estimasi yang bias untuk *error* yang tidak normal. Untuk mengatasinya, digunakan metode regresi kuantil dalam regresi tobit yang dapat digunakan pada data yang heterogen. Pengembangan regresi Tobit Kuantil (TK) dengan menggunakan pendekatan bayesian memberikan kemudahan secara komputasi dalam estimasi. Namun, model Tobit Kuantil Bayesian (TKB) tersebut tidak dapat diaplikasikan pada penelitian yang melibatkan variabel *endogenous* karena estimasi yang dihasilkan akan bias. Berdasarkan uraian tersebut, diperlukan penerapan dari pengembangan suatu metode baru yang fleksibel seperti model TKB *Endogenous*. Dalam model tersebut, variabel *endogenous* dan variabel *exogenous* akan diregresikan pada tahap pertama. Agar lebih fleksibel, regresi dilakukan secara non-parametrik dengan memanfaatkan *Dirichlet Process Mixture*. Pada tahap selanjutnya, regresi kuantil ke-p akan dimodelkan secara parametrik mengikuti kaidah regresi kuantil bayesian. Penelitian ini bertujuan untuk menerapkan model Tobit Kuantil Bayesian (TKB) *Endogenous* dengan mengaplikasikannya pada data pengeluaran rumah tangga untuk konsumsi susu. Estimasi parameter model akan menggunakan metode MCMC dengan algoritma *Gibbs Sampler*. Hasil dari penelitian ini adalah bahwa pada data pengeluaran rumah tangga untuk konsumsi susu di Kabupaten Sidoarjo, endogenitas signifikan pada kuantil di sebelah kiri kuantil 0,55 dan pada kuantil di sebelah kanan kuantil 0,80. Selanjutnya, model TKB *Endogenous* dibandingkan dengan model TK dan TKB dengan menggunakan RMSE dan hasilnya adalah model TKB *Endogenous* memiliki performa yang lebih baik ketika masalah endogenitas muncul.

Kata Kunci: Tobit Kuantil Bayesian (TKB), variabel *Endogenous*, *Gibbs Sampler*, konsumsi susu.

**ENDOGENOUS BAYESIAN TOBIT QUANTILE (BTQ)
MODEL
(Application on Household Expenditure
for Milk Consumption in Sidoarjo)**

Name : Sartika Ayu Wulandari
NRP : 06211650017014
Supervisors : Dr. Dra. Ismaini Zain, M.Si
Santi Puteri Rahayu, M.Si, Ph. D

ABSTRACT

In surveys conducted by Badan Pusat Statistik (BPS) with household observation units, many households do not allocate expenditures for certain types of consumer goods. This resulted a lot of mixed data, a mixture of discrete and continuous data, categorized as censored data. A model that can be used to analyze problems with response variable that are censored data is tobit model. However, the tobit estimator has the disadvantage of generating a biased estimate of non-normal errors. To overcome this problem, a quantile regression is used in tobit regression that can be used in heterogeneous data. The development of the Quantile Tobit (QT) regression using the bayesian approach provides computational ease in estimation. However, the Bayesian Tobit Quantile (BTQ) model can not be applied to studies involving endogenous variable because the estimator would be biased. Based on these descriptions, the application of the development of a new flexible method such as endogenous BTQ model is required. In the model, the endogenous and exogenous variables will be regressed in the first stage. To be more flexible, the regression is done non-parametrically utilizing Dirichlet Process Mixture. In the next step, the p -th quantile regression will be modeled parametrically following bayesian quantile regression rules. This study aims to apply the endogenous BTQ model into the household expenditure for milk consumption data in Sidoarjo. Estimation of the model parameters is done using MCMC method with Gibbs Sampler algorithm. It was found that in the household expenditure for milk consumption data in Sidoarjo, the significant endogeneity was found in the left side of quantile 0,55 and on the right side of quantile 0,80. In addition, the endogenous BTQ model is compared with the QT and BTQ models using RMSE and the result is the endogenous BTQ model performs better when the endogenous problem arises.

Keywords : Bayesian Tobit Quantile (BTQ), endogenous variable, gibbs sampler, milk consumption.

KATA PENGANTAR

Segala puji bagi Allah SWT yang Mahapengasih lagi Mahapenyayang, yang telah melimpahkan karunia dan nikmat sehingga dengan izin-Nya penyusunan tesis dengan judul **“Model Tobit Kuantil Bayesian *Endogenous* (Aplikasi pada Pengeluaran Rumah Tangga Untuk Konsumsi Susu di Kabupaten Sidoarjo)”** dapat terselesaikan.

Pada kesempatan ini penulis ingin menyampaikan ucapan terima kasih dan penghargaan yang setinggi-tingginya kepada berbagai pihak yang telah membantu penyelesaian tesis ini, baik secara moril maupun materiil:

1. Badan Pusat Statistik (BPS) yang telah memberi kesempatan serta beasiswa melalui Pusat Pendidikan dan Pelatihan (Pusdiklat) BPS kepada penulis untuk mengikuti Program Magister Statistika di ITS.
2. Ibu Dr. Dra. Ismaini Zain, M.Si dan Dr. Santi Puteri Rahayu, M. Si selaku pembimbing yang telah meluangkan banyak waktu untuk memberikan bimbingan, arahan dan petunjuk dalam menyelesaikan tesis ini dengan segala keterbatasan penulis.
3. Bapak Prof. Dr. I Nyoman Budiantara, M. Si dan Bapak Dr. Kadarmanto, M. A. selaku penguji yang telah memberikan banyak saran dan masukan untuk perbaikan penulisan tesis.
4. Bapak Dr. rer. pol. Dedy Dwi Prastyo, M.Si selaku dosen wali yang telah meluangkan waktu untuk memberikan bimbingan, arahan dan petunjuk selama perkuliahan di ITS.
5. Bapak dan Ibu dosen Statistika ITS yang telah mencurahkan ilmu dan pengalamannya selama proses studi dan seluruh staff jurusan Statistika, FMKSD ITS yang telah memberikan bantuan selama proses studi.
6. Seluruh keluargaku tercinta, terutama kedua anakku Ali Zakir Maulana dan Ahmad Quthbi Zayyan, terima kasih untuk pengertian kalian pada ibumu yang tidak sempurna.

7. Teman seperjuangan Batch 10 BPS, Ratna, Reni, Imra, Ratih, Mike, Prih, Rafael, Fendi, Sony, Taufik, Umam, Aniq, Suprpto, dan Fieldry, yang telah bersama-sama dan saling memotivasi selama menempuh pendidikan.
8. Bapak Syahrul, Ibu Arifatus, dan Ibu Sarni yang tengah menyelesaikan pendidikan doktor, yang turut membantu dan tidak pernah bosan menjawab pertanyaan penulis selama proses penyusunan tesis. Semoga Allah memberikan balasan yang terbaik.
9. Semua pihak yang telah membantu penyelesaian tesis ini.

Akhirnya, doa dan harapan selalu dipanjatkan kepada Allah SWT agar ilmu yang telah diperoleh menjadi barokah dan bermanfaat bagi sesama serta dapat menjadi sarana meraih ridho Allah. Aamiin Ya Robbal Alamin.

Surabaya, Januari 2016

Penulis

DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN.....	ix
ABSTRAK	viii
ABSTRACT.....	ix
KATA PENGANTAR.....	xi
DAFTAR ISI.....	xiii
DAFTAR TABEL	xv
DAFTAR GAMBAR.....	xvii
DAFTAR LAMPIRAN.....	xix
BAB 1 PENDAHULUAN.....	1
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah.....	5
1.3. Tujuan Penelitian	5
1.4. Manfaat Penelitian	6
1.5. Batasan Masalah	6
BAB 2 TINJAUAN PUSTAKA.....	7
2.1. Regresi Tobit	7
2.1.1. Konsep Data Tersensor	7
2.1.2. Model Regresi Tobit.....	8
2.2. Regresi Kuantil	9
2.3. Model Tobit Kuantil (TK)	11
2.4. Model Tobit Kuantil Bayesian (TKB).....	11
2.4.1. <i>Posterior Inference</i> Model TKB	12
2.4.2. Inferensia Model TKB dengan <i>Scale Parameter</i>	13
2.5. Model Tobit Kuantil Bayesian (TKB) <i>Endogenous</i>	14
2.5.1. Endogenitas	14
2.5.2. Bias pada Model TKB dengan Adanya Endogenitas	16
2.5.3. Model Tobit Kuantil Bayesian (TKB) <i>Endogenous</i>	16
2.6. Estimasi dengan Metode MCMC (Markov Chain Monte Carlo): Metode <i>Gibbs Sampler</i>	18
2.6.1. Algoritma <i>Gibbs Sampler</i>	19
2.6.2. Algoritma <i>Gibbs Sampler</i> Untuk ALDP	20
2.7. <i>Credible Interval</i> (CI)	20
2.8. Uji <i>Outlier</i>	20
2.9. <i>Root Mean Square Error</i> (RMSE).....	20
2.10. Faktor-Faktor yang Mempengaruhi Pengeluaran Rumah Tangga untuk Konsumsi Susu	23

BAB 3 METODE PENELITIAN	29
3.1. Sumber Data.....	29
3.2. Variabel Penelitian.....	29
3.3. Metode Penelitian.....	33
BAB 4 HASIL DAN PEMBAHASAN.....	37
4.1. Model TK dan TKB untuk Pengeluaran Rumah Tangga untuk Konsumsi Susu di Kabupaten Sidoarjo	37
4.1.1. Karakteristik data Pengeluaran Rumah Tangga untuk Konsumsi Susu di Kabupaten Sidoarjo	39
4.1.2. Model TK pada Pengeluaran Rumah Tangga untuk Konsumsi Susu di Kabupaten Sidoarjo	42
4.1.3. Model TKB pada Pengeluaran Rumah Tangga untuk Konsumsi Susu di Kabupaten Sidoarjo	43
4.2. Model TKB Endogenous pada Pengeluaran Rumah Tangga untuk Konsumsi Susu di Kabupaten Sidoarjo	44
4.2.1. Uji Endogenitas	44
4.2.2. Kajian Distribusi Prior Model TKB Endogenous	47
4.3. Interpretasi Model TKB Endogenous untuk Pengeluaran Rumah Tangga untuk Konsumsi Susu di Kabupaten Sidoarjo.....	47
BAB 5 KESIMPULAN DAN SARAN.....	59
5.1 Kesimpulan	59
5.2 Saran	60
DAFTAR PUSTAKA	61
LAMPIRAN-LAMPIRAN.....	65
BIOGRAFI PENULIS.....	104

DAFTAR TABEL

Tabel 3. 1. Variabel Prediktor/ <i>Exogenous</i>	30
Tabel 3. 2. Struktur Data Untuk Analisis	32
Tabel 4. 1. Statistik Deskriptif Variabel Penelitian.....	39
Tabel 4. 2. Nilai RSME Model TKB pada Kuantil ke-p.....	58
Tabel 4. 3. Nilai Eta pada Berbagai Titik Kuantil.....	48
Tabel 4. 4. Signifikansi Variabel Penelitian pada Setiap Kuantil	49
Tabel 4. 5. Ringkasan Posterior pada 30.000 kali Iterasi dan 10.000 <i>Burn In</i>	52
Tabel 4. 6. Nilai RMSE Model TKB <i>Endogenous</i> pada Kuantil ke-p.....	58

DAFTAR GAMBAR

Gambar 2. 1. Jenis Umum Hubungan Variabel dalam Persamaan Simultan.....	15
Gambar 2. 2. Hubungan antar Variabel dalam Penelitian.....	15
Gambar 2. 3. Diagram Alir Algoritma <i>Gibbs Sampler</i>	19
Gambar 2. 4. Kerangka Konseptual Penelitian	26
Gambar 3. 1. Diagram Alir Penelitian	36
Gambar 4. 1. Histogram Data Pengeluaran Rumah Tangga untuk Konsumsi Susu di Kabupaten Sidoarjo	38
Gambar 4. 2. <i>Outlier</i> pada Variabel Respon	38
Gambar 4. 3. Pendidikan KRT di rumah tangga sampel di Kabupaten Sidoarjo..	41
Gambar 4. 4. Ilustrasi <i>Prior, Likelihood, dan Posterior</i>	47
Gambar 4. 5. Nilai Eta dan 95% <i>Credible Interval</i> pada Setiap Kuantil	49
Gambar 4. 6. <i>Trace Plot</i> pada Kuantil 0,95	50
Gambar 4. 7. <i>Correlogram</i> pada Kuantil 0,95	51
Gambar 4. 8. <i>Density</i> pada Kuantil 0,95	51
Gambar 4. 9. <i>Posterior Mean</i> dari $(\beta'_p, \delta'_p)'$ untuk Model TKB <i>Endogenous</i> pada Setiap Kuantil	55

DAFTAR LAMPIRAN

Lampiran 1. Plot Variabel Respon Pengeluaran Rumah Tangga untuk Konsumsi Susu.....	65
Lampiran 2. <i>Output</i> Uji Endogenitas dan Eksogenitas untuk Regresi Tobit dengan Software Stata12	66
Lampiran 3. Uji Multikolinieritas dengan <i>Software</i> Stata12	67
Lampiran 4. <i>Syntax</i> ALDP dengan <i>Software</i> OxMetrics v6	68
Lampiran 5. <i>Syntax Subroutine</i> dengan <i>Software</i> OxMetrics v6	82
Lampiran 6. <i>Syntax Main Program</i> TKB <i>Endogenous</i> dengan <i>Software</i> Oxmetrics v6	86
Lampiran 7. <i>Output</i> TKB <i>Endogenous</i> pada Masing-Masing Kuantil	87
Lampiran 8. <i>Trace plot</i> , <i>Corelogram</i> , dan <i>Density</i> parameter Beta.....	100

BAB 1

PENDAHULUAN

1.1. Latar Belakang

Pada survei-survei yang diselenggarakan oleh Badan Pusat Statistik (BPS) dimana rumah tangga adalah sebagai unit observasinya, misalnya Survei Sosial Ekonomi Nasional (SUSENAS), banyak tercatat rumah tangga yang tidak mengalokasikan pengeluaran untuk beberapa jenis barang konsumsi. Hal ini menyebabkan banyak ditemukannya data dengan struktur bernilai nol untuk sebagian observasi dan bernilai tertentu yang bervariasi untuk sebagian observasi lainnya. Menurut Greene (2000), struktur data dengan variabel respon yang mempunyai sifat campuran (*mixture*) antara diskrit dengan kontinyu, diskrit untuk yang bernilai nol dan kontinyu untuk yang tidak nol, dikategorikan sebagai data tersensor.

Long (1997) menjelaskan bahwa jika seluruh data tersensor digunakan dalam suatu model linier, maka akan dihasilkan nilai parameter yang *overestimates* pada *slope* dan *underestimates* pada *intercept*. Sebaliknya, jika informasi yang nilainya nol dihilangkan dan tidak digunakan, maka model akan memiliki *slope* yang *underestimate* dan *intercept* yang *overestimates*. Data yang terpotong akan menimbulkan korelasi antara variabel prediktor dengan residual sehingga estimasi yang dihasilkan menjadi tidak konsisten.

Sehubungan dengan hal tersebut, maka diperlukan model khusus untuk menjelaskan hubungan antara variabel prediktor dan variabel respon yang bersifat tersensor. Jika variabel respon bersifat biner, maka hubungan antara keduanya dapat dijelaskan melalui model probit atau logit. Namun nilai variabel respon yang bervariasi di dalam batasan menjadi tidak efisien karena tidak akan digunakan dalam kedua model tersebut. Dengan demikian, metode statistik yang dapat mengakomodirnya adalah regresi linier. Namun asumsi dalam regresi linier menjadi tidak terpenuhi dan hasil estimasi parameter menjadi bias (Greene, 2008). Untuk mengatasi permasalahan pelanggaran asumsi pada OLS, maka James Tobin

(1958) mengusulkan penggunaan model Tobit. Zain dan Suhartono (1997) juga menjelaskan bahwa model yang dapat digunakan untuk menganalisis masalah dengan variabel respon yang merupakan data tersensor adalah model regresi tobit.

Model regresi tobit telah banyak dikembangkan dan digunakan dalam berbagai bidang penelitian. Diantaranya adalah Fitriani (2011) yang menganalisis faktor-faktor yang mempengaruhi pengeluaran rumah tangga untuk konsumsi susu dengan menggunakan regresi tobit dan Suhardi dan Llewelyn (2011) yang menggunakan model regresi tobit untuk menganalisis faktor-faktor yang berpengaruh terhadap Kepuasan Konsumen untuk Jasa Pengangkutan Barang.

Estimasi parameter regresi tobit diperoleh dengan menggunakan metode *Maximum Likelihood Estimation* (MLE) yang berbasis *conditional mean*. Metode ini hanya mampu menduga model dari fungsi bersyarat *mean* dan tidak mewakili keseluruhan data dari distribusi (Hao dan Naiman, 2007). Terdapat metode lain yang mampu menggambarkan hubungan antara variabel prediktor terhadap variabel respon pada berbagai titik kuantil (*conditional quantile*). Metode ini dapat memberikan hasil yang tepat dan stabil pada kehadiran *outlier* serta dapat membatasi pengaruh dari *outlier* (Furno, 2014). Dengan demikian, penggunaan regresi kuantil pada regresi tobit dapat digunakan pada kondisi data yang heterogen dengan menggambarkan hubungan variabel prediktor dengan variabel respon pada berbagai titik kuantil dari distribusi variabel respon tersebut.

Regresi Tobit Kuantil (TK) muncul secara bertahap sebagai pendekatan yang komprehensif, diawali oleh Powell (1986). Penelitian yang mengkaji dan mengaplikasikan model tersebut adalah Cintiani (2017) yang memodelkan pengeluaran rumah tangga untuk konsumsi rokok dengan model regresi kuantil tersensor.

Pada kerangka kerja berbasis bayesian, Yu dan Stander (2007) memperkenalkan regresi Tobit Kuantil Bayesian (TKB) sebagai pengembangan atas model regresi kuantil bayesian oleh Yu dan Moyeed (2001) dengan mengaplikasikan metode estimasi menggunakan *Markov Chain Monte Carlo* (MCMC). Kozumi dan Kobayashi (2011) kemudian mengaplikasikan metode *Gibbs Sampler* yang hasil estimasinya lebih efisien. Dari bidang non-parametrik, Taddy dan Kottas (2010) memeriksa model semi-parametrik dan non-parametrik

menggunakan *Dirichlet process mixture models*. Ogasawara dan Kobayashi (2015) memanfaatkan model parametrik sederhana menggunakan dua distribusi *Asymmetric Laplace* untuk regresi kuantil panel. Meskipun demikian, metode ini hanya dapat diaplikasikan pada data tak tersensor dan terbatas pada bentuk distribusi *Asymmetric Laplace* yang dapat mempengaruhi estimasi.

Pada beberapa penelitian, sering terdapat variabel *endogenous*, yaitu variabel yang didalamnya mencakup semua variabel *instrumental* dan variabel *exogenous*. Belum banyak studi yang mengestimasi regresi kuantil tersensor dengan prediktor *endogenous*, sehingga mengembangkan model regresi kuantil yang menghitung secara simultan untuk variabel respon dengan melibatkan variabel *endogenous* menjadi tantangan (Kobayashi, 2017).

Kobayashi (2017) mengembangkan suatu metode parametrik baru yang fleksibel yaitu model TKB *endogenous*. Dalam penelitiannya diperkenalkan model regresi tobit kuantil ke-p dengan melibatkan variabel *endogenous*. Pada tahap pertama, kuantil ke- α dari error pada regresi variabel *endogenous* dan variabel *exogenous*, yang disebut regresi tahap pertama, diasumsikan nol. Selanjutnya, error dari tahap pertama akan dilibatkan pada model regresi kuantil ke-p dimana *quantile conditional* ke-p dari error yang baru adalah nol. Distribusi error tahap pertama dimodelkan dengan pendekatan parametrik dan semiparametrik. Dalam pendekatan semiparametriknya, digunakan *Dirichlet process mixture* dengan tujuan mendapatkan model yang lebih fleksibel. Nilai α adalah prior yang tidak diketahui dan dianggap sebagai parameter yang tidak pasti dan akan diestimasi dari data. Hasil penelitian tersebut adalah bahwa model *Asymmetric Laplace* (AL), *Asymmetric Laplace Dirichlet Process* (ALDP), dan *Skew Normal Dirichlet Process* (SNDP) memberikan hasil yang baik pada berbagai situasi. Di lain pihak, model *Skew Normal* (SN) tidak dapat mengakomodir ekor yang tebal pada error tahap pertama dan *Asymmetric Exponential Power* (AEP) tidak cukup fleksibel.

Penelitian yang mengaplikasikan model TKB adalah penelitian oleh Lusiana (2015) yang membahas kajian tentang estimasi parameter dan metode pemilihan model regresi TKB terbaik, kemudian hasil estimasinya dibandingkan dengan estimator Powell (metode BRCENS) dengan teknik simulasi. Model yang

diperoleh diaplikasikan pada data pengeluaran rumah tangga untuk konsumsi susu. Kesimpulan yang diperoleh dari penelitian tersebut adalah estimator TKB lebih baik daripada estimator tobit standar untuk model yang memiliki *error* berdistribusi bukan normal, kecuali pada kuantil-kuantil bawah. Namun, penelitian tersebut tidak memasukkan variabel pendapatan rumah tangga ke dalam modelnya karena bersifat nonlinier.

Pemodelan data yang dilakukan Lusiana (2015) tersebut belum memisahkan variabel *endogenous* dan variabel *exogenous* pada regresinya. Karena itu, pada pemodelan data penelitian ini akan memisahkan variabel *endogenous* yang akan diregresikan dengan variabel-variabel *exogenous* dan variabel *instrumental* terlebih dahulu pada tahap pertama dan diregresikan dalam model TKB pada tahap kedua.

Lusiana mengasumsikan error berdistribusi *Asymmetric Laplace* (AL). Namun Kobayashi (2017) menyebutkan bahwa distribusi AL mempunyai keterbatasan, seperti densitas yang puncaknya runcing, ekor yang mempunyai batasan, dan kemencengan. Ketika model tidak sesuai dengan data, maka estimasi titik kuantil akan jauh dari nilai yang sebenarnya. Jika pada penelitian tersebut mengasumsikan variabel kontrol dari regresi tahap pertama adalah homoskedastis, maka estimasi *intercept* akan menjadi bias. Konsekuensinya, hasil estimasi dari parameter di tahap kedua akan terpengaruh.

Berdasarkan uraian di atas, maka penelitian ini akan mengaplikasikan model TKB *endogenous* dengan memanfaatkan model alternatif untuk distribusi error tahap pertama regresi TKB *endogenous*, yaitu *Asymmetric Laplace Dirichlet Process* (ALDP) sebagai pengembangan AL dengan memanfaatkan *Dirichlet process mixture*. Regresi kuantil ke-p akan dimodelkan secara parametrik mengikuti kaidah regresi kuantil bayesian seperti pada Kobayashi (2017).

Hal yang cukup penting dalam penggunaan metode bayesian adalah penentuan *prior* yang sesuai. Interpretasi dari *prior* adalah sebagai bentuk distribusi frekuensi, sebagai bentuk representasi normatif dan objektif pada suatu parameter yang lebih rasional untuk dipercayai, dan Sebagai suatu representasi subjektifitas seseorang dalam memandang sebuah parameter menurut penilainnya sendiri. Permasalahan pokok agar *prior* dapat interpretatif adalah bagaimana

memilih distribusi prior untuk suatu parameter yang tidak diketahui namun sesuai dengan permasalahan fisik yang ada. Dengan demikian kajian penentuan prior menjadi hal yang penting.

Berdasarkan uraian di atas, maka tujuan dari penelitian ini adalah untuk memodelkan data pengeluaran rumah tangga untuk konsumsi susu di Kabupaten Sidoarjo dengan model TK, TKB, dan TKB *Endogenous* dengan menggunakan variabel-variabel yang berdasar pada kajian teori dan penelitian-penelitian sebelumnya. Untuk model TKB *Endogenous*, dilakukan kajian *prior* yang akan digunakan untuk estimasi *posterior*. Dari model TK, TKB, dan TKB *Endogenous* yang diperoleh, masing-masing diukur performanya dengan kriteria nilai *Root Mean Square Error* (RMSE).

1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka dirumuskan dua hal, yaitu: (1) bagaimana model TK dan TKB untuk data pengeluaran rumah tangga untuk konsumsi susu di Kabupaten Sidoarjo? (2) bagaimana penentuan *prior* yang akan digunakan untuk estimasi posterior pada model TKB *Endogenous* untuk data pengeluaran rumah tangga untuk konsumsi susu di Kabupaten Sidoarjo? dan (2) bagaimana model TKB *Endogenous* untuk data pengeluaran rumah tangga untuk konsumsi susu di Kabupaten Sidoarjo dan bagaimana endogenitas di setiap kuantilnya?

1.3. Tujuan Penelitian

Tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut:

1. Memodelkan TK dan TKB untuk data pengeluaran rumah tangga untuk konsumsi susu di Kabupaten Sidoarjo.
2. Mengkaji penentuan *prior* yang akan digunakan untuk estimasi posterior pada model TKB *Endogenous* untuk data pengeluaran rumah tangga untuk konsumsi susu di Kabupaten Sidoarjo.

3. Memodelkan TKB *Endogenous* untuk data pengeluaran rumah tangga untuk konsumsi susu di Kabupaten Sidoarjo dan mendapatkan penjelasan masalah endogenitas di setiap kuantilnya.

1.4. Manfaat Penelitian

Manfaat dari penelitian ini adalah sebagai berikut:

1. Memperoleh model TK dan TKB untuk data pengeluaran rumah tangga untuk konsumsi susu di Kabupaten Sidoarjo.
2. Mendapatkan kajian penentuan *prior* yang akan digunakan untuk estimasi posterior pada model TKB *Endogenous* untuk data pengeluaran rumah tangga untuk konsumsi susu di Kabupaten Sidoarjo.
3. Memperoleh model TKB *Endogenous* untuk data pengeluaran rumah tangga untuk konsumsi susu di Kabupaten Sidoarjo dan mendapatkan penjelasan masalah endogenitas di setiap kuantilnya, sehingga diharapkan dapat digunakan sebagai bahan perencanaan, monitoring dan sebagai tolok ukur evaluasi keberhasilan, khususnya terkait peningkatan gizi melalui konsumsi susu.

1.5. Batasan Masalah

Batasan dalam penelitian ini adalah sebagai berikut:

1. Ukuran yang digunakan untuk mengevaluasi kebaikan model adalah RMSE.
2. Kuantil yang akan dimodelkan adalah kuantil 0,025, 0,05, 0,25, 0,50, 0,75, 0,95, 0,975, dan 0.99.

BAB 2

TINJAUAN PUSTAKA

Pada bagian ini akan disajikan kajian teori yang terkait dalam penelitian. Di antaranya adalah pengantar mengenai konsep data tersensor yang digunakan dalam penelitian, model regresi tobit, model regresi kuantil, model TK, TKB, konsep endogenitas, bias pada model dengan adanya masalah endogenitas, dan estimasi model TKB *Endogenous*, serta RMSE sebagai kriteria kebaikan model. Dari sisi non-statistik, disajikan kajian teori terkait pengeluaran rumah tangga untuk konsumsi susu.

2.1. Regresi Tobit

2.1.1. Konsep Data Tersensor

Jenis data tersensor merupakan data yang memuat nilai nol pada sebagian pengamatan sedangkan untuk sebagian lain mempunyai nilai tertentu yang bervariasi. Variabel respon y_i^* dikatakan tersensor bila mengikuti persamaan:

$$y_i = \begin{cases} y_i^* & , \text{ jika } y_i^* > 0 \\ 0 & , \text{ jika } y_i^* \leq 0 \end{cases} \quad (2.1)$$

dimana $i = 1, 2, \dots, n$ dan n adalah banyaknya observasi dan y_i^* adalah variabel respon tersensor.

Long (1997) menjelaskan bahwa jika seluruh data tersensor digunakan dalam suatu model linier, maka akan dihasilkan nilai parameter yang *overestimates* pada *slope* dan *underestimates* pada *intercept*. Sebaliknya, jika informasi yang nilainya nol dihilangkan dan tidak digunakan, maka model akan memiliki *slope* yang *underestimate* dan *intercept* yang *overestimates*. Data yang terpotong akan menimbulkan korelasi antara variabel prediktor dengan *residual* sehingga estimasi yang dihasilkan menjadi tidak konsisten. Untuk mengatasi permasalahan-permasalahan tersebut, maka James Tobin (1958) mengusulkan penggunaan model Tobit.

2.1.2. Model Regresi Tobit

Model Tobit standart yang diperkenalkan oleh James Tobin (1958) adalah sebagai berikut:

$$y_i^* = \mathbf{x}'_i \boldsymbol{\beta} + \epsilon_i \quad , \quad \epsilon_i \sim \text{iid } N(0, \sigma^2)$$

atau dapat dituliskan:

$$y_i = \max(y_i^*, 0) \quad \text{atau} \quad y_i = \max(\mathbf{x}'_i \boldsymbol{\beta} + \epsilon_i, 0) \quad (2.2)$$

dimana:

y_i^* : variabel respon,

\mathbf{x}'_i : vektor variabel prediktor,

$\boldsymbol{\beta}$: matriks koefisien parameter berukuran $k \times 1$ yang tidak diketahui dengan k adalah banyaknya parameter,

ϵ_i : error yang berdistribusi normal tersensor (identik dan independen) dengan *mean* nol dan varian σ^2 .

Dengan $y_i^* \sim N(\mu, \sigma^2)$, maka *probability density function* (pdf)-nya adalah:

$$f(y^* | \mu, \sigma) = \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{y^* - \mu}{\sigma} \right)^2 \right] \quad (2.3)$$

Jika dituliskan dalam bentuk pdf Normal Standar, maka menjadi:

$$f(y^* | \mu, \sigma) = \frac{1}{\sigma} \left\{ \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{y^* - \mu}{\sigma} \right)^2 \right] \right\} = \frac{1}{\sigma} \phi \left(\frac{y^* - \mu}{\sigma} \right) \quad (2.4)$$

dimana $\phi(\cdot)$ merupakan bentuk pdf dari distribusi normal standar.

Karena $\mu = E(y_i^*) = \mathbf{x}'_i \boldsymbol{\beta}$, maka:

$$f(y^* | \mathbf{x}'_i \boldsymbol{\beta}, \sigma) = \frac{1}{\sigma} \phi \left(\frac{y^* - \mathbf{x}'_i \boldsymbol{\beta}}{\sigma} \right) \quad (2.5)$$

Sedangkan *cummulative distribution funcion* (cdf)-nya adalah:

$$f(y^* = 0 | \mathbf{x}'_i \boldsymbol{\beta}, \sigma) = \Phi \left(\frac{\mathbf{x}'_i \boldsymbol{\beta}}{\sigma} \right) \quad (2.6)$$

dengan $\Phi(\cdot)$ merupakan bentuk cdf dari distribusi normal standar dan $\epsilon_i / \sigma \sim N(0,1)$.

Metode Tobit mengasumsikan bahwa variabel-variabel prediktor tidak terbatas nilainya (*non-censored*), variabel respon bersifat *censored*, tidak ada *autocorrelation*, tidak ada *heteroscedascity*, dan tidak ada multikolinearitas yang sempurna (Gujarati, 2004).

Parameter model regresi Tobit biasanya diestimasi dengan metode MLE (*Maximum Likelihood Estimation*). Fungsi likelihood untuk model regresi tobit dibedakan menjadi dua yaitu untuk $y_i^* > 0$ dan untuk $y_i^* \leq 0$ dan dapat dituliskan sebagai berikut:

$$L(\boldsymbol{\beta}) = \prod_{i=1}^n \left[1(y_i > 0) \cdot \frac{1}{\sigma} \phi\left(\frac{y_i - \mathbf{x}_i' \boldsymbol{\beta}}{\sigma}\right) \right] \left[1(y_i = 0) \left\{ 1 - \Phi\left(\frac{\mathbf{x}_i' \boldsymbol{\beta}}{\sigma}\right) \right\} \right] \quad (2.7)$$

dan fungsi *log-likelihood*nya adalah:

$$\begin{aligned} l(\boldsymbol{\beta}) &= \log L(\boldsymbol{\beta}) \\ &= \sum_{i=1}^n 1(y_i > 0) \cdot \log \left\{ \frac{1}{\sigma} \phi\left(\frac{y_i - \mathbf{x}_i' \boldsymbol{\beta}}{\sigma}\right) \right\} + 1(y_i = 0) \cdot \log \left\{ 1 - \Phi\left(\frac{\mathbf{x}_i' \boldsymbol{\beta}}{\sigma}\right) \right\} \end{aligned} \quad (2.8)$$

Dengan mengasumsikan bahwa σ^2 diketahui, maka estimator parameter diperoleh dari turunan pertama persamaan (2.8) disamadengankan dengan nol. Namun penyelesaian persamaan ini bersifat nonlinier sehingga digunakan metode iterasi Newton-Raphson untuk menyelesaikannya.

2.2. Regresi Kuantil

Regresi Kuantil adalah regresi yang dapat menjelaskan hubungan antar variabel respon dan variabel prediktor pada berbagai titik kuantil dari distribusi variabel respon tersebut. Hal ini menjadikan regresi ini dapat digunakan pada kondisi data yang heterogen, berbeda dengan regresi linier berganda yang hanya mampu menjelaskan hubungan sebab akibat pada *mean (conditional mean)* variabel respon (Koenker dan Hallock, 2001).

Metode regresi kuantil diperkenalkan pertama kali oleh Roger Koenker dan Gilbert Basset (1978). Regresi ini dapat memberikan gambaran yang lebih lengkap tentang bagaimana variabel prediktor berhubungan dengan distribusi bersyarat yang mendasari variabel respon, terutama ketika distribusi tidak mengikuti distribusi normal standar. Regresi Kuantil sangat berguna ketika data tidak homogen dan tidak simetris, terdapat ekor pada sebaran atau *truncated distribution*.

Pada regresi linier $E(y|\mathbf{x}_i) = \mathbf{x}_i' \boldsymbol{\beta}$, sementara dalam regresi kuantil $Q_{y|\mathbf{x}_i}(p) = \mathbf{x}_i' \boldsymbol{\beta}_p$ yang dapat diuraikan menjadi:

$$y_i = \beta_{p,0} + \sum_{k=1}^r \beta_{pk} x_{ik} + \varepsilon_{pi} \quad , \quad i = 1, 2, \dots, n \quad (2.9)$$

dimana $\beta_{p0}, \beta_{p1}, \dots, \beta_{pk}$ merupakan parameter yang tidak diketahui pada kuantil ke- p dan ε_{pi} merupakan residual dari model regresi pada sampel sebanyak dan pada kuantil ke- p.

Dalam bentuk matriks, regresi kuantil dapat ditulis sebagai berikut:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & \dots & x_{1r} \\ 1 & x_{21} & \dots & x_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \dots & x_{nr} \end{bmatrix} \begin{bmatrix} \beta_{p0} \\ \beta_{p1} \\ \vdots \\ \beta_{pr} \end{bmatrix} + \begin{bmatrix} \varepsilon_{p1} \\ \varepsilon_{p2} \\ \vdots \\ \varepsilon_{pk} \end{bmatrix} \quad (2.10)$$

Bentuk umum model regresi kuantil linier adalah sebagai berikut:

$$y_i = \mathbf{x}'_i \boldsymbol{\beta}_p + \varepsilon_p, \quad 0 < p < 1 \quad (2.11)$$

dimana:

y_i : nilai variabel respon ke-i

\mathbf{x}'_i : $(1, x_{1i}, x_{2i}, \dots, x_{ki})$

$\boldsymbol{\beta}_p$: parameter model regresi pada kuantil ke-p

ε_p : error model regresi kuantil ke-p

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

Menurut Chen (2005), estimasi parameter model regresi kuantil dilakukan dengan menyatakan fungsi peluang kumulatif dari variabel random Y, $F(y) = P(Y \leq y)$ sehingga kuantil ke-p dari variabel tersebut dapat dituliskan sebagai berikut:

$$Q_p(Y) = \inf\{y: F(y) \geq p\} \quad (2.12)$$

Kemudian pada kuantil ke-p dari $F(y)$ dapat diperoleh dengan meminimumkan fungsi tersebut terhadap Q, yaitu:

$$p \int_{y>q} |y - q| dF_Y(y) + (1 - p) \int_{y<q} |y - q| dF_Y(y) \quad (2.13)$$

Jika Y merupakan fungsi X yang diketahui memiliki fungsi probabilitas $F_{Y|X}(y)$, maka kuantil ke-p dari fungsi tersebut dapat dituliskan menjadi $Q_{Y_i}(p|x)$ yang merupakan fungsi dari X dan diselesaikan dengan persamaan berikut:

$$\min_q p \int_{i=1, y>q}^n |y - q| dF_Y(y) + (1 - p) \int_{i=1, y<q}^n |y - q| dF_Y(y) \quad (2.14)$$

Dalam regresi kuantil terdapat fungsi kuantil bersyarat ke-p yang mempertimbangkan penduga β_p sehingga diperoleh solusi untuk permasalahan tersebut yang dinyatakan sebagai berikut:

$$\beta(p)^* = \min_{\beta} \left\{ p \sum_{i=1, y > x}^n |y - \mathbf{x}'\beta| + (1 - p) \sum_{i=1, y < x}^n |y - \mathbf{x}'\beta| \right\} \quad (2.15)$$

Seperti pada metode OLS yang meminimumkan jumlah kuadrat residual untuk estimasi, dengan metode *Least Absolute Deviation* (LAD) estimasi β dalam regresi kuantil pada persamaan (2.11) dilakukan dengan meminimumkan *loss function*. Nilai β yang meminimumkan kuadrat error dengan metode OLS yaitu:

$$\min_{\beta_p} \sum_{i=1}^n \rho_p(y_i - \mathbf{x}'_i \beta_p) \quad (2.16)$$

dimana $\rho_p(\cdot)$ merupakan *check (or loss) function* yang didefinisikan dengan

$$\rho_p(u) = u(p - \mathbf{I}(u < 0)) \text{ atau } \rho_p(u) = \begin{cases} pu & , u \geq 0 \\ (p - 1)u & , u < 0 \end{cases} \text{ dimana } p \in (0,1) \text{ dan}$$

$\mathbf{I}(\cdot)$ merupakan fungsi indikator (Powell, 1986).

Persamaan (2.14) tidak *close form*, maka metode iterasi numerik biasa tidak dapat digunakan untuk menyelesaikan persamaan tersebut. Sehingga digunakan metode pemrograman linier yaitu metode simpleks (Koenker dan Machado, 1999).

2.3. Model Tobit Kuantil (TK)

Jika variabel respon dituliskan dengan notasi $y_i = c(y_i^*) = \max\{y_i^*, 0\}$, dengan $i = 1, \dots, n$, maka model regresi kuantil ke- p untuk y_i^* dapat dituliskan:

$$y_i^* = \mathbf{x}'_i \beta_p + \epsilon_i \quad (2.17)$$

dimana y_i^* adalah variabel respon, \mathbf{x}'_i merupakan vektor variabel prediktor, β_p merupakan koefisien, dan ϵ_i merupakan error yang memiliki kuantil ke- p sama dengan nol sebagaimana $\int_{-\infty}^0 f_p(\epsilon_i) d\epsilon_i = p$.

Kuantil ke- p dari y^* dapat dituliskan dengan $Q_{y^*|x}(p) = \mathbf{x}'\beta_p$. Di bawah transformasi $c(\cdot)$ dari kuantil, maka kuantil ke- p dari y diberikan oleh persamaan:

$$Q_{y|x}(p) = c\left(Q_{y^*|x}(p)\right) \quad (2.18)$$

sehingga model tobit kuantil dapat diestimasi dengan solusi dari:

$$\min_{\beta_p} \sum_{i=1}^n \rho_p(y_i - \max\{\mathbf{x}'_i \beta_p\}) \quad (2.19)$$

Estimator regresi TK bisa diperoleh menggunakan solusi minimasi metode pemrograman linear. Kemudian dikembangkan tiga tahap algoritma yang

dikenalkan oleh Chernozhukov dan Hong (2002) dimana tahapan yang digunakan adalah mengambil sub sampel dengan cara pemisahan dari probabilitas tersensor, dan melakukan dua kali estimasi menggunakan kuantil regresi. Hasil dari estimasi pertama adalah mendapatkan sub sampel yang sesuai, kemudian hasil dari estimasi yang kedua adalah untuk membuat estimasi yang efisien.

2.4. Model Tobit Kuantil Bayesian (TKB)

Pendekatan Bayesian mengasumsikan ϵ mengikuti distribusi *Asymmetric Laplace*, karena dengan meminimumkan persamaan (2.19) sama dengan memaksimumkan fungsi likelihood dari distribusi *Asymmetric Laplace* (Koenker dan Machado, 1999).

2.4.1. Posterior Inference Model TKB

Menurut Kozumi dan Kobayashi (2011) dengan mengikuti Yu dan Moyeed (2001), $\epsilon_i \sim \text{Asymmetric Laplace}$ dengan *density*:

$$f_p(\epsilon) = p(1-p)\exp\{-\rho_p(\epsilon)\} \quad (2.20)$$

dengan *mean* dan varian diperoleh dari penurunan sebagai berikut:

$$\begin{aligned} \varphi(t) &= E[e^{it\epsilon}] \\ &= \int_{-\infty}^{\infty} e^{it\epsilon} f_p(\epsilon) d\epsilon \\ &= \int_{-\infty}^0 p(1-p)e^{it\epsilon+(1-p)\epsilon} d\epsilon + \int_0^{\infty} p(1-p)e^{it\epsilon-p\epsilon} d\epsilon \\ &= p(1-p) \left\{ \frac{1}{it+(1-p)} + \frac{1}{p-it} \right\} \\ &= \left\{ \frac{1}{p(1-p)} t^2 - i \frac{1-2}{p(1-p)} t + 1 \right\}^{-1} \quad \text{dimana } i^2 = -1 \end{aligned} \quad (2.21)$$

Dengan $\epsilon = \theta z + \tau\sqrt{z}u$ dimana z variabel *exponential* standar dan u variabel normal standar, fungsi karakteristik dari $\epsilon' = \theta z + \tau\sqrt{z}u$ dapat dituliskan:

$$\phi(t) = E[e^{it(\theta z + \tau\sqrt{z}u)}] = \int_0^{\infty} e^{it\theta z} E[e^{it\tau\sqrt{z}u}] e^{-z} dz. \quad (2.22)$$

Karena u berdistribusi normal standar, maka:

$$E[e^{it\tau\sqrt{z}u}] = e^{-\frac{1}{2}t^2\tau^2z}. \quad (2.23)$$

Oleh karena itu, diperoleh fungsi karakteristik dari ϵ' sebagaimana berikut:

$$\phi(t) = \int_0^\infty e^{-z(1+\frac{1}{2}t^2\tau^2-i\theta t)} dz = \left(\frac{1}{2}t^2\tau^2 - i\theta t + 1\right)^{-1}. \quad (2.24)$$

Kedua fungsi karakteristik sama ketika $E[\epsilon] = \theta = \frac{1-2p}{p(1-p)}$ dan $\text{Var}(\epsilon) = \tau^2 = \frac{1-2p+2p^2}{p^2(1-p)^2}$.

Dari hasil tersebut, variabel respon y_i dapat dituliskan:

$$y_i = \mathbf{x}'_i \boldsymbol{\beta} + \theta z_i + \tau \sqrt{z_i} u_i \quad (2.25)$$

dimana $z_i \sim \mathcal{E}(1)$ dan $u_i \sim N(0,1)$ saling independen, dan $\mathcal{E}(\psi)$ adalah distribusi eksponensial dengan *mean* ψ . Conditional distribution dari y_i given z_i adalah normal dengan *mean* $\mathbf{x}'_i \boldsymbol{\beta} + \theta z_i$ dan varian $\tau^2 z_i$, sehingga *joint density* dari $\mathbf{y} = (y_1, y_2, \dots, y_n)'$ adalah:

$$f(\mathbf{y} | \boldsymbol{\beta}_p, \mathbf{z}) \propto \left(\prod_{i=1}^n z_i^{-\frac{1}{2}} \right) \exp \left\{ - \sum_{i=1}^n \frac{(y_i - \mathbf{x}'_i \boldsymbol{\beta}_p - \theta z_i)^2}{2\tau^2 z_i} \right\} \quad (2.26)$$

dimana $\mathbf{z} = (z_1, z_2, \dots, z_n)'$.

2.4.2. Inferensia Model TKB dengan *Scale Parameter*

Dengan memasukkan *scale parameter* $\sigma > 0$, maka variabel respon dapat dituliskan kembali:

$$y_i^* = \mathbf{x}'_i \boldsymbol{\beta}_p + \sigma \theta z_i + \sigma \tau \sqrt{z_i} u_i = \mathbf{x}'_i \boldsymbol{\beta}_p + \theta v_i + \tau \sqrt{\sigma v_i} u_i \quad (2.27)$$

dimana $v_i \sim \mathcal{E}(\sigma)$ dan $u_i \sim N(0,1)$.

Probability Density Function (pdf) dari distribusi *Asymmetric Laplace* dengan *scale parameter* $\sigma > 0$ dan $p \in (0,1)$ adalah *shape parameter* yang dinotasikan $AL(\sigma, p)$ adalah:

$$f_{AL(\epsilon|\sigma,p)} = \frac{p(1-p)}{\sigma} \exp \left\{ - \frac{\rho_p(\epsilon)}{\sigma} \right\}, \quad -\infty < x < \infty \quad (2.28)$$

Sesuai dengan Kozumi dan Kobayashi (2011) di atas, maka *mean* dan varian dari distribusi $AL(\sigma, p)$ adalah $E[\epsilon] = \sigma \frac{1-2p}{p(1-p)}$ dan $\text{Var}(\epsilon) = \sigma^2 \frac{1-2p+2p^2}{p^2(1-p)^2}$.

Kuantil ke- p dari distribusi ini adalah nol, $\int_{-\infty}^0 f(\epsilon) = p$. Dengan mengasumsikan distribusi *prior*-nya, maka estimasi parameter dilakukan dengan metode MCMC (Yu dan Stander, 2007).

Karena y_i^* pada persamaan (2.27) memiliki regresi dengan bentuk *conditional* pada v_i , maka metode yang dikembangkan oleh Chib (1992) dapat diaplikasikan untuk melakukan sampling terhadap y_i^* sebagaimana:

$$y_i^* | \mathbf{y}, \boldsymbol{\beta}_p, \mathbf{v}, \sigma \sim y_i I(y_i > 0) + \mathcal{TN}_{(-\infty, 0)}(\mathbf{x}'_i \boldsymbol{\beta}_p + \theta v_i, \tau^2 \sigma v_i) I(y_i = 0) \quad (2.29)$$

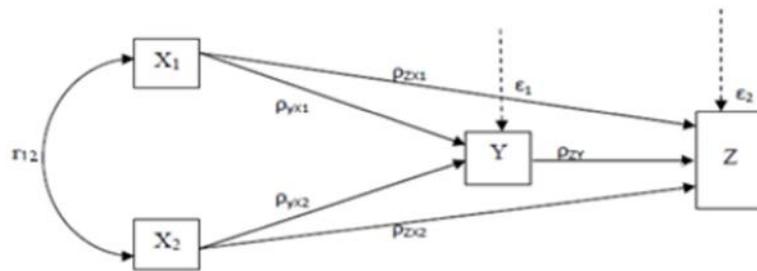
dimana $\mathcal{TN}_{(a,b]}(\mu, \sigma^2)$ adalah distribusi normal dengan mean μ dan varian σ^2 *truncated* pada interval $(a, b]$.

2.5. Model Tobit Kuantil Bayesian (TKB) *Endogenous*

2.5.1. Endogenitas

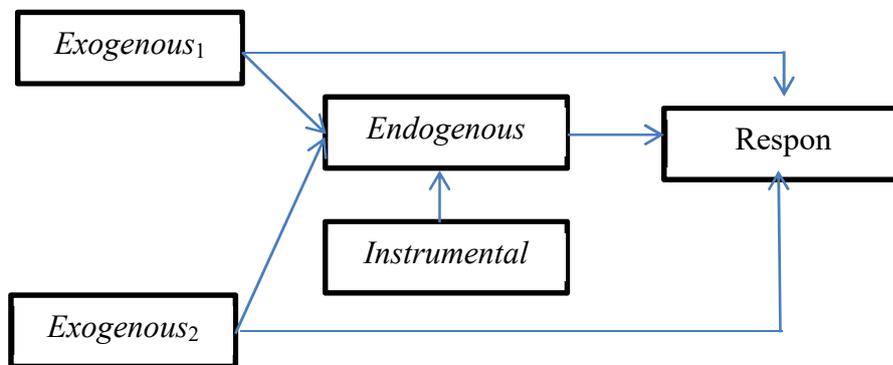
Masalah endogenitas selalu menyebabkan masalah simultanitas. Gujarati (2004) menjelaskan bahwa dalam persamaan simultan terdapat beberapa jenis variabel, di antaranya variabel *exogenous*, variabel *endogenous* dan variabel *instrumental*. Variabel *exogenous* adalah variabel yang dianggap memiliki pengaruh terhadap variabel yang lain, namun tidak dipengaruhi oleh variabel lain dalam model. Sedangkan Variabel *endogenous* adalah variabel yang dianggap dipengaruhi oleh variabel lain dalam model. Variabel *endogenous* merupakan variabel perantara yang dapat dipandang sebagai efek dari variabel *exogenous* atau variabel perantara lainnya, dan merupakan penyebab terhadap variabel perantara lainnya dan variabel respon, serta dapat berfungsi sebagai variabel respon itu sendiri. Sedangkan variabel *instrumental* adalah variabel yang mempengaruhi variabel *endogenous* di luar variabel *eksogenous*, namun tidak mempengaruhi variabel respon.

Hubungan antara Variabel *Exogenous*, Variabel *Endogenous*, dan Variabel Respon sebagaimana yang dicontohkan oleh Riduwan dan Kuncoro (2011) sebagai Jenis Umum pada *Path Analysis* dapat dilihat dari Gambar 2.1.



Gambar 2. 1. Jenis Umum Hubungan Variabel dalam Persamaan Simultan

Jika alur tersebut diterjemahkan dalam penelitian ini, hubungan antar variabel dapat dijelaskan sebagaimana Gambar 2.2.



Gambar 2. 2. Hubungan antar Variabel dalam Penelitian

Uji endogenitas yang biasanya digunakan untuk model tobit adalah uji eksogenitas seperti yang diusulkan oleh Smith dan Blundell (1986). Dengan model berikut:

$$y_i^* = \beta_0 + \beta_1 x_i + \delta d_i + u_i \text{ dan} \\ d_i = \gamma_0 + \gamma_1 x_i + \gamma_2 w_i + v_i, \text{ untuk } i = 1, 2, \dots, n \quad (2.30)$$

maka hipotesis pada pengujian ini adalah:

H_0 : $\text{corr}(u_i, v_i) = 0$ atau tidak terdapat variabel *endogenous*.

H_1 : $\text{corr}(u_i, v_i) \neq 0$ atau terdapat variabel *endogenous* yang merupakan fungsi dari satu set variabel *exogenous* dan *instrumental*.

Keputusan H_0 ditolak jika $p\text{-value} < 0,05$. Dengan menolak H_0 , maka model tobit standar tidak bisa digunakan karena akan menghasilkan estimasi yang bias untuk parameternya.

2.5.2. Bias pada Model TKB dengan Adanya Endogenitas

Estimasi untuk model TKB akan bias ketika variabel *endogenous* dimasukkan sebagai prediktor. Dengan model sebagaimana persamaan (2.30), dimana $(\beta_0, \beta_1, \delta) = (1, 1, 1)$, $(\gamma_0, \gamma_1, \gamma_2) = (1, 1, 1)$, $x_i, w_i \sim N(0, 1)$ dan $\begin{pmatrix} u_i \\ v_i \end{pmatrix} \sim N(\mathbf{0}, \Sigma)$ dengan $\Sigma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$, dan ρ menunjukkan tingkat endogenitas, maka variabel d bersifat *exogenous* jika $\rho = 0$ dan bersifat *endogenous* jika $\rho \neq 0$.

Karena $u|v \sim N(\rho v, 1 - \rho^2)$ model dapat ditulis sebagai berikut:

$$y_i^* = \beta_0 + \beta_1 x_i + \delta d_i + \rho v + \sqrt{1 - \rho^2} u_i. \quad (2.31)$$

Oleh karena itu, model standar yang memodelkan *conditional quantile* dari y^* dengan $\beta_0 + \beta_1 x_i + \delta d_i$ menghasilkan estimasi yang bias.

2.5.3. Model Tobit Kuantil Bayesian (TKB) *Endogenous*

Model TKB dengan variabel *endogenous* (Kobayashi, 2017) dapat dituliskan sebagai berikut:

$$d_i = \mathbf{z}_i' \boldsymbol{\gamma} + v_i \quad \text{dengan} \quad \mathbf{z}_i = (\mathbf{x}_i', w_i) ', \quad i = 1, \dots, n \quad (2.32)$$

$$y_i^* = \mathbf{x}_i' \boldsymbol{\beta}_p + \delta_p d_i + \eta_p (d_i - \mathbf{z}_i' \boldsymbol{\gamma}) + e_i \quad (2.33)$$

dimana \mathbf{x}_i adalah vektor variabel *exogenous* dengan elemen pertama adalah 1, d_i adalah variabel *endogenous*, w_i adalah variabel prediktor yang tidak termasuk dalam \mathbf{x}_i (disebut juga variabel *instrumental*), dan $v_i = d_i - \mathbf{z}_i' \boldsymbol{\gamma}$ diperkenalkan sebagai variabel kontrol. Dengan $\eta_p \neq 0$, maka hal ini mengindikasikan bahwa d_i adalah variabel *endogenous*.

Berdasarkan Lee (2007), error ϵ_i dari model TKB standar didekomposisikan menjadi $\eta_p (d_i - \mathbf{z}_i' \boldsymbol{\gamma})$ dan e_i . Diasumsikan bahwa persamaan pada regresi tahap pertama dispesifikasikan dengan benar dan kebebasan kuantil dari e_i pada \mathbf{z}_i *conditional* pada v_i adalah sebagai berikut:

$$Q_{\epsilon|d, \mathbf{z}}(p) = Q_{\epsilon|v, \mathbf{z}}(p) = Q_{\epsilon|v}(p) = \eta_p (d - \mathbf{z}' \boldsymbol{\gamma}) \quad (2.34)$$

diasumsikan juga:

$$Q_{v|\mathbf{z}}(\alpha) = 0 \quad (2.35)$$

dimana *conditional quantile* ke- α dari v_i adalah nol untuk beberapa $\alpha \in (0, 1)$.

Adapun tahapan regresi TKB *endogenous* adalah sebagai berikut:

a. Tahap pertama

Ogasawara dan Kobayashi (2015) mengasumsikan $v_i \sim AL(\phi, \alpha)$, $i = 1, \dots, n$. Namun distribusi AL mempunyai keterbatasan, seperti densitas yang puncaknya runcing, ekor yang mempunyai batasan, dan kemencengan. Ketika model tidak sesuai dengan data, maka estimasi *conditional quantile* akan jauh dari nilai yang sebenarnya (Kobayashi, 2017). Selain itu, dengan mengasumsikan v_i homoskedastis, estimasi *intercept* (γ_0) akan menjadi bias. Konsekuensinya, estimasi β_{p0} akan dipengaruhi mengingat adanya $\eta_p(d_i - \mathbf{z}_i' \boldsymbol{\gamma})$. Ketika v_i adalah heteroskedastis, maka keseluruhan vektor koefisien tidak akan terpengaruh. Oleh karena itulah, perlu model alternatif untuk distribusi *error* tahap pertama, misalkan ALDP.

Pengembangan model ALDP berdasarkan distribusi AL dengan menggunakan *dirichlet process mixture* dapat dituliskan sebagai berikut (Kobayashi, 2017):

$$f_{ALDP}(v|G) = \int f_{AL}(v|\phi, \alpha) dG(\phi), \quad G \sim DP(a, G_0) \quad (2.36)$$

dengan $DP(a, G_0)$ adalah *Dirichlet process* dengan parameter presisi $a > 0$ dan ukuran dasar G_0 .

b. Tahap kedua

Di tahap kedua, seperti pada regresi kuantil bayesian, diasumsikan $e_i \sim AL(\sigma, p)$, $i = 1, \dots, n$ dan *conditional quantile* ke- p dari e_i adalah nol. Model diekspresikan dalam bentuk hirarki sebagai berikut:

$$y_i = \max\{y_i^*, 0\}$$

$$y_i^* \sim N(\tilde{\mathbf{x}}_i' \tilde{\boldsymbol{\beta}}_p + \theta_p g_i, \tau_p^2 \sigma g_i)$$

$$g_i \sim \varepsilon(\sigma), \quad \text{untuk } i = 1, \dots, n$$

dimana:

$$\tilde{\mathbf{x}}_i = (\mathbf{x}_i', d_i, d_i - \mathbf{z}_i' \boldsymbol{\gamma})'$$

$$\tilde{\boldsymbol{\beta}}_p = (\boldsymbol{\beta}_p', \delta_p, \eta_p)'$$

$\varepsilon(\sigma)$ adalah distribusi eksponensial dengan mean σ , dan

$$\theta_p = \frac{1-2p}{p(1-p)}, \quad \text{dan } \tau_p^2 = \frac{2}{p(1-p)} \quad (2.37)$$

2.6. Estimasi dengan Metode MCMC (Markov Chain Monte Carlo): Metode *Gibbs Sampler*

Salah satu solusi untuk mengatasi masalah penentuan distribusi posterior yang sulit dilakukan karena melibatkan persamaan integral yang sangat kompleks adalah dengan pendekatan numerik, yaitu MCMC (Carlin dan Chib, 1995). Pendekatan MCMC sangat efektif untuk mengurangi beban komputasi dalam menyelesaikan persamaan integrasi yang kompleks. Selain itu, metode ini memungkinkan proses simulasi dengan mengambil sampel random dari model stokastik yang sangat rumit. Ide dasarnya adalah membangkitkan data sampel dari distribusi *posterior* sesuai proses *Markov Chain* dengan menggunakan simulasi *Monte Carlo* secara iteratif sehingga diperoleh kondisi yang konvergen terhadap *posterior*. Selanjutnya, sampel parameter dalam *Markov Chain* diambil setelah kondisi stasioner tercapai sehingga sampel yang terambil dijamin merupakan sampel dari distribusi *posterior* dari parameter tersebut.

Terdapat dua algoritma utama dalam MCMC, yaitu algoritma *Metropolis-Hastings* dan algoritma *Gibbs Sampler*. *Gibbs Sampler* dapat didefinisikan sebagai suatu teknik simulasi untuk membangkitkan variabel random dari suatu fungsi distribusi tertentu tanpa harus menghitung fungsi densitasnya. *Gibbs Sampler* bisa diterapkan apabila distribusi probabilitas bersama (*joint probability distribution*) tidak diketahui secara eksplisit, tetapi distribusi bersyarat (*conditional distribution*) dari tiap-tiap variabel diketahui. *Gibbs Sampler* merupakan generator yang sangat efisien sehingga sering digunakan sebagai generator variabel random pada analisis data yang menggunakan MCMC. Proses ini dilakukan dengan mengambil sampel dengan cara membangkitkan rangkaian *gibbs* variabel random berdasarkan sifat-sifat dasar proses *Markov Chain*. Dalam menjalankan program yang menggunakan rantai markov dilakukan pada kondisi bersyarat penuh. Ini merupakan salah satu kelebihan dari *Gibbs Sampler* karena variabel random tersebut dibangkitkan dengan menggunakan konsep distribusi unidimensional yang terstruktur sebagai distribusi *full conditional*. *Gibbs Sampler* sangat berguna dalam mengestimasi suatu parameter dalam suatu model kompleks yang mempunyai tingkat kerumitan dalam proses integritasi yang kompleks pula dan sulit diselesaikan secara analitis.

2.6.1. Algoritma Gibbs Sampler

Algoritma *Gibbs Sampler* untuk setiap parameter β bisa dituliskan sebagai berikut:

Langkah 1: Tentukan nilai awal β pada $t=0$, yaitu $\beta^0 = (\beta_1^{(0)}, \dots, \beta_p^{(0)})'$

Langkah 2: Untuk $t=1, \dots, M$, ulangi langkah:

- a. Tentukan $\beta = \beta^{(t-1)}$
- b. Untuk $s=1, 2, \dots, p$, *update* β_s dari $\beta_s \sim f(\beta_s | \beta_{\setminus s}, \mathbf{y})$
- c. Tentukan $\beta^{(t)} = \beta$ dan gunakan untuk membangkitkan iterasi ke $t+1$.

Proses sampling untuk mendapatkan nilai $\beta^{(t)}$ adalah:

$\beta_1^{(t)}$ dari $f(\beta_1 | \beta_2^{(t-1)}, \beta_3^{(t-1)}, \dots, \beta_p^{(t-1)}, \mathbf{y})$

$\beta_2^{(t)}$ dari $f(\beta_2 | \beta_1^{(t)}, \beta_3^{(t-1)}, \dots, \beta_p^{(t-1)}, \mathbf{y})$

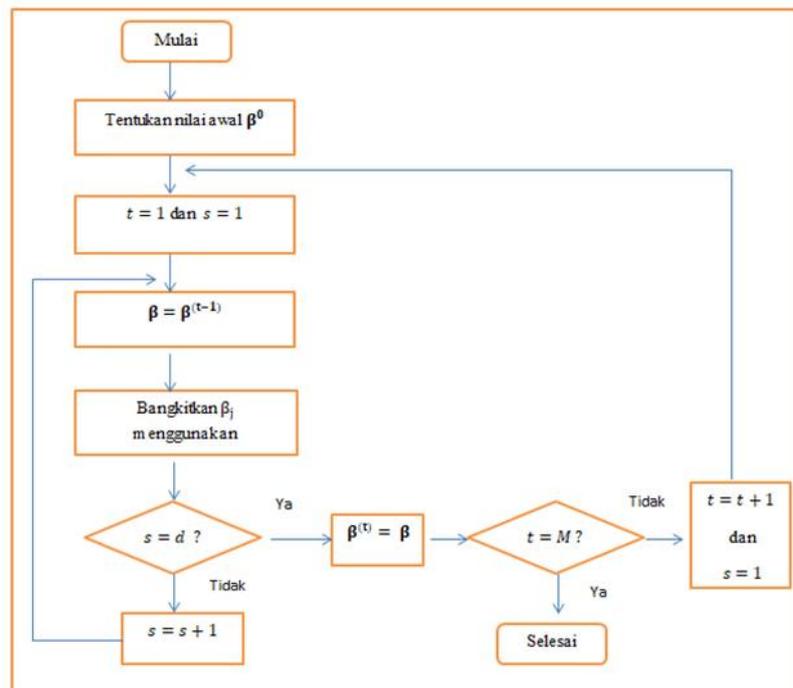
⋮

$\beta_s^{(t)}$ dari $f(\beta_s | \beta_1^{(t)}, \beta_2^{(t-1)}, \dots, \beta_{s-1}^{(t)}, \beta_{s+1}^{(t-1)}, \dots, \beta_p^{(t-1)}, \mathbf{y})$

⋮

$\beta_p^{(t)}$ dari $f(\beta_p | \beta_1^{(t)}, \beta_2^{(t)}, \dots, \beta_{p-1}^{(t-1)}, \mathbf{y})$

Untuk lebih jelasnya, algoritma tersebut dapat dilihat dari Gambar 2.3.



Gambar 2. 3. Diagram Alir Algoritma Gibss Sampler

Pembangkitan nilai dari:

$$f(\beta_s | \beta_{\setminus s}, \mathbf{y}) = f(\beta_s | \beta_1^{(t)}, \beta_2^{(t-1)}, \dots, \beta_{s-1}^{(t)}, \beta_{s+1}^{(t-1)}, \dots, \beta_p^{(t-1)}, \mathbf{y})$$

adalah relatif mudah karena merupakan distribusi univariat dan dapat ditulis sebagai $f(\theta_s | \beta_{\setminus s}, \mathbf{y}) \propto f(\beta | \mathbf{y})$ dimana variabel lain selain β_j adalah konstan.

2.6.2. Algoritma *Gibbs Sampler* Untuk ALDP

Pada penelitian Kobayashi (2017), estimasi dilakukan dengan menggunakan algoritma *Gibbs Sampler* sebagai pengembangan dari *Gibbs Sampler* yang digunakan Ogasawara dan Kobayashi (2015). Variabel-variabel yang terlibat dalam *Dirichlet process* disampel dengan menggunakan *retrospective sampler* dan *slice sampler*. Pertama diperkenalkan $u_i \sim \mathcal{U}(0,1)$ dan $k_i, i = 1, \dots, n$ dengan $\pi_l = \Pr(k_i = l), l = 1, \dots, \infty$.

Gibbs Sampler dibangun dengan mengikuti *joint densities*:

$$f_{\text{ALDP}}(v_i, u_i) = \sum_{l=1}^{\infty} I(u_i, \omega_l) f_{\text{AL}}(v_i | \phi_l, \alpha) \quad (2.38)$$

dimana $\phi_l \sim G_0$, $\pi_l = \omega_l \prod_{l < r} (1 - \tau \omega_r)$, $\omega_l \sim \mathcal{B}(1, a)$, dan $\mathcal{B}(a, b)$ adalah notasi dari distribusi Beta dengan parameter a dan b (Sethuraman, 1994). Dengan k^* adalah bilangan integer minimum, maka $\sum_{l=1}^{k^*} \pi_l > 1 - \min\{u_1, \dots, u_n\}$.

Untuk model ALDP, Kobayashi (2017) memanfaatkan representasi *mixture* untuk distribusi AL pada sampel $\boldsymbol{\gamma}$ secara efisien dengan $v_i | h_i \sim N(\theta_\alpha h_i, \tau_\alpha^2 \phi_i h_i)$ dimana $h_i \sim \varepsilon(\phi_i)$, untuk $i = 1, \dots, n$ dengan θ_p dan τ_p^2 seperti pada persamaan (2.33).

Proses *Gibbs Sampler* dilakukan dengan melakukan *sampling* terhadap $\{u_i\}_{i=1}^n$, $\{\omega_l\}_{l=1}^{k^*}$, $\{k_i\}_{i=1}^{k^*}$, $\{\phi_l\}_{l=1}^{k^*}$, a , γ , $\{h_i\}_{i=1}^n$, α , $\{y_i^*\}_{i=1}^n$, $\tilde{\boldsymbol{\beta}}_p$, σ , dan $\{g_i\}_{i=1}^n$ seperti pada Kobayashi (2017).

2.7. Credible Interval (CI)

Credible Interval (CI) merupakan estimasi interval dari parameter yang digunakan dalam bayesian. Estimasi *credible interval* mendefinisikan probabilitas posterior sehingga parameter terletak dalam interval tersebut (Raudenbush dan Bryk, 2002). King, dkk (2009) menyatakan jika β adalah parameter yang akan

diestimasi dengan distribusi posterior, $f(\beta|y)$, maka $100(1-\alpha)\%$ *credible interval* untuk β adalah:

$$P(\beta \in [a, b]|y) = \int_a^b f(\beta|y)d\beta = 1 - \alpha, \quad 0 < \alpha < 1 \quad (2.39)$$

Persamaan (2.35) menunjukkan bahwa β adalah variabel random dengan selang yang *fixed*. *Credible Interval* akan unik jika memenuhi kondisi sebagai berikut:

$$\int_{-\infty}^a f(\beta|y)d\beta = \int_b^{\infty} f(\beta|y)d\beta = \frac{\alpha}{2} \text{ dengan batas bawah (a) adalah kuantil } \frac{\alpha}{2} \text{ dan}$$

batas atas (b) adalah kuantil $1 - \frac{\alpha}{2}$.

Credible Interval dapat digunakan sebagai uji signifikansi untuk β dengan dasar hipotesis:

$H_0 : \beta = 0$ (parameter β signifikan)

$H_1 : \beta \neq 0$ (parameter β tidak signifikan)

Keputusan didasarkan pada kriteria tolak H_0 jika *credible interval* tidak mengandung nilai nol (Koop, 2003).

2.8. Uji *Outlier*

Outlier adalah pengamatan yang nilai sisaannya jauh lebih besar daripada sisaan pengamatan lain. *Outlier* ini terletak tiga atau empat simpangan baku, atau lebih jauh lagi dari rata-rata sisaannya. Keberadaan *outlier* akan menimbulkan beberapa masalah, diantaranya akan mengubah atau mengaburkan kesimpulan yang dibuat oleh peneliti karena nilai estimator parameternya bersifat bias (Draper dan Smith, 1992). *Outlier* pada variabel respon dideteksi menggunakan *Studentized Deleted Residual (TRES) Test*. Hipotesis dari uji ini adalah sebagai berikut:

H_0 : pengamatan ke-i bukan merupakan *outlier*

H_1 : pengamatan ke-i merupakan *outlier*

Statistik Uji yang digunakan adalah:

$$|TRES|_i = \frac{Y_i - \hat{Y}_{i(i)}}{\sqrt{\frac{MSE(i)}{1-h_{ii}}}}, \quad i=1,2,\dots, n \quad (2.40)$$

dengan MSE adalah jumlah kuadrat residual tanpa pengamatan ke- i dan h_{ii} adalah elemen diagonal ke- i dari matriks $\mathbf{H}=\mathbf{X}'_i(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_i$. Keputusan tolak H_0 adalah jika $|\text{TRES}|_i > t_{\alpha/2, n-k-2}$ dengan k adalah jumlah variabel yang digunakan. Sebaliknya, H_0 gagal ditolak jika $|\text{TRES}|_i \leq t_{\alpha/2, n-k-2}$.

Jenis *outlier* dibedakan menjadi *Outlier* berpengaruh dan tidak berpengaruh. *Outlier* berpengaruh merupakan pengamatan yang berpengaruh besar pada tanda dan besar nilai estimator parameter apabila dibuang. *Outlier* berpengaruh dapat diketahui menggunakan *The Difference in Fit Statistics (DFFITS) Test*. Hipotesis dari uji ini adalah sebagai berikut:

H_0 : pengamatan ke- i bukan merupakan *outlier* berpengaruh

H_1 : pengamatan ke- i merupakan *outlier* berpengaruh

Statistik Uji yang digunakan adalah:

$$|\text{DFITS}|_i = \left[\frac{Y_i - \hat{Y}_{i(i)}}{\sqrt{\frac{\text{MSE}_{(i)}}{1-h_{ii}}}} \right] \left[\frac{h_{ii}}{1-h_{ii}} \right]^{1/2}, \quad i=1,2,\dots, n \quad (2.41)$$

dengan MSE adalah jumlah kuadrat residual tanpa pengamatan ke- i dan h_{ii} adalah elemen diagonal ke- i dari matriks $\mathbf{H}=\mathbf{X}'_i(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_i$. Keputusan tolak H_0 adalah jika $|\text{DFITS}|_i > 2\sqrt{\frac{k+1}{n}}$ dengan k adalah jumlah variabel yang digunakan.

Sebaliknya, H_0 gagal ditolak jika $|\text{TRES}|_i \leq 2\sqrt{\frac{k+1}{n}}$.

2.9. Root Mean Square Error (RMSE)

Kebaikan suatu model dapat diukur dari tingkat kesalahannya. Semakin kecil tingkat kesalahan, maka berarti semakin baik model tersebut. Terdapat beberapa kriteria untuk menentukan tingkat kesalahan, salah satu yang sering digunakan untuk evaluasi pada model tobit adalah *Root Mean Square Error (RMSE)*.

RMSE merupakan salah satu ukuran yang merupakan selisih antara nilai yang diprediksi oleh suatu model dengan nilai yang sebenarnya. Model dengan

nilai RMSE yang paling kecil merupakan model yang performanya lebih baik dibandingkan dengan yang lain. Formula RMSE adalah sebagaimana persamaan:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n e_i^2} \quad (2.42)$$

2.10. Faktor-Faktor yang Mempengaruhi Pengeluaran Rumah Tangga untuk Konsumsi Susu

Konsumsi adalah kegiatan menghabiskan *utility* (nilai guna) barang dan jasa. Barang meliputi barang tahan lama dan barang tidak tahan lama. Barang konsumsi menurut kebutuhannya, yaitu kebutuhan primer, kebutuhan sekunder, dan kebutuhan tersier.

Badan Pusat Statistik (BPS) mendefinisikan pengeluaran konsumsi rumah tangga adalah pengeluaran atas barang dan jasa oleh rumah tangga untuk tujuan konsumsi. Data konsumsi secara tidak langsung dapat menunjukkan informasi pendapatan masing-masing rumah tangga. Selain itu, pemanfaatan data konsumsi juga bisa dipergunakan untuk mengukur penentuan status kemiskinan penduduk (BPS Provinsi Jawa Timur, 2016).

Susu merupakan salah satu komoditi penting dalam usaha peningkatan gizi masyarakat untuk meningkatkan kecerdasan generasi muda dan sehat di masa mendatang. Namun, hingga saat ini upaya peningkatan gizi di Indonesia ini masih terkendala oleh rendahnya konsumsi susu masyarakat Indonesia. Konsumsi ideal susu segar dua gelas tiap hari belum menjadi budaya di masyarakat Indonesia. Dengan kata lain, konsumsi susu per kapita di Indonesia masih rendah, bahkan jika dibandingkan dengan negara ASEAN yang lain. Pada 2015, tercatat konsumsi susu segar di Indonesia hanya 12,1 liter per kapita per tahun, kalah jauh dibanding dengan Malaysia yang sudah 36,2 liter per kapita per tahun, Myanmar yang mencapai 26,7 kilogram per kapita per tahun, Thailand 22,2 kilogram per kapita per tahun, dan Philipina 17,8 kilogram per kapita per tahun.

Masih rendahnya kesadaran masyarakat Indonesia untuk mengkonsumsi susu tersebut terindikasi dari banyaknya rumah tangga yang sedikit atau sama sekali tidak mengalokasikan pengeluaran untuk konsumsi susu. Faktor yang

mempengaruhinya diantaranya adalah faktor sosial ekonomi dan demografi, antara lain usia, pendapatan, pendidikan, pekerjaan dan budaya.

Menurut Hatirli, Ozkan, dan Aktas (2004), jumlah anak, ukuran rumah tangga, tingkat pendidikan dan pendapatan merupakan faktor penting yang mempengaruhi pengeluaran untuk konsumsi susu. Cecep (2000), pernah melakukan penelitian tentang konsumsi susu pada remaja dengan menggunakan variabel pendapatan orang tua, pendidikan ayah, pendidikan ibu, pekerjaan ibu, pengetahuan gizi, sikap dan uang jajan terhadap konsumsi susu. Sebelumnya Purnomo (2008) pernah melakukan penelitian mengenai faktor-faktor yang mempengaruhi pengeluaran konsumsi daging dan susu dengan menggunakan data hasil SUSENAS provinsi Jambi tahun 2006 dengan menggunakan faktor-faktor antara lain pendapatan rumah tangga, tingkat pendidikan kepala rumah tangga, persentase pengeluaran makanan, jumlah anggota rumah tangga, persentase anggota rumah tangga yang bekerja, persentase anggota rumah tangga usia dibawah 12 tahun, rata-rata pengeluaran per kapita dan daerah tempat tinggal.

Lusiana (2015) pernah meneliti pengeluaran rumah tangga untuk konsumsi susu dan memodelkannya dengan pendapatan rumah tangga, pendidikan kepala rumah tangga, persentase pengeluaran untuk makanan, jumlah anggota rumah tangga, persentase anggota rumah tangga ≤ 12 tahun, persentase anggota rumah tangga yang bekerja, dan rata-rata pengeluaran perkapita, serta membedakan pola konsumsi antara rumah tangga pedesaan dan perkotaan.

Pendidikan merupakan suatu investasi yang penting. Dengan mendapatkan pendidikan yang baik, maka seseorang berpeluang untuk mendapatkan pekerjaan yang baik pula. Maka dari itu, dengan pendidikan seseorang atau rumah tangga dapat meningkatkan kesejahteraannya. Menurut Todaro (2000) alasan pokok mengenai pengaruh dari pendidikan formal terhadap distribusi pendapatan adalah adanya korelasi positif antara pendidikan seseorang dengan penghasilan yang akan diperolehnya. Di lain pihak, Menurut Rahardja dkk (2005) semakin tinggi pendidikan seseorang pengeluaran konsumsinya juga akan semakin tinggi, sehingga mempengaruhi pola konsumsi dan hubungannya positif. Pada saat seseorang atau keluarga memiliki pendidikan yang tinggi, kebutuhan hidupnya semakin banyak.

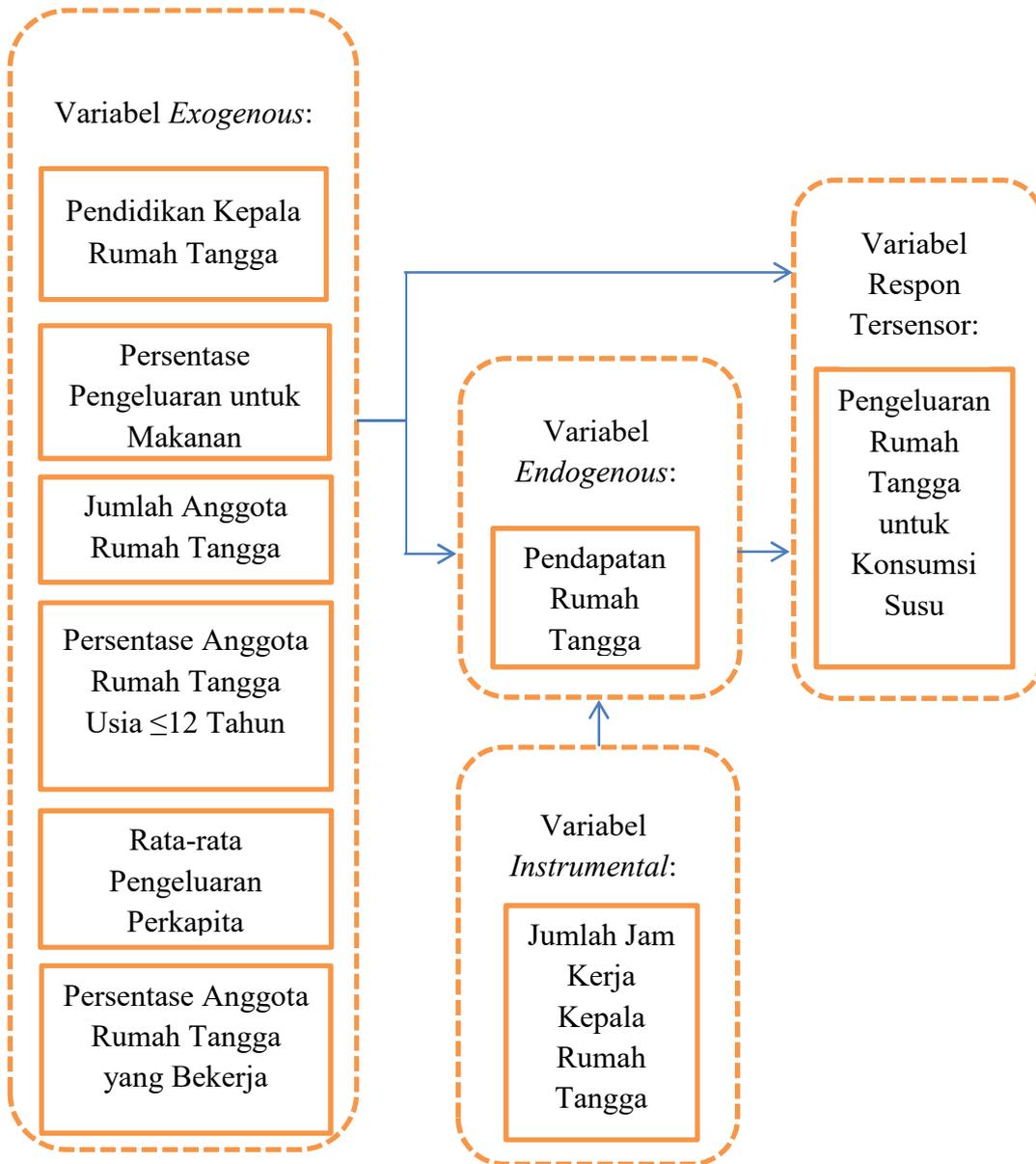
Rumah tangga dengan pengeluaran untuk makanan yang tinggi tergolong rumah tangga dengan tingkat kesejahteraan rendah dibandingkan rumah tangga yang proporsi pengeluaran untuk pangannya rendah. Sesuai dengan Hukum Engel, bagian pendapatan yang digunakan untuk belanja makanan cenderung menurun jika pendapatannya meningkat. Hal ini mengindikasikan bahwa struktur pola konsumsi merupakan salah satu indikator tingkat kesejahteraan rumah tangga. Dengan kata lain konsumsi rumah tangga untuk makanan, berpengaruh pada pendapatan rumah tangga tersebut dan akhirnya berpengaruh positif pada pengeluaran konsumsi rumah tangga.

Jumlah tanggungan rumah tangga merupakan salah satu faktor yang mempengaruhi pola konsumsi rumah tangga. Dengan banyaknya anggota keluarga, maka pola konsumsinya semakin bervariasi karena masing-masing anggota rumah tangga belum tentu mempunyai selera yang sama. Jumlah anggota keluarga berkaitan dengan pendapatan rumah tangga secara negatif yang akhirnya akan mempengaruhi pengeluaran rumah tangga untuk konsumsi.

Umur anggota keluarga berperan dalam penentuan jenis kebutuhan yang diperlukan. Keberadaan anak usia ≤ 12 tahun menyebabkan pola konsumsi yang berbeda mengingat usia tersebut merupakan interval usia untuk tumbuh kembang anak. Pola kecepatan pertumbuhan badan anak menurut Salimo (1994) yaitu masa pertumbuhan cepat (0-2 tahun), masa pertumbuhan lambat (3-12 tahun), masa Pertumbuhan cepat kembali (13-18 tahun), dan masa pertumbuhan lambat sampai dengan berhenti (19 tahun-berhenti). Pada usia sekolah (6-12 tahun) yang berada pada fase pertumbuhan lambat, anak membutuhkan energi dan asupan gizi yang lebih baik. Hal ini menjelaskan bahwa keberadaan anak usia ≤ 12 tahun akan berpengaruh secara positif terhadap pengeluaran rumah tangga untuk konsumsi susu dimana susu merupakan salah satu penyempurna kebutuhan gizi penunjang tumbuh kembang anak.

Teori Konsumsi Keynes terkenal dengan Teori Konsumsi dengan Hipotesis Pendapatan Absolut (*Absolute Income Hypothesis*) yang pada intinya menjelaskan bahwa konsumsi seseorang dan atau masyarakat secara absolut ditentukan oleh tingkat pendapatan, kalau ada faktor lain yang juga menentukan, maka menurut Keynes semuanya tidak terlalu berpengaruh. Semakin besar

pendapatan, semakin besar pula pengeluaran untuk konsumsi. Namun demikian data pendapatan yang akurat sulit diperoleh, sehingga dalam berbagai survei data pendapatan didekati melalui pengeluaran rumah tangga.



Gambar 2. 4. Kerangka Konseptual Penelitian

Berdasarkan uraian tersebut, maka secara umum dapat disimpulkan bahwa pendapatan rumah tangga bersifat *endogenous*. Pendapatan rumah tangga dipengaruhi oleh pendidikan kepala rumah tangga, persentase pengeluaran untuk makanan, jumlah anggota rumah tangga, persentase anggota rumah tangga ≤ 12

tahun, persentase anggota rumah tangga yang bekerja, dan rata-rata pengeluaran perkapita. Kemudian secara bersama-sama mempengaruhi pengeluaran rumah tangga untuk konsumsi susu.

Siswanta (2011) menyebutkan bahwa curahan jam kerja adalah waktu dengan satuan jam yang digunakan untuk bekerja pada kegiatan yang menghasilkan pendapatan atau penghasilan secara langsung maupun tidak. Berdasarkan hal tersebut, maka diperoleh informasi bahwa seringkali pendapatan rumah tangga akan dipengaruhi oleh banyak tidaknya jam kerja, terutama jam kerja kepala rumah tangga. Semakin besar curahan jam kerja kepala rumah tangga, maka pendapatan akan semakin tinggi. Hubungan antar faktor yang diuraikan di atas dapat dijelaskan melalui kerangka konseptual pada Gambar 2.4.

“Halaman ini sengaja dikosongkan”

BAB 3

METODE PENELITIAN

3.1. Sumber Data

Sumber data yang akan digunakan dalam penelitian ini adalah data *cross section* yang merupakan data mikro hasil pencacahan Survei Sosial Ekonomi Nasional (SUSENAS) yang dilaksanakan bulan Maret 2016 di Kabupaten Sidoarjo yang diselenggarakan oleh Badan Pusat Statistik (BPS). Unit observasinya adalah 960 rumah tangga pada 96 Blok Sensus (BS) yang tersebar di Kabupaten Sidoarjo.

Data yang digunakan adalah data hasil pencacahan dengan kuesioner SUSENAS Kor (VSEN15.K) untuk karakteristik individu dan rumah tangga serta dari Konsumsi Pengeluaran (VSEN15.KP) untuk data konsumsi rumah tangga, diantaranya adalah data tentang kuantitas dan nilai konsumsi makanan, minuman dan tembakau seminggu yang lalu, data tentang pengeluaran untuk barang-barang bukan makanan selama sebulan dan setahun terakhir (dalam rupiah), serta keterangan tentang pendapatan, penerimaan dan pengeluaran bukan konsumsi selama setahun terakhir.

3.2. Variabel Penelitian

Variabel penelitian yang digunakan beserta definisi operasionalnya adalah sebagai berikut:

1. Variabel Respon dengan notasi Y adalah pengeluaran rumah tangga untuk konsumsi susu dimana:

$y_i = 0$ jika pada rumah tangga tidak ada pengeluaran rumah tangga untuk konsumsi susu, dan

$y_i = y_i^*$ adalah pengeluaran rumah tangga untuk konsumsi susu.

Definisi pengeluaran rumah tangga untuk konsumsi susu adalah nilai susu (dalam satuan ribuan rupiah) yang benar-benar habis dikonsumsi oleh anggota rumah tangga selama seminggu terakhir.

Konsep rumah tangga yang digunakan dalam penelitian ini secara umum adalah seorang atau sekelompok orang yang mendiami sebagian atau seluruh bangunan fisik atau sensus, dan biasanya tinggal bersama serta makan dari satu dapur. Rumah tangga dalam SUSENAS terbagi menjadi dua, yaitu rumah tangga biasa dan rumah tangga khusus. Rumah tangga biasa umumnya terdiri atas ibu, bapak dan anak. Rumah tangga khusus umumnya mencakup orang-orang yang tinggal di asrama; yang tinggal di lembaga pemasyarakatan, panti asuhan, rumah tahanan, dan sejenisnya; sekelompok orang yang indekos yang berjumlah besar atau sama dengan sepuluh orang.

2. Variabel Prediktor/*Exogenous* dengan notasi X adalah sebagai berikut:

Tabel 3. 1. Variabel Prediktor/*Exogenous*

Nama Variabel	Satuan	Kategori	Skala	Notasi
Pendidikan Kepala Rumah Tangga	-	(0,0) = SD/ tidak tamat SD (1,0) = SLTP dan SMU (0,1) = Perguruan Tinggi	Ordinal	- X ₁₁ X ₁₂
Persentase Pengeluaran untuk Makanan	Persen	-	Rasio	X ₂
Jumlah Anggota Rumah Tangga	Orang	-	Interval	X ₃
Persentase Anggota Rumah Tangga Usia ≤12 Tahun	Persen	-	Rasio	X ₄
Rata-rata Pengeluaran Perkapita	Ribuan Rupiah	-	Rasio	X ₅
Persentase Anggota Rumah Tangga yang Bekerja	Persen	-	Rasio	X ₆

Definisi operasional dari masing-masing variabel prediktor tersebut adalah sebagai berikut:

- a. Pendidikan Kepala Rumah Tangga adalah pendidikan terakhir yang ditamatkan oleh kepala rumah tangga.

Kepala rumah tangga adalah salah seorang dari anggota rumah tangga yang bertanggung jawab atas pemenuhan kebutuhan sehari-hari di rumah tangga atau orang yang dianggap/ditunjuk sebagai kepala rumah tangga.

- b. Persentase pengeluaran untuk makanan adalah Jumlah pengeluaran rumah tangga untuk makanan selama sebulan dibagi dengan jumlah total pengeluaran rumah tangga sebulan dikalikan seratus persen. Data pengeluaran untuk makanan pada SUSENAS adalah data pengeluaran selama seminggu, sehingga dilakukan penyesuaian terlebih dahulu. Data dikonversi menjadi data persentase pengeluaran untuk makanan selama seminggu.

- c. Jumlah Anggota Rumah Tangga adalah jumlah total semua orang yang biasanya bertempat tinggal di suatu rumah tangga, baik yang sedang berada di rumah pada waktu pencacahan maupun yang sementara tidak berada di rumah.

- d. Persentase Anggota Rumah Tangga Usia ≤ 12 adalah jumlah anggota rumah tangga usia ≤ 12 tahun dibagi dengan jumlah total anggota rumah tangga dikalikan 100 persen.

- e. Pengeluaran Rata-rata Perkapita sebulan adalah biaya yang dikeluarkan untuk konsumsi semua anggota rumah tangga selama sebulan dibagi dengan banyaknya anggota rumah tangga. Data dikonversi menjadi pengeluaran rata-rata perkapita seminggu.

- f. Persentase Anggota Rumah Tangga yang bekerja adalah Jumlah anggota rumah tangga yang bekerja dibagi dengan jumlah total anggota rumah tangga dikalikan seratus persen.

Bekerja adalah melakukan pekerjaan dengan maksud memperoleh atau membantu memperoleh penghasilan atau keuntungan paling sedikit selama satu jam dalam seminggu yang lalu. Bekerja selama satu jam tersebut harus dilakukan berturut-turut dan tidak terputus. Kegiatan bekerja ini

mencakup baik yang sedang bekerja maupun yang punya pekerjaan tetapi dalam seminggu yang lalu sementara tidak aktif bekerja, misalnya karena sakit, cuti, menunggu panen, mogok kerja, tugas belajar dan sejenisnya.

3. Variabel *Endogenous* dengan notasi d adalah Pendapatan Rumah Tangga. Pendapatan Rumah Tangga adalah total pendapatan (dalam satuan ribuan rupiah) dari seluruh anggota rumah tangga selama sebulan yang mencakup upah/gaji/pendapatan, termasuk semua tunjangan dan bonus bagi pekerja/karyawan/pegawai dan hasil usaha berupa sewa, bunga atau keuntungan, baik berupa uang atau barang bagi pengusaha. Mengingat kendala konseptual, kendala praktis, dan ketidakersediaan data yang diperlukan, pendekatan berbasis *output* perlu dilakukan. Oleh sebab itu nilai pendapatan rumah tangga dalam penelitian ini didekati dengan pengeluaran total rumah tangga selama sebulan (dalam satuan ribuan rupiah) yang dikonversi menjadi pengeluaran total rumah tangga selama seminggu.
4. Variabel *instrumental* dengan notasi w adalah Jumlah Jam Kerja Kepala Rumah Tangga. Jumlah Jam Kerja Kepala Rumah Tangga adalah lama waktu (dalam jam) yang digunakan kepala rumah tangga untuk bekerja pada pekerjaan utama selama seminggu berturut-turut.

Struktur data yang digunakan dalam penelitian ini adalah sebagaimana Tabel 3.2.

Tabel 3. 2. Struktur Data Untuk Analisis

Sampel ke- i	y	x_{11}	x_{12}	x_2	x_3	x_4	x_5	x_6	d	w
1	y_1	x_{111}	x_{121}	x_{21}	x_{31}	x_{41}	x_{51}	x_{61}	d_1	w_1
2	y_2	x_{112}	x_{122}	x_{22}	x_{32}	x_{42}	x_{52}	x_{62}	d_2	w_2
3	y_3	x_{113}	x_{123}	x_{23}	x_{33}	x_{43}	x_{53}	x_{63}	d_3	w_3
...
n	y_n	x_{11n}	x_{12n}	x_{2n}	x_{3n}	x_{4n}	x_{5n}	x_{6n}	d_n	w_n

Adapun model TKB *Endogenous* yang akan dibentuk pada masing-masing kuantil ke-p adalah sebagai berikut:

$$\hat{y}_p = \widehat{const}_p + \hat{\beta}_{p,11}x_{11} + \hat{\beta}_{p,12}x_{12} + \hat{\beta}_{p,2}x_2 - \hat{\beta}_{p,3}x_3 + \hat{\beta}_{p,4}x_4 - \hat{\beta}_{p,5}x_5 + \hat{\beta}_{p,6}x_6 + \hat{\delta}_p d \quad , y^* > 0$$

$$= 0 \quad , y^* \leq 0$$

3.3. Metode Penelitian

Pemodelan data pengeluaran rumah tangga untuk konsumsi susu di Kabupaten Sidoarjo dilakukan langkah sebagai berikut:

Langkah 1: Bentuk model TK dan TKB untuk data pengeluaran rumah tangga untuk konsumsi susu di Kabupaten Sidoarjo dengan langkah sebagai berikut:

- a. Lakukan analisis deskriptif atas seluruh variabel yang digunakan dalam penelitian.
- b. Bentuk model TK untuk setiap kuantil dengan mendapatkan estimasi parameter model TK.
- c. Bentuk model TKB untuk setiap kuantil dengan mendapatkan estimasi parameter model TKB.
- d. Hitung RMSE untuk model TK dan TKB setiap kuantil.

Langkah 2: Tentukan distribusi *prior* beserta nilai masing-masing parameter dengan melakukan kajian atas model TKB *Endogenous*.

Langkah 3: Bentuk model TKB *Endogenous* terhadap data pengeluaran rumah tangga untuk konsumsi susu di Kabupaten Sidoarjo dengan langkah sebagai berikut:

- a. Lakukan analisis deskriptif terhadap seluruh variabel yang digunakan dalam penelitian, baik variabel respon, variabel *exogenous*/prediktor, variabel *endogenous*, maupun variabel *instrumental*.
- b. Dapatkan estimasi parameter model TKB *endogenous* dengan algoritma *Gibbs Sampler* untuk ALDP.

1. *Sampling* $\{u_i\}_{i=1}^n$: Bangkitkan u_i dari $\mathcal{U}(0, \pi_{k_i})$ untuk $i=1, \dots, n$.

2. *Sampling* $\{\omega_l\}_{l=1}^{k^*}$: Bangkitkan ω_l dari $\mathcal{B}(1+n_l, n-\sum_{r \leq l} n_r + a)$ dimana $n_l = \sum_{i=1}^n I(k_i = l)$ untuk $l=1, \dots, k^*$.

3. *Sampling* $\{k_i\}_{i=1}^{k^*}$: Bangkitkan k_i dari distribusi multinomial dengan probabilitas:

$$\Pr(k_i = l) \propto f_{AL}(d_i - \mathbf{z}'_i \boldsymbol{\gamma} | \phi_l, \alpha) I(u_i < \pi_l), l=1, \dots, k^*$$

untuk $i=1, \dots, n$.

4. *Sampling* $\{\phi_l\}_{l=1}^{k^*}$: Bangkitkan ϕ_l dari $\mathcal{JG}(c_l, d_l)$ dimana:

$$c_l = 1.5n_l + c_0$$

$$d_l = \sum_{i:k_i=l} \left[h_i + \frac{(d_i - \mathbf{z}'_i \boldsymbol{\gamma} - \theta_\alpha h_i)^2}{2\tau_\alpha^2 h_i} \right] + d_0$$

5. *Sampling a* : Dengan prior gamma, $\mathcal{G}(a_0, b_0)$ dan $c \sim \mathcal{B}(a+1, n)$, maka *full conditional distribution* dari a adalah *mixture* dari dua distribusi gamma yaitu:

$$\varphi \mathcal{G}(a_0 + n^*, b_0 - \log c) + (1 - \varphi) \mathcal{G}(a_0 + n^* - 1, b_0 - \log c)$$

dimana n^* adalah jumlah dari cluster yang berbeda dan $\frac{\varphi}{(1-\varphi)} =$

$$\frac{(a_0 + n^* - 1)}{(n(b_0 - \log c))}$$

6. *Sampling $\boldsymbol{\gamma}$* : Dengan $\boldsymbol{\gamma} \sim N(\mathbf{g}_0, \mathbf{G}_0)$, maka $\boldsymbol{\gamma}$ *disampling* dari $N(\mathbf{g}_1, \mathbf{G}_1)$ dimana

$$\mathbf{G}_1 = \left[\sum_{i=1}^n \mathbf{z}_i \left(\frac{\eta_p^2}{\tau_p^2 \sigma_{g_i}} + \frac{1}{\tau_\alpha^2 \phi_{k_i} h_i} \right) \mathbf{z}'_i + \mathbf{G}_0^{-1} \right]^{-1}$$

$$\mathbf{g}_1 = \mathbf{G}_1 \left[\sum_{i=1}^n \mathbf{z}_i \left(-\frac{\eta_p (y_i^* - \mathbf{x}'_i \boldsymbol{\beta}_p - \eta_p d_i - \theta_p g_i)}{\tau_p^2 \sigma_{g_i}} + \frac{d_i - \theta_\alpha h_i}{\tau_\alpha^2 \phi_{k_i} h_i} \right) + \mathbf{G}_0^{-1} \mathbf{g}_0 \right]$$

dengan densitas dari *full conditional distribution* dengan notasi $\pi(\boldsymbol{\gamma} | -)$ adalah sebagai berikut:

$$\begin{aligned} \pi(\boldsymbol{\gamma} | -) &\propto \exp \left\{ - \sum_{i=1}^n \frac{(y_i^* - \mathbf{x}'_i \boldsymbol{\beta}_p - \delta_p d_i - \eta_p (d_i - \mathbf{z}'_i \boldsymbol{\gamma}) - \theta_p g_i)^2}{2\tau_p^2 \sigma_{g_i}} \right\} \\ &\quad \times \exp \left\{ - \sum_{i=1}^n \frac{(d_i - \mathbf{z}'_i \boldsymbol{\gamma})^2}{2\tau_\alpha^2 \phi_{k_i} h_i} \right\} \exp \left\{ - \frac{1}{2} (\boldsymbol{\gamma} - \mathbf{g}_0)' \mathbf{G}_0^{-1} (\boldsymbol{\gamma} - \mathbf{g}_0) \right\} \\ &\propto \exp \left\{ - \frac{1}{2} (\boldsymbol{\gamma} - \mathbf{g}_1)' \mathbf{G}_1^{-1} (\boldsymbol{\gamma} - \mathbf{g}_1) \right\} \end{aligned}$$

7. *Sampling* $\{h_i\}_{i=1}^n$: *Full conditional distribution* untuk h_i adalah *Generalised Inverse Gaussian Distribution* yang dinotasikan dengan $\mathcal{GIG}(v, \xi, \chi)$ dengan pdf (*Probability Density Function*) sebagai berikut:

$$f(x|v, \xi, \chi) = \frac{\left(\frac{x}{\xi}\right)^v}{2K_v(\xi\chi)} x^{v-1} \exp\left\{-\frac{1}{2}(\xi^2 x^{-1} + \chi^2 x)\right\},$$

$$x > 0, -\infty < v < \infty, \xi, \chi \geq 0$$

dimana $K_v(\cdot)$ adalah fungsi Bessel jenis ke-3 yang dimodifikasi. Untuk

$i=1, \dots, n$, maka h_i *disampling* dari $\mathcal{GIG}(1/2, \xi_i, \chi_i)$ dengan $\xi_i^2 = \frac{(d_i - \mathbf{z}'_i \boldsymbol{\gamma})^2}{\tau_\alpha^2 \phi_{k_i}}$

$$\text{dan } \chi_i^2 = \frac{\theta_\alpha^2}{\tau_\alpha^2 \phi_{k_i}} + \frac{2}{\phi_{k_i}}.$$

8. *Sampling* α : Densitas dari *full conditional distribution* dari α adalah $\pi(\alpha | -) \propto \pi(\alpha) \prod_{i=1}^n f_{AL}(d_i - \mathbf{z}'_i \boldsymbol{\gamma} | \phi_{k_i}, \alpha)$ dimana $\pi(\alpha)$ adalah *prior density* dari α . Digunakan algoritma *random walk Metropolis-Hastings* (MH) untuk melakukan *sampling* dari distribusi ini.

9. *Sampling* $\{y_i^*\}_{i=1}^n$: *Full conditional distribution* dari y_i^* adalah sebagai berikut:

$$y_i I(y_i > 0) + \mathcal{TN}_{(-\infty, 0)}(\tilde{\mathbf{x}}_i' \tilde{\boldsymbol{\beta}}_p + \theta_p g_i, \tau_p^2 \sigma g_i) I(y_i = 0), \quad i = 1, \dots, n.$$

10. *Sampling* $\tilde{\boldsymbol{\beta}}_p$: Lakukan *sampling* $\tilde{\boldsymbol{\beta}}_p = (\boldsymbol{\beta}'_p, \delta_p, \eta_p)'$ sekaligus. Dengan asumsi $\tilde{\boldsymbol{\beta}}_p \sim N(\tilde{\mathbf{b}}_0, \tilde{\mathbf{B}}_0)$, maka *full conditional distribution*nya mengikuti $N(\tilde{\mathbf{b}}_1, \tilde{\mathbf{B}}_1)$ dimana

$$\tilde{\mathbf{B}}_1 = \left[\sum_{i=1}^n \frac{\tilde{\mathbf{x}}_i \tilde{\mathbf{x}}_i'}{\tau_p^2 \sigma g_i} + \tilde{\mathbf{B}}_0^{-1} \right]^{-1} \text{ dan } \tilde{\mathbf{b}}_1 = \tilde{\mathbf{B}}_1 \left[\sum_{i=1}^n \frac{\tilde{\mathbf{x}}_i (y_i^* - \theta_p g_i)}{\tau_p^2 \sigma g_i} + \tilde{\mathbf{B}}_0^{-1} \tilde{\mathbf{b}}_0 \right].$$

11. *Sampling* σ : Dengan $\sigma \sim \mathcal{IG}(m_0, s_0)$, maka lakukan *sampling* terhadap σ dari $\mathcal{IG}(m_1, s_1)$ dimana:

$$m_1 = 1.5n + m_0 \text{ dan } s_1 = \sum_{i=1}^n g_i + \sum_{i=1}^n \frac{(y_i - \tilde{\mathbf{x}}_i' \tilde{\boldsymbol{\beta}}_p - \theta_p g_i)^2}{2\tau_p^2 g_i} + s_0$$

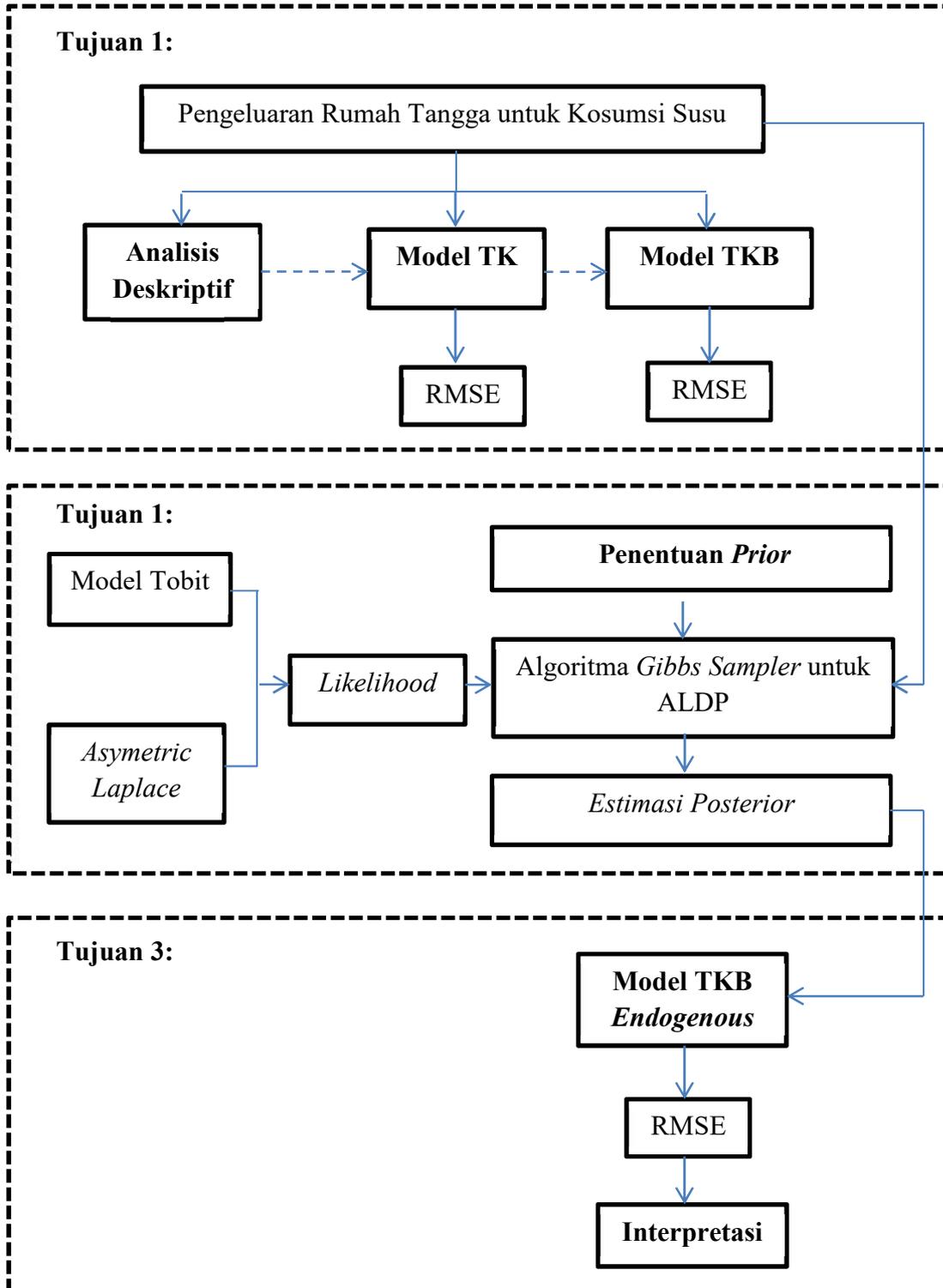
12. *Sampling* $\{g_i\}_{i=1}^n$: *Sampling* g_i sebagaimana melakukan *sampling* terhadap

$$h_i, \text{ yaitu dari } \mathcal{GIG}(1/2, \lambda_i, \psi) \text{ dimana } \lambda_i^2 = \frac{(y_i^* - \tilde{\mathbf{x}}_i' \tilde{\boldsymbol{\beta}}_p)^2}{\tau_p^2 \sigma} \text{ dan } \psi^2 = \frac{\theta_p^2}{\tau_p^2 \sigma} + \frac{2}{\sigma},$$

$i=1, \dots, n$.

- c. Interpretasikan hasil dan hitung RMSE untuk model TKB *Endogenous* di setiap kuantil.

Adapun diagram alir penelitian adalah sebagaimana Gambar 3.1.



Gambar 3. 1. Diagram Alir Penelitian

BAB 4

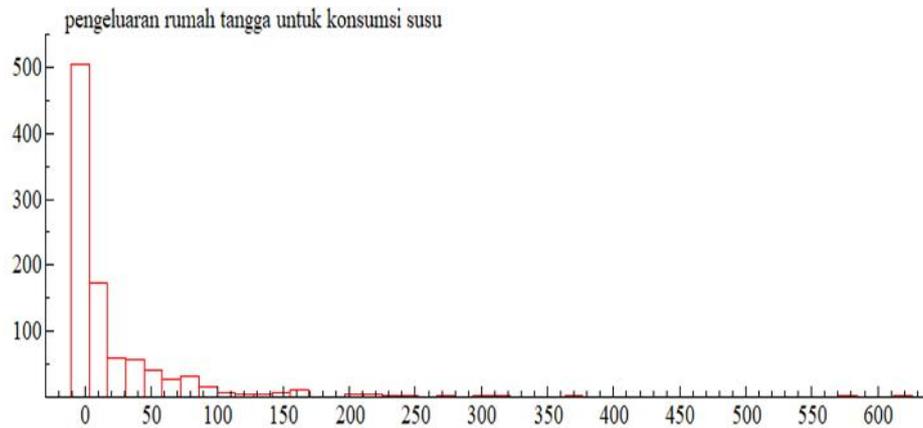
HASIL DAN PEMBAHASAN

Pada bab ini akan dibahas karakteristik data pengeluaran rumah tangga untuk konsumsi susu di Kabupaten Sidoarjo beserta faktor-faktor yang mempengaruhinya sebagai deskripsi awal. Selanjutnya disajikan hasil pemodelan data dengan model TK, TKB, dan TKB *Endogenous* beserta nilai RMSE masing-masing model untuk kemudian dilakukan penarikan kesimpulan mengenai model terbaik yang dapat digunakan pada data yang mengandung masalah endogenitas.

4.1. Model TK dan TKB untuk Pengeluaran Rumah Tangga untuk Konsumsi Susu di Kabupaten Sidoarjo

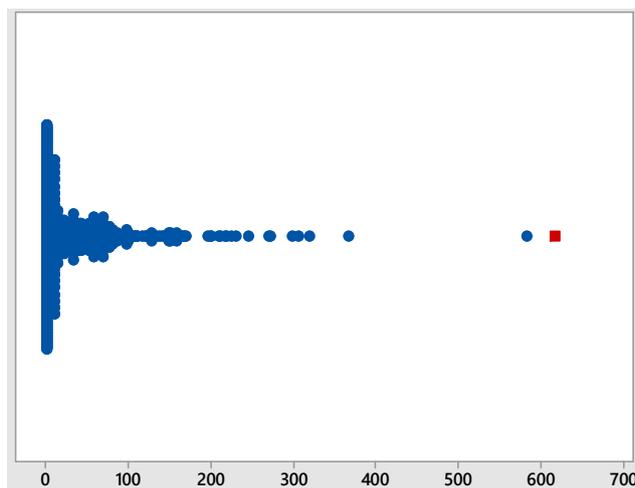
Penelitian ini melibatkan 4 (empat) jenis variabel, yaitu variabel respon, variabel *exogenous*, variabel *endogenous*, dan variabel *instrumental*. Variabel respon dari data yang digunakan, yaitu data pengeluaran rumah tangga untuk konsumsi susu, merupakan data *mixture* antara data diskret dan kontinyu. Oleh karena itu, data variabel respon tersebut dapat dikategorikan sebagai data tersensor. Dari histogram sebagaimana Gambar 4.2, terlihat bahwa data tersensor di sebelah kiri (*left censored*) yaitu pada data $y < 0$. Regresi yang dapat digunakan untuk memodelkan variabel respon dengan tipe sebagaimana uraian di atas adalah regresi tobit.

Censoring rate pada model adalah sebesar 0,52. Artinya sebanyak 52 persen data bernilai nol dan 48 persen data memiliki nilai yang bervariasi dengan nilai maksimum pengeluaran rumah tangga untuk konsumsi susu dalam kurun waktu seminggu adalah Rp. 617.000,-. Hal ini didukung oleh histogram data pada Gambar 4.1 dimana terlihat bahwa frekuensi tertinggi berada pada nilai nol. Hal ini menyebabkan plot densitas akan sangat runcing pada nilai tersebut sekaligus memberikan informasi awal bahwa *error* dari regresi akan berdistribusi *asymetric laplace*.



Gambar 4. 1. Histogram Data Pengeluaran Rumah Tangga untuk Konsumsi Susu di Kabupaten Sidoarjo

Dari histogram tersebut juga terlihat bahwa terdapat ekor yang panjang di sebelah kanan yang memberikan indikasi awal akan adanya data *outlier*. *Outlier* adalah pengamatan yang nilai sisaannya jauh lebih besar daripada sisaan pengamatan lain. *Outlier* ini terletak tiga atau empat simpangan baku, atau lebih jauh lagi dari *mean* residualnya. Keberadaan *outlier* dapat mengubah atau mengaburkan kesimpulan yang dibuat oleh peneliti karena estimator parameternya akan bersifat bias. Deteksi lanjutan terhadap variabel respon atas adanya data *outlier*, secara sederhana dapat ditunjukkan melalui plot sebagaimana Gambar 4.2.



Gambar 4. 2. *Outlier* pada Variabel Respon

Selanjutnya, melalui uji *Studentized Deleted Residual* (TRES) dengan hipotesis sebagai berikut:

H_0 : pengamatan ke- i bukan merupakan *outlier*

H_1 : pengamatan ke-i merupakan *outlier*

data menunjukkan bahwa bahwa terdapat 44 pengamatan yang merupakan data *outlier*.

Selanjutnya, untuk melihat apakah observasi tersebut berpengaruh atau tidak, dilakukan uji *The Difference in Fits Statistics* (DFITS) dengan hipotesis:

H_0 : pengamatan ke-i bukan merupakan *outlier* berpengaruh

H_1 : pengamatan ke-i merupakan *outlier* berpengaruh

diperoleh hasil bahwa pengamatan yang ditemukan sebagai *outlier* tersebut merupakan pengamatan yang berpengaruh.

Dengan adanya data *outlier* yang berpengaruh, maka asumsi *error* normal yang diperlukan pada regresi tobit standar tidak dapat terpenuhi. Akibat dari pelanggaran asumsi tersebut adalah hasil estimasi yang dihasilkan akan bias. Dengan demikian diperlukan model kuantil yang bersifat tegar (*robust*) terhadap adanya data *outlier* yang dapat menyajikan analisa data pada kuantil-kuantil yang diperlukan.

4.1.1. Karakteristik data Pengeluaran Rumah Tangga untuk Konsumsi Susu di Kabupaten Sidoarjo

Berdasarkan Tabel 4.1. pengeluaran total rumah tangga dalam waktu seminggu, sebagai pendekatan dari pendapatan rumah tangga, memiliki nilai *mean* sebesar Rp. 1.229.810,- dengan nilai minimum Rp. 73.990,- dan nilai maksimum Rp. 11.268.160,-.

Tabel 4. 1. Statistik Deskriptif Variabel Penelitian

Variabel	Mean	Std	Min	5%	50%	95%	Max
y	23,943	51,554	0	0	0	159,025	617
x2	47,870	13,739	7,520	17,458	49,370	71,317	84,180
x3	3,659	1,359	1	1	4	6	11
x4	8,474	14,692	0	0	0	50	75
x5	359,983	346,456	53,700	83,537	269,985	1182,481	6479,190
x6	27,403	20,768	0	0	25	100	100
d	1229,810	1119,032	73,990	240,346	946,655	3940,186	11268,160

Pada rumah tangga sampel di Kabupaten Sidoarjo, nilai persentase pengeluaran rumah tangga untuk makanan terbesar adalah 84,18 persen. Sedangkan *mean*-nya adalah 47,87 persen dengan standar deviasi sebesar 13,739. Hal ini menunjukkan bahwa pada rumah tangga sampel secara rata-rata rumah tangganya sudah memiliki pengeluaran yang nyaris seimbang antara pengeluaran makanan dengan pengeluaran non-makanan. Dapat dikatakan bahwa kesejahteraan masyarakat di Kabupaten Sidoarjo berdasarkan data sampel sudah cukup baik karena konsumsi untuk makanan sudah tidak terlalu menjadi fokus perhatian masyarakat.

Jumlah ART pada rumah tangga sampel secara rata-rata adalah 4 (empat) orang, sesuai dengan konsep rumah tangga yang ideal. Dari Tabel 4.1 di atas dapat juga dijelaskan bahwa tidak banyak rumah tangga yang memiliki anggota rumah tangga besar atau lebih dari 6 orang karena data jumlah ART pada kuantil 0,95 bernilai 6,00 walaupun nilai *maximum* data adalah 11,00.

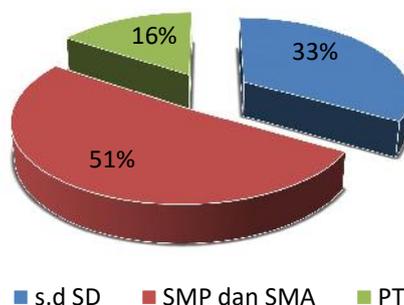
Terdapat 71,56 persen rumah tangga sampel di Kabupaten Sidoarjo dimana di dalam rumah tangga tersebut tidak terdapat anggota rumah tangga berumur ≤ 12 tahun. Dari rumah tangga sampel yang tidak memiliki ART ≤ 12 tahun tersebut, 39,01 persennya mengalokasikan pengeluaran rumah tangganya untuk konsumsi susu. Sedangkan pada rumah tangga sampel yang mempunyai ART ≤ 12 tahun, 71,43 persennya mengalokasikan pengeluaran rumah tangganya untuk konsumsi susu. Hal ini mengindikasikan bahwa dengan adanya ART ≤ 12 tahun, maka rumah tangga akan cenderung memiliki pengeluaran untuk konsumsi susu. Dari analisa tersebut, dapat dikatakan juga bahwa kesadaran masyarakat untuk memberikan susu sebagai penunjang gizi bagi tumbuh kembang anak dengan kriteria usia ≤ 12 tahun di Kabupaten Sidoarjo cenderung sudah baik.

Mean pengeluaran perkapita rumah tangga sampel selama seminggu bernilai Rp. 359.980,- dengan standar deviasi 346,456. Tidak banyak rumah tangga yang pengeluaran perkapitanya bernilai jauh lebih tinggi dari sebagian besar rumah tangga lainnya. Hal ini dapat dilihat dari nilai pengeluaran perkapita pada kuantil 0,95 adalah Rp. 1.182.481,- walaupun nilai *maximum* data adalah sebesar Rp. 6.479.190,-.

Pada rumah tangga dimana ada anggota rumah tangga bekerja, 50,30 persennya mengalokasikan pengeluaran rumah tangga untuk konsumsi susu. Hal ini memberikan indikasi bahwa persentase anggota rumah tangga yang bekerja akan tidak berpengaruh secara signifikan dalam pengambilan keputusan rumah tangga untuk mengkonsumsi susu.

Jumlah jam kerja KRT selama seminggu terakhir yang dalam hal ini adalah variabel *instrumental* pada data sampel di Kabupaten Sidoarjo memiliki nilai mean sebesar 37,967. Di samping itu, diperoleh informasi bahwa lima persen KRT di Kabupaten Sidoarjo memiliki jumlah jam kerja yang lebih dari 82 jam selama seminggu.

Dari Gambar 4.3 terlihat bahwa pendidikan yang ditamatkan KRT sampel di Kabupaten Sidoarjo didominasi di kelompok pendidikan menengah (SMP dan SMA). Selanjutnya, berdasarkan pendidikan Kepala Rumah Tangga (KRT) sampel di Sidoarjo, terdapat 31,25 persen rumah tangga dari pendidikan KRT sampai dengan SD, 54,08 persen rumah tangga dari pendidikan KRT Menengah (SMP dan SMA), dan 65,33 persen rumah tangga dari pendidikan KRT yang menamatkan Perguruan Tinggi yang mengalokasikan pengeluaran untuk konsumsi susu. Hal ini menjelaskan teori bahwa semakin tinggi pendidikan KRT di rumah tangga, maka kesadaran akan konsumsi susu untuk anggota rumah tangganya semakin besar. Dengan demikian hubungan pendidikan KRT dengan besarnya konsumsi rumah tangga untuk konsumsi susu di dalam rumah tangga tersebut adalah kuat dan positif.



Gambar 4. 3. Pendidikan KRT di rumah tangga sampel di Kabupaten Sidoarjo

4.1.2. Model TK pada Pengeluaran Rumah Tangga untuk Konsumsi Susu di Kabupaten Sidoarjo

Regresi Kuantil adalah regresi yang dapat menjelaskan hubungan antar variabel respon dan variabel prediktor pada berbagai titik kuantil dari distribusi variabel respon tersebut. Dengan demikian, pembagian data ke dalam kuantil-kuantil dilakukan berdasarkan distribusi variabel respon.

Hal ini menjadikan regresi ini dapat digunakan pada kondisi data yang heterogen, berbeda dengan regresi linier berganda yang hanya mampu menjelaskan hubungan sebab akibat pada *mean (conditional mean)* variabel respon. Hal ini memberikan jaminan tidak adanya informasi rumah tangga yang terbuang karena data yang jauh dari *mean* akan dapat dimodelkan tersendiri.

Model TK untuk pengeluaran rumah tangga untuk konsumsi susu di Kabupaten Sidoarjo yang dibentuk berdasarkan hasil estimasi parameternya adalah sebagai berikut:

$$\begin{aligned}
 \hat{y}_{0,50} &= - 15,879 + 3,278 x_{11} + 7,624 x_{12} + 0,238 x_2 - 0,794 x_3 + 0,300 x_4 \\
 &\quad - 0,007 x_5 - 0,072 x_6 + 0,012 d && , y^* > 0 \\
 &= 0 && , y^* \leq 0 \\
 \hat{y}_{0,75} &= - 50,909 + 18,790 x_{11} + 40,478 x_{12} + 0,781 x_2 - 1,461 x_3 + 0,849 x_4 \\
 &\quad - 0,013 x_5 - 0,360 x_6 + 0,033 d && , y^* > 0 \\
 &= 0 && , y^* \leq 0 \\
 \hat{y}_{0,95} &= - 88,709 + 45,632 x_{11} + 102,306 x_{12} + 1,654 x_2 + 4,072 x_3 + 1,129 x_4 \\
 &\quad + 0,017 x_5 - 1,167 x_6 + 0,057 d && , y^* > 0 \\
 &= 0 && , y^* \leq 0 \\
 \hat{y}_{0,975} &= - 96,061 + 60,167 x_{11} + 104,582 x_{12} + 1,864 x_2 + 3,900 x_3 + 1,375 x_4 \\
 &\quad + 0,028 x_5 - 1,192 x_6 + 0,056 d && , y^* > 0 \\
 &= 0 && , y^* \leq 0 \\
 \hat{y}_{0,99} &= - 226,476 + 39,361 x_{11} + 63,100 x_{12} + 4,005 x_2 + 8,418 x_3 + 0,500 x_4 \\
 &\quad + 0,014 x_5 - 0,615 x_6 + 0,115 d && , y^* > 0 \\
 &= 0 && , y^* \leq 0
 \end{aligned}$$

Nilai RMSE untuk model pada kuantil 0,5, 0,75, 0,95, 0,975, dan 0,99 secara berturut-turut adalah 50,06, 45,03, 84,58, 92,58, dan 138,33. Model TK

yang merupakan regresi dalam frekuentis tidak dapat mengestimasi parameter pada kuantil-kuantil bawah, yaitu pada kuantil 0,025, 0,05, dan 0,25. Oleh karena itu, pendekatan bayesian diperlukan untuk dapat memodelkan data pada kuantil bawah.

4.1.3. Model TKB pada Pengeluaran Rumah Tangga untuk Konsumsi Susu di Kabupaten Sidoarjo

Pemodelan data pengeluaran rumah tangga untuk konsumsi susu di Kabupaten Sidoarjo dengan model TKB dilakukan dengan iterasi sebanyak 30.000 kali dan *burn in* 10.000. Hasil pemodelan adalah sebagai berikut:

$$\begin{aligned}
 \hat{y}_{0,025} &= 52,911 - 100,504 x_{11} - 126,516 x_{12} + 7,475 x_2 - 51,730 x_3 \\
 &\quad - 82,155 x_4 - 59,472 x_5 - 74,627 x_6 - 183,361 d \quad , y^* > 0 \\
 &= 0 \quad , y^* \leq 0 \\
 \hat{y}_{0,05} &= - 60,220 - 218,68 x_{11} + 47,020 x_{12} - 10,530 x_2 - 63,45 x_3 \\
 &\quad + 49,630 x_4 - 46,480 x_5 - 10,210 x_6 - 218,680 d \quad , y^* > 0 \\
 &= 0 \quad , y^* \leq 0 \\
 \hat{y}_{0,25} &= 149,050 + 143,630 x_{11} + 145,830 x_{12} - 73,580 x_2 - 146,040 x_3 \\
 &\quad - 193,020 x_4 - 75,090 x_5 - 41,660 x_6 - 47,21 d \quad , y^* > 0 \\
 &= 0 \quad , y^* \leq 0 \\
 \hat{y}_{0,50} &= - 130,109 - 116,137 x_{11} + 110,448 x_{12} + 22,226 x_2 + 179,197 x_3 \\
 &\quad - 12,806 x_4 + 9,232 x_5 - 79,332 x_6 - 87,817 d \quad , y^* > 0 \\
 &= 0 \quad , y^* \leq 0 \\
 \hat{y}_{0,75} &= 5,545 - 62,447 x_{11} - 34,908 x_{12} - 27,852 x_2 - 17,642 x_3 \\
 &\quad - 44,372 x_4 - 66,135 x_5 - 89,183 x_6 - 9,035 d \quad , y^* > 0 \\
 &= 0 \quad , y^* \leq 0 \\
 \hat{y}_{0,95} &= - 9,473 + 4,768 x_{11} + 2,983 x_{12} + 0,275 x_2 + 0,989 x_3 + 1,206 x_4 \\
 &\quad - 0,123 x_5 + 0,044 x_6 + 0,112 d \quad , y^* > 0 \\
 &= 0 \quad , y^* \leq 0
 \end{aligned}$$

$$\begin{aligned}
\hat{Y}_{0,975} &= 4,533 - 4,133 x_{11} - 2,179 x_{12} + 0,567 x_2 - 0,225 x_3 + 0,936 x_4 \\
&\quad - 0,065 x_5 - 0,215 x_6 + 0,102 d && , y^* > 0 \\
&= 0 && , y^* \leq 0 \\
\hat{Y}_{0,99} &= - 5,544 - 2,829 x_{11} - 0,225 x_{12} + 0,833 x_2 + 3,991 x_3 + 1,273 x_4 \\
&\quad + 0,041 x_5 - 0,434 x_6 + 0,102 d && , y^* > 0 \\
&= 0 && , y^* \leq 0
\end{aligned}$$

Hasil estimasi model TKB menunjukkan adanya perbedaan tanda parameter pada kuantil yang berbeda. Hal ini menyebabkan sulitnya penarikan kesimpulan mengenai hubungan variabel prediktor dengan variabel responnya. Berdasarkan nilai RMSE pada Tabel 4.2, model TKB dapat memodelkan data pengeluaran rumah tangga untuk konsumsi susu di Kabupaten Sidoarjo dengan baik hanya pada kuantil-kuantil atas, namun tidak pada kuantil-kuantil bawah sampai dengan tengah.

Tabel 4. 2. Nilai RMSE Model TKB pada Kuantil ke-p

Kuantil ke-p	TKB
0,025	268266,74
0,050	311393,72
0,250	92313,45
0,500	115742,81
0,750	36882,93
0,950	101,93
0,975	108,73
0,990	150,99

4.2. Model TKB *Endogenous* pada Pengeluaran Rumah Tangga untuk Konsumsi Susu di Kabupaten Sidoarjo

Pada pemodelan data pengeluaran rumah tangga untuk konsumsi susu di beberapa penelitian sebelumnya, seringkali variabel pendapatan rumah tangga tidak dapat digunakan dalam analisis dan interpretasi modelnya. Hal ini memberikan dugaan bahwa variabel pendapatan rumah tangga bersifat

endogenous. Jika terdapat masalah endogenitas, maka model TKB standar sebagaimana ulasan pada Subbab sebelumnya tidak dapat digunakan dalam analisa data. Hal ini dikarenakan estimasi parameter nya akan bersifat bias, sebagaimana dibahas pada Subbab 2.5.2.

4.2.1. Uji Endogenitas

Untuk memperoleh informasi awal mengenai sifat endogenitas pada variabel yang diduga merupakan fungsi dari satu set variabel *exogenous* dan *instrumental*, maka dilakukan uji eksogenitas pada regresi tobit. Hipotesis untuk pengujian tersebut adalah:

$H_0: \text{corr}(u_i, v_i) = 0$ atau tidak terdapat variabel *endogenous*.

$H_1: \text{corr}(u_i, v_i) \neq 0$ atau terdapat variabel *endogenous* yang merupakan fungsi dari satu set variabel *exogenous* dan *instrumental*.

Data sampel rumah tangga di Kabupaten Sidoarjo menunjukkan bahwa hipotesis *null* sebagaimana dirumuskan diatas ditolak, yang berarti variabel *endogenous* merupakan fungsi dari satu set variabel *exogenous* dan *instrumental*. Dengan demikian, terdapat masalah endogenitas dalam regresi tobit. Hal ini menjadi dasar untuk melanjutkan analisis dengan memperhatikan pengaruh endogenitas suatu variabel terhadap variabel lainnya.

Variabel pengeluaran total rumah tangga sebagai pendekatan untuk pendapatan rumah tangga di Kabupaten Sidoarjo merupakan variabel yang bersifat *endogenous* yang juga dipengaruhi oleh variabel *exogenous*, yaitu pendidikan KRT, persentase pengeluaran untuk makanan, jumlah ART, persentase ART usia ≤ 12 tahun, rata-rata pengeluaran perkapita, persentase ART yang bekerja, dan variabel *instrumental* yaitu jumlah jam kerja KRT.

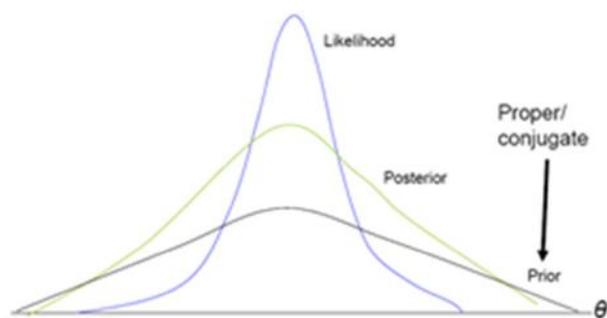
Dengan adanya informasi bahwa terdapat masalah endogenitas melalui hasil pengujian ini dan berdasarkan tinjauan pustaka pada bab sebelumnya, maka dapat disimpulkan bahwa hasil estimasi parameter akan menjadi bias jika pemodelan data menggunakan model TKB standar. Selanjutnya, perlu dilakukan pemodelan yang dapat mengakomodir masalah endogenitas tersebut, yaitu model TKB *Endogenous*.

4.2.2. Kajian Distribusi Prior Model TKB *Endogenous*

Hal yang perlu diperhatikan dalam penggunaan model TKB *Endogenous* adalah penentuan prior yang sesuai. Sebagaimana dijelaskan pada bab sebelumnya bahwa kegunaan dari *prior* adalah sebagai bentuk distribusi frekuensi, sebagai bentuk representasi normatif dan objektif pada suatu parameter yang lebih rasional untuk dipercayai, dan sebagai suatu representasi subjektifitas peneliti dalam memandang sebuah parameter. Permasalahan pokok agar *prior* dapat interpretatif adalah bagaimana memilih distribusi *prior* untuk suatu parameter yang tidak diketahui namun sesuai dengan permasalahan fisik yang ada. Jika parameter hanya mempunyai nilai-nilai pada *range* tertentu saja, hal ini disebabkan oleh penggunaan *prior* yang mempunyai densitas yang *equally likelly*. Artinya, setiap kondisi diberi kesempatan yang sama untuk dapat terpilih sebagai pendukung *likelihood* dalam membentuk *posterior*nya. Dengan demikian, *prior* akan memiliki arti yang janggal jika salah dalam pemilihannya. Selain itu, pemilihan *prior* yang tidak tepat akan menyebabkan *misleading*, yaitu *posterior* akan menjadi tidak jelas.

Berdasarkan uraian di atas, maka *prior* yang digunakan dalam penelitian ini adalah sebagai berikut:

1. γ yang merupakan koefisien regresi pada tahap pertama diasumsikan memiliki *prior* normal karena sesuai untuk AL maupun ALDP secara perhitungannya. Karena tidak ada informasi tentang nilai koefisien, maka *varians* disesuaikan menjadi $\gamma \sim N(\mathbf{0}, 500\mathbf{I})$ dimana nilai tersebut diperoleh dengan cara *trial and error* dimulai dari nilai 100, 200, 300, 400, 500, 600, 700, dan 800, dengan memperhatikan sigifikansi hasil estimasi parameter koefisien regresi dan kesesuaiannya dengan teori. Gambar 4.4 merupakan ilustrasi yang dapat menggambarkan hubungan *prior*, *likelihood*, dan *posterior* sehingga dapat dilakukan penetapan nilai *varians* pada *prior* sesuai bagi *likelihood*.



Gambar 4. 4. Ilustrasi *Prior, Likelihood, dan Posterior*

2. *Scale parameter* σ untuk distribusi AL pada tahap kedua diasumsikan berdistribusi *inverse gamma* sehingga $\sigma \sim \mathcal{IG}(0.1, 0.1)$.
3. *Scale parameter* ϕ untuk distribusi ALDP pada tahap pertama diasumsikan berdistribusi *inverse gamma* sehingga $\phi \sim \mathcal{IG}(2, 0.5)$.
4. *Prior* parameter presisi a dari *Dirichlet Process* adalah $a \sim \mathcal{G}(2, 2)$.
5. *Prior* untuk parameter koefisien pada tahap kedua, β_p dan δ_p diasumsikan berdistribusi Normal dimana $(\beta_p', \delta_p)' \sim N(\mathbf{0}, 500\mathbf{I})$. Penentuan nilai varians dilakukan dengan cara yang sama sebagaimana penentuan *prior* untuk koefisien pada regresi tahap pertama. Untuk data pada penelitian ini ditetapkan bahwa *prior* yang digunakan adalah $N(\mathbf{0}, 500\mathbf{I})$ karena pada varians tersebut, tanda dari parameter telah sesuai dengan teori dan banyaknya signifikansi variabel prediktor.
6. *Prior* untuk koefisien *error* tahap pertama yang melibatkan pada persamaan tahap kedua, yang merupakan indikator endogenitas, yaitu $\eta_p \sim N(0, 5)$.
7. Untuk ALDP, diasumsikan $\alpha \sim \mathcal{U}(0, 1)$.

4.3. Interpretasi Model TKB *Endogenous* untuk Pengeluaran Rumah Tangga untuk Konsumsi Susu di Kabupaten Sidoarjo

Pemodelan data dengan model TKB *Endogenous* dilakukan dengan penentuan distribusi dan nilai parameter pada *prior* sebagaimana pembahasan sebelumnya. Model dibentuk dengan melakukan iterasi dan membuang data hasil *sampling* di awal (*burn in*) untuk mendapatkan model yang konvergen. Model yang telah konvergen dapat dilihat dari *trace plot* masing-masing parameter setelah periode *burn in*.

Berdasarkan *prior* seperti pada pembahasan sebelumnya, dengan melakukan iterasi sebanyak 10.000 kali dan *burn in* 5.000 data awal, diperoleh hasil bahwa masalah endogenitas muncul pada kuantil di sebelah kiri kuantil ke-0,55 dan kuantil di sebelah kanan kuantil ke-0,80. Endogenitas disimpulkan dengan memperhatikan signifikansi nilai parameter η . Signifikansi η ditentukan dengan kriteria *credible interval* (CI) dimana η dikatakan signifikan jika 95% CI untuk parameter η tidak memuat nilai nol. Pada kuantil $0,55 < p < 0,80$ endogenitas variabel pengeluaran total rumah tangga pada data rumah tangga di Kabupaten Sidoarjo tidak signifikan karena 95% CI untuk parameter η pada kuantil-kuantil pada *range* tersebut memuat nilai nol.

Untuk mendapatkan estimasi dan model yang lebih baik, maka dengan *setting prior* seperti pada pembahasan sebelumnya, menambah iterasi menjadi sebanyak 30.000 kali dan *burn in* 10.000 data awal, diperoleh hasil bahwa masalah endogenitas muncul pada semua kuantil yang sama, yaitu pada kuantil-kuantil di sebelah kiri kuantil 0,55 dan pada kuantil-kuantil di sebelah kanan kuantil 0,80. Ringkasan nilai *mean* dan 95% CI untuk parameter η pada kuantil 0,025, 0,05, 0,25, 0,50, 0,75, 0,95, 0,975, dan 0,99 adalah sebagaimana Tabel 4.3.

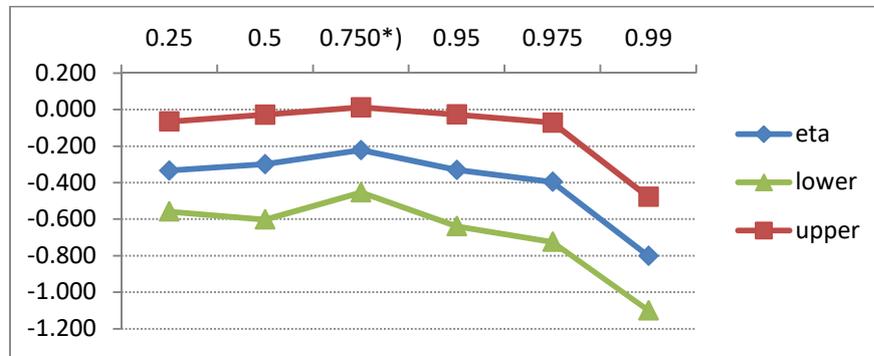
Tabel 4. 3. Nilai Eta pada Berbagai Titik Kuantil

Kuantil ke-p	Estimasi Eta	95% <i>Credible Interval</i>	
		<i>Upper</i>	<i>Lower</i>
0,025	-0,301	-0,555	-0,060
0,250	-0,334	-0,559	-0,066
0,500	-0,299	-0,602	-0,028
0,750^{*)}	-0,222	-0,454	0,012
0,950	-0,331	-0,639	-0,027
0,975	-0,396	-0,725	-0,072
0,990	-0,801	-1,100	-0,478

*) : Eta tidak signifikan

Dari Tabel 4.3 dan Gambar 4.5 terlihat bahwa CI nilai eta pada kuantil 0,75 berada pada rentang nilai yang melewati titik nol. Hal ini menunjukkan

bahwa pada kuantil tersebut, endogenitas variabel pengeluaran total rumah tangga sebagai pendekatan pendapatan rumah tangga tidak muncul.



Gambar 4. 5. Nilai Eta dan 95% *Credible Interval* pada Setiap Kuantil

Walaupun kesimpulan untuk masalah endogenitas tidak berbeda pada kedua *setting* iterasi dan *burn in* tersebut, namun variabel yang signifikan menjadi lebih banyak pada kuantil 0,05, 0,25, dan 0,50, dan 0,95. Signifikansi masing-masing variabel di masing-masing kuantil diperoleh dengan kriteria CI. Adapun ringkasan signifikansi variabel-variabel prediktor dengan adanya endogenitas adalah sebagaimana Tabel 4.4.

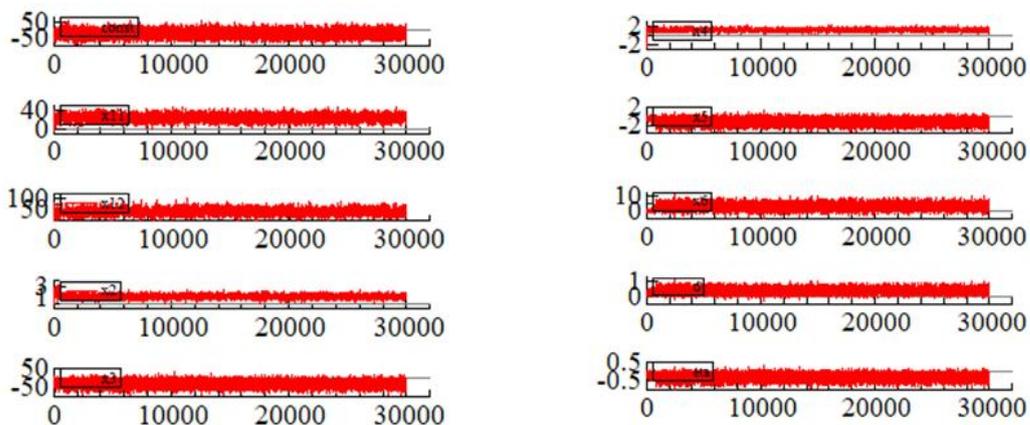
Tabel 4. 4. Signifikansi Variabel Penelitian pada Kuantil ke-p

Kuantil ke-p	Variabel Prediktor							
	d	X ₁₁	X ₁₂	X ₂	X ₃	X ₄	X ₅	X ₆
0,025	sig	sig	sig	tidak sig	sig	sig	sig	sig
0,050	sig	sig	sig	tidak sig	sig	sig	sig	sig
0,250	sig	sig	sig	sig	sig	sig	sig	sig
0,500	sig	sig	sig	sig	sig	sig	sig	sig
0,750	sig	sig	sig	sig	tidak sig	sig	sig	tidak sig
0,950	sig	sig	sig	sig	sig	sig	sig	tidak sig
0,975	sig	sig	sig	sig	sig	sig	sig	sig
0,990	sig	sig	sig	sig	sig	tidak sig	sig	sig

Pada kuantil bawah, yaitu kuantil 0,025 dan 0,05, variabel persentase pengeluaran rumah tangga untuk makanan tidak signifikan mempengaruhi pengeluaran rumah tangga untuk konsumsi susu. Sedangkan pada kuantil tengah, seluruh variabel memiliki pengaruh yang signifikan. Berbeda dengan kuantil 0,75, dimana variabel jumlah ART dan persentase ART yang bekerja tidak signifikan mempengaruhi pengeluaran rumah tangga untuk konsumsi susu. Hal ini menunjukkan bahwa pada kuantil tersebut pengeluaran untuk konsumsi susu tidak bergantung pada jumlah beban tanggungan ART, namun lebih bergantung pada keberadaan anak usia ≤ 12 tahun dan faktor lainnya. Pada kuantil 0,95, kembali variabel persentase ART yang bekerja tidak signifikan mempengaruhi pengeluaran rumah tangga untuk konsumsi susu. Namun pada kuantil 0,99 justru variabel persentase banyaknya anak usia ≤ 12 tahun yang tidak berpengaruh secara signifikan.

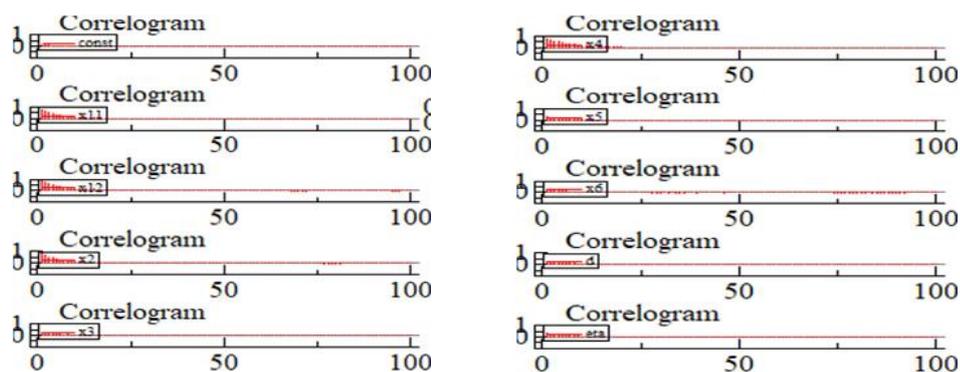
Tabel 4.4. juga menunjukkan bahwa variabel yang signifikan di semua kuantil adalah variabel pengeluaran total rumah tangga sebagai pendekatan dari variabel pendapatan rumah tangga, variabel pendidikan KRT, dan variabel rata-rata pengeluaran perkapita. Artinya, ketiga variabel tersebut secara umum berpengaruh kuat terhadap pengeluaran rumah tangga untuk konsumsi susu.

Sebagaimana dijelaskan sebelumnya bahwa kekonvergenan dalam konsep bayesian dapat dilihat dari *trace plot*. Gambar 4.6 adalah sebagai contoh bahwa proses *chain* telah konvergen pada kuantil ke 0,95 (*trace plot* pada masing-masing kuantil pada batasan penelitian terlampir).



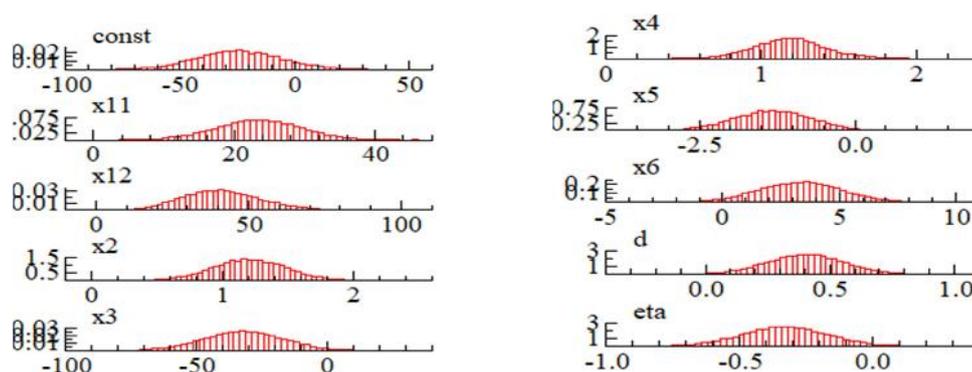
Gambar 4. 6. Trace Plot pada Kuantil 0,95

Correlogram merupakan plot yang memberikan informasi tentang autokorelasi sampel yang dibangkitkan dari masing-masing parameter. *Correlogram* parameter beta sebagaimana Gambar 4.7 menunjukkan bahwa dataset pada kuantil 0,95 bersifat random dan saling bebas (*correlogram* pada masing-masing kuantil pada batasan penelitian terlampir).



Gambar 4. 7. *Correlogram* pada Kuantil 0,95

Sedangkan plot densitas parameter beta pada kuantil 0.95 adalah sebagaimana Gambar 4.8. Dari plot tersebut, terlihat bahwa densitas sudah membentuk kurva normal, sesuai dengan distribusi likelihood yang telah ditentukan untuk parameter beta (*density plot* ada masing-masing kuantil pada batasan penelitian terlampir).



Gambar 4. 8. *Density* pada Kuantil 0,95

Untuk membentuk model TKB *Endogenous* pada masing-masing kuantil, maka disajikan ringkasan *posterior* untuk data pengeluaran rumah tangga untuk konsumsi susu di Kabupaten Sidoarjo sebagaimana Tabel 4.5.

Tabel 4. 5. Ringkasan *Posterior* pada 30.000 kali Iterasi dan 10.000 *Burn In*

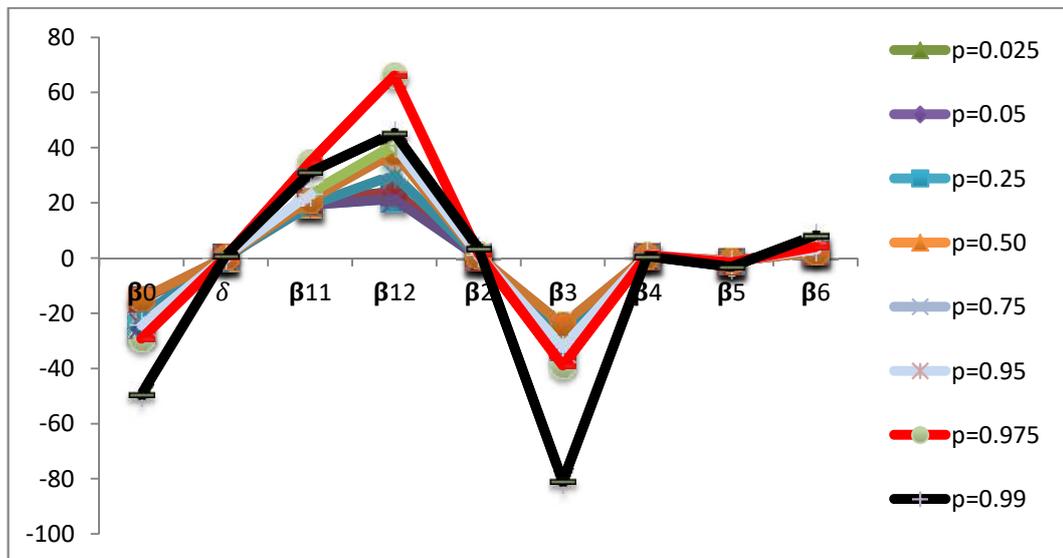
Kuantil ke-p	Parameter	Mean	95% CI	
			Upper	Lower
0,025	$\beta_{0,025,0}$	-25.079	-53.041	23.021
	$\beta_{0,025,11}$	21.150	13.066	29.148
	$\beta_{0,025,12}$	23.984	14.802	33.775
	$\beta_{0,025,2}$	0.008	-0.201	0.288
	$\beta_{0,025,3}$	-29.724	-49.109	-2.806
	$\beta_{0,025,4}$	0.562	0.429	0.692
	$\beta_{0,025,5}$	-1.229	-2.213	-0.235
	$\beta_{0,025,6}$	3.413	0.199	4.619
	$\delta_{0,025}$	0.309	0.061	0.556
	$\eta_{0,025}$	-0.309	-0.555	-0.060
0,050	$\beta_{0,050,0}$	-15.070	-53.041	23.021
	$\beta_{0,050,11}$	20.664	13.066	29.148
	$\beta_{0,050,12}$	23.801	14.802	33.775
	$\beta_{0,050,2}$	0.051	-0.201	0.288
	$\beta_{0,050,3}$	-25.762	-49.109	-2.806
	$\beta_{0,050,4}$	0.565	0.429	0.692
	$\beta_{0,050,5}$	-1.199	-2.213	-0.235
	$\beta_{0,050,6}$	2.338	0.199	4.619
	$\delta_{0,050}$	0.302	0.061	0.556
	$\eta_{0,050}$	-0.301	-0.555	-0.060
0,250	$\beta_{0,250,0}$	-24.564	-60.728	11.497
	$\beta_{0,250,11}$	19.228	10.956	27.628
	$\beta_{0,250,12}$	21.415	10.008	32.911
	$\beta_{0,250,2}$	0.325	0.054	0.602
	$\beta_{0,250,3}$	-32.636	-60.284	-4.875
	$\beta_{0,250,4}$	0.659	0.470	0.839
	$\beta_{0,250,5}$	-1.337	-2.397	-0.270
	$\beta_{0,250,6}$	3.496	0.456	6.517
	$\delta_{0,250}$	0.343	0.076	0.608
	$\eta_{0,250}$	-0.334	-0.599	-0.066

Kuantil ke-p	Parameter	Mean	95% CI	
			Upper	Lower
0,500	$\beta_{0,500,0}$	-20.914	-57.750	16.940
	$\beta_{0,500,11}$	19.197	11.353	27.251
	$\beta_{0,500,12}$	30.159	18.443	41.776
	$\beta_{0,500,2}$	0.456	0.116	0.799
	$\beta_{0,500,3}$	-27.935	-56.529	-0.869
	$\beta_{0,500,4}$	0.714	0.485	0.945
	$\beta_{0,500,5}$	-1.204	-2.412	-0.126
	$\beta_{0,500,6}$	2.908	0.008	6.384
	$\delta_{0,500}$	0.316	0.046	0.618
	$\eta_{0,500}$	-0.299	-0.602	-0.028
0,75 ^{*)}	$\beta_{0,750,0}$	-14.709	-52.688	23.629
	$\beta_{0,750,11}$	21.107	12.300	29.979
	$\beta_{0,750,12}$	38.242	23.645	52.559
	$\beta_{0,750,2}$	0.641	0.268	1.021
	$\beta_{0,750,3}$	-24.153	-49.208	0.900
	$\beta_{0,750,4}$	0.852	0.588	1.126
	$\beta_{0,750,5}$	-0.887	-1.804	0.039
	$\beta_{0,750,6}$	2.009	-0.537	4.545
	$\delta_{0,750}$	0.250	0.018	0.482
	$\eta_{0,750}$	-0.222	-0.454	0.012
0,950	$\beta_{0,950,0}$	-24.214	-60.607	13.188
	$\beta_{0,950,11}$	23.480	11.528	35.480
	$\beta_{0,950,12}$	41.065	18.569	67.017
	$\beta_{0,950,2}$	1.204	0.643	1.764
	$\beta_{0,950,3}$	-32.212	-62.570	-2.041
	$\beta_{0,950,4}$	1.181	0.700	1.679
	$\beta_{0,950,5}$	-1.333	-2.550	-0.130
	$\beta_{0,950,6}$	3.350	-0.110	6.844
	$\delta_{0,950}$	0.405	0.103	0.709
	$\eta_{0,950}$	-0.331	-0.639	-0.027

Kuantil ke-p	Parameter	Mean	95% CI	
			Upper	Lower
0,975	$\beta_{0,975,0}$	-29.215	-66.448	8.827
	$\beta_{0,975,11}$	34.714	19.799	48.772
	$\beta_{0,975,12}$	66.082	40.088	86.390
	$\beta_{0,975,2}$	1.658	1.019	2.295
	$\beta_{0,975,3}$	-38.940	-70.087	-7.641
	$\beta_{0,975,4}$	1.111	0.630	1.647
	$\beta_{0,975,5}$	-1.565	-2.842	-0.291
	$\beta_{0,975,6}$	3.972	0.280	7.706
	$\delta_{0,975}$	0.468	0.150	0.789
	$\eta_{0,975}$	-0.396	-0.725	-0.072
0,990	$\beta_{0,990,0}$	-49.487	-89.023	-10.098
	$\beta_{0,990,11}$	31.006	5.522	54.711
	$\beta_{0,990,12}$	45.311	16.441	71.742
	$\beta_{0,990,2}$	3.565	2.784	4.337
	$\beta_{0,990,3}$	-80.949	-110.756	-50.067
	$\beta_{0,990,4}$	0.600	-0.016	1.159
	$\beta_{0,990,5}$	-3.132	-4.305	-1.874
	$\beta_{0,990,6}$	8.263	4.683	11.611
	$\delta_{0,990}$	0.893	0.573	1.191
	$\eta_{0,990}$	-0.801	-1.100	-0.478

*) : tidak terdapat masalah endogenitas

Berdasarkan Tabel 4.3 dan Tabel 4.4, dapat disimpulkan bahwa penggunaan regresi kuantil pada model sudah tepat. Hal ini dilihat dari bervariasinya signifikansi variabel prediktor pada setiap kuantil. Selain itu, nilai posterior hasil estimasi dengan model TKB *endogenous* menunjukkan nilai yang bervariasi antar kuantil yang diamati. Untuk lebih jelasnya dapat dilihat dari plot *Posterior Mean* dari $(\beta'_p, \delta'_p)'$ untuk Model TKB *Endogenous* sebagaimana Gambar 4.9.



Gambar 4. 1. *Posterior Mean* dari $(\beta'_p, \delta'_p)'$ untuk Model TKB *Endogenous* pada Setiap Kuantil

Posterior dari parameter yang merupakan koefisien dari model regresi TKB *endogenous* pada setiap kuantilnya memiliki nilai yang bervariasi, namun memiliki pola yang sama untuk kuantil-kuantil bawah, kuantil-kuantil tengah, dan untuk kuantil-kuantil atas. Perbedaan yang terlihat sangat menonjol adalah pada kuantil atas, yaitu kuantil 0,975 dan 0,99.

Pada kuantil 0,975, parameter β_{12} memiliki nilai positif yang jauh lebih tinggi dibanding pada kuantil lainnya. Hal ini menjelaskan bahwa dengan pendidikan KRT yang tinggi, maka kesadaran untuk konsumsi susu bagi ART menjadi semakin besar. Di sisi lain, pada kuantil 0,99, parameter β_3 memiliki nilai negatif dan jauh lebih kecil dibandingkan nilai parameter yang sama pada kuantil yang lain. Hal yang sama juga terjadi pada kuantil 0,975. Hal ini menjelaskan bahwa dengan jumlah ART yang semakin banyak pada kuantil-kuantil atas, maka nilai konsumsi susu pada rumah tangga akan semakin kecil.

Hal yang menarik untuk diperhatikan adalah variabel *dummy* untuk pendidikan. Keduanya signifikan pada seluruh kuantil. Hal ini menjelaskan bahwa tingkatan pendidikan KRT memiliki pengaruh yang berbeda terhadap pengeluaran rumah tangga untuk konsumsi susu di Kabupaten Sidoarjo, walaupun pengaruhnya sama-sama bernilai positif. Nilai posterior yang lebih besar untuk pendidikan KRT yang tinggi (perguruan tinggi) menunjukkan bahwa semakin tinggi pendidikan

KRT, maka nilai pengeluaran rumah tangga untuk konsumsi susu menjadi semakin besar.

Berdasarkan uraian pada pembahasan ini, maka model TKB *Endogenous* untuk pengeluaran rumah tangga untuk konsumsi susu di Kabupaten Sidoarjo adalah sebagai berikut:

1. Model untuk Kuantil 0,025

$$\begin{aligned} \hat{y}_{0,025} &= -25.079 + 21.150 x_{11} + 23.984 x_{12} + 0.008x_2 - 29.724 x_3 + 0.562 x_4 \\ &\quad -1.229 x_5 + 3.413 x_6 + 0.309 d && , y^* > 0 \\ &= 0 && , y^* \leq 0 \end{aligned}$$

2. Model untuk Kuantil 0,5

$$\begin{aligned} \hat{y}_{0,05} &= - 15.070 + 20.664 x_{11} + 23.801 x_{12} + 0.051 x_2 - 25.762 x_3 + 0.565 x_4 \\ &\quad - 1.199 x_5 + 2.338 x_6 + 0.302 d && , y^* > 0 \\ &= 0 && , y^* \leq 0 \end{aligned}$$

3. Model untuk Kuantil 0,25

$$\begin{aligned} \hat{y}_{0,25} &= - 24.564 + 19.228 x_{11} + 21.415 x_{12} + 0.325 x_2 - 32.636 x_3 + 0.659 x_4 \\ &\quad - 1.337 x_5 + 3.496 x_6 + 0.343 d && , y^* > 0 \\ &= 0 && , y^* \leq 0 \end{aligned}$$

4. Model untuk Kuantil 0,50

$$\begin{aligned} \hat{y}_{0,50} &= - 20.914 + 19.197 x_{11} + 30.159 x_{12} + 0.456 x_2 - 27.935 x_3 + 0.714 x_4 \\ &\quad - 1.204 x_5 + 2.908 x_6 + 0.316 d && , y^* > 0 \\ &= 0 && , y^* \leq 0 \end{aligned}$$

5. Model untuk Kuantil 0,95

$$\begin{aligned} \hat{y}_{0,95} &= - 24.214 + 23.480 x_{11} + 41.065 x_{12} + 1.204 x_2 - 32.212 x_3 + 1.181 x_4 \\ &\quad - 1.333 x_5 + 3.350 x_6 + 0.405 d && , y^* > 0 \\ &= 0 && , y^* \leq 0 \end{aligned}$$

6. Model untuk Kuantil 0,975

$$\begin{aligned} \hat{y}_{0,975} &= - 29.215 + 34.714 x_{11} + 66.082 x_{12} + 1.658 x_2 - 38.940 x_3 + 1.111 x_4 \\ &\quad - 1.565 x_5 + 3.972 x_6 + 0.468 d && , y^* > 0 \end{aligned}$$

$$= 0 \quad , y^* \leq 0$$

7. Model untuk Kuantil 0,99

$$\begin{aligned} \hat{y}_{0,99} &= -49.487 + 31.006 x_{11} + 45.311 x_{12} + 3.565 x_2 - 80.949 x_3 + 0.600 x_4 \\ &\quad - 3.132 x_5 + 8.263 x_6 + 0.893 d \quad , y^* > 0 \\ &= 0 \quad , y^* \leq 0 \end{aligned}$$

Variabel yang memiliki tanda negatif adalah variabel jumlah ART dan rata-rata pengeluaran perkapita. Hal ini menunjukkan bahwa pada rumah tangga di Kabupaten Sidoarjo, jika jumlah ART lebih banyak, maka kecenderungan rumah tangga untuk mengkonsumsi susu akan turun. Demikian halnya jika rata-rata pengeluaran perkapita rumah tangga meningkat, maka nilai susu yang dikonsumsi rumah tangga akan turun. Dengan demikian, dalam usaha peningkatan gizi masyarakat melalui konsumsi susu, maka hal yang perlu diperhatikan pemerintah adalah jumlah ART. Hal yang bisa dilakukan untuk mendukung peningkatan konsumsi susu rumah tangga misalnya adalah dengan semakin menggalakkan program KB.

Variabel yang memiliki tanda positif adalah pengeluaran total rumah tangga, pendidikan KRT, persentase pengeluaran untuk makanan, persentase ART bekerja, dan persentase ART usia ≤ 12 tahun. Dengan demikian, dalam usaha peningkatan gizi masyarakat melalui konsumsi susu, maka hal yang perlu diperhatikan pemerintah adalah fokus pada kesejahteraan masyarakat melalui pengendalian upah/gaji buruh/karyawan/pegawai, peningkatan taraf pendidikan masyarakat, dan ketersediaan lapangan pekerjaan untuk masyarakat.

Dari model di atas terlihat bahwa tanda setiap parameter di setiap kuantil tidak ada perbedaan. Hal ini menunjukkan bahwa model TKB *Endogenous* baik untuk memodelkan data. Nilai RMSE terkecil untuk model TKB *Endogenous* pada Tabel 4.6 adalah nilai RMSE untuk model pada kuantil ke 0,75. Namun sebagaimana dibahas sebelumnya bahwa tidak terdapat masalah endogenitas pada kuantil tersebut, maka pemodelan untuk data pengeluaran rumah tangga untuk konsumsi susu di Kabupaten Sidoarjo lebih baik dilakukan dengan model TKB standar.

Tabel 4. 6. Nilai RMSE Model Regresi TKB *Endogenous* pada Kuantil ke-p

Kuantil ke-p	TKB <i>Endogenous</i>
0,025	884,74
0,050	135,76
0,250	139,55
0,500	124,19
0,750	94,98
0,950	153,05
0,975	176,41
0,990	315,47

Dengan membandingkan nilai RMSE model TKB pada Tabel 4.2 dan nilai RMSE TKB *Endogenous* pada Tabel 4.6 diperoleh kesimpulan bahwa model TKB *Endogenous* secara umum baik digunakan pada data dimana terdapat variabel *endogenous* yang juga dipengaruhi oleh variabel *exogenous* dan *instrumental*.

BAB 5

KESIMPULAN DAN SARAN

5.1. Kesimpulan

Beberapa hal yang dapat disimpulkan terkait hasil penelitian ini yaitu:

1. Model TK standar tidak bisa mengestimasi parameter pada kuantil-kuantil bawah untuk data rumah tangga untuk konsumsi susu di Kabupaten Sidoarjo. Sedangkan model TKB tidak cukup baik digunakan untuk memodelkan kuantil-kuantil bawah jika dilihat berdasarkan nilai RMSE-nya.
2. *Prior* yang digunakan adalah $\gamma \sim N(\mathbf{0}, 500\mathbf{I})$, $\sigma \sim \mathcal{IG}(0.1, 0.1)$, $\phi \sim \mathcal{IG}(2, 0.5)$, $a \sim \mathcal{G}(2, 2)$, $(\beta_p', \delta_p)' \sim N(\mathbf{0}, 500\mathbf{I})$, $\eta_p \sim N(0, 5)$, dan $\alpha \sim \mathcal{U}(0, 1)$.
3. Selanjutnya, dengan iterasi menjadi sebanyak 30.000 kali dan *burn in* 10.000 data awal, diperoleh hasil bahwa endogenitas variabel pengeluaran total rumah tangga sebagai pendekatan pendapatan rumah tangga muncul pada kuantil-kuantil di sebelah kiri kuantil 0,55 dan kuantil-kuantil di sebelah kanan kuantil 0,80. Dengan adanya endogenitas, seluruh variabel signifikan pada kuantil tengah, namun terdapat beberapa variabel yang tidak signifikan pada kuantil bawah yaitu variabel pengeluaran rumah tangga untuk makanan. Selain itu, variabel jumlah ART dan persentase ART bekerja tidak signifikan pada kuantil-kuantil atas. Berdasarkan uraian tersebut, maka model TKB *Endogenous* untuk pengeluaran rumah tangga untuk konsumsi susu di Kabupaten Sidoarjo adalah sebagai berikut:

$$\hat{y}_{0,025} = - 25.079 + 21.150 x_{11} + 23.984 x_{12} + 0.008x_2 - 29.724 x_3 + 0.562 x_4 - 1.229 x_5 + 3.413 x_6 + 0.309 d, y^* > 0$$

$$= 0, y^* \leq 0$$

$$\hat{y}_{0,05} = - 15.070 + 20.664 x_{11} + 23.801 x_{12} + 0.051 x_2 - 25.762 x_3 + 0.565 x_4 - 1.199 x_5 + 2.338 x_6 + 0.302 d, y^* > 0$$

$$= 0, y^* \leq 0$$

$$\hat{y}_{0,25} = - 24.564 + 19.228 x_{11} + 21.415 x_{12} + 0.325 x_2 - 32.636 x_3 + 0.659 x_4 - 1.337 x_5 + 3.496 x_6 + 0.343 d, y^* > 0$$

$$\begin{aligned}
&= 0 && , y^* \leq 0 \\
\hat{y}_{0,50} &= - 20.914 + 19.197 x_{11} + 30.159 x_{12} + 0.456 x_2 - 27.935 x_3 + \\
&0.714 x_4 - 1.204 x_5 + 2.908 x_6 + 0.316 d && , y^* > 0 \\
&= 0 && , y^* \leq 0 \\
\hat{y}_{0,95} &= - 24.214 + 23.480 x_{11} + 41.065 x_{12} + 1.204 x_2 - 32.212 x_3 + \\
&1.181 x_4 - 1.333 x_5 + 3.350 x_6 + 0.405 d && , y^* > 0 \\
&= 0 && , y^* \leq 0 \\
\hat{y}_{0,975} &= - 29.215 + 34.714 x_{11} + 66.082 x_{12} + 1.658 x_2 - 38.940 x_3 + \\
&1.111 x_4 - 1.565 x_5 + 3.972 x_6 + 0.468 d && , y^* > 0 \\
&= 0 && , y^* \leq 0 \\
\hat{y}_{0,99} &= - 49.487 + 31.006 x_{11} + 45.311 x_{12} + 3.565 x_2 - 80.949 x_3 + \\
&0.600 x_4 - 3.132 x_5 + 8.263 x_6 + 0.893 d && , y^* > 0 \\
&= 0 && , y^* \leq 0
\end{aligned}$$

5.2. Saran

Saran-saran yang dapat disampaikan berdasarkan hasil penelitian adalah:

1. Pada penelitian ini hanya menggunakan satu variabel *endogenous*, sehingga perlu dilakukan kajian lebih lanjut mengenai model TKB *Endogenous* untuk lebih dari satu variabel yang bersifat *endogenous*.
2. Pada penelitian ini, variabel *endogenous* yang digunakan bersifat kontinyu. Sehingga perlu pengembangan kajian lebih lanjut untuk variabel *endogenous* yang bersifat kategorik.

DAFTAR PUSTAKA

- Badan Pusat Statistik Provinsi Jawa Timur (2016), *Ringkasan Eksekutif Pengeluaran untuk Konsumsi Rumah Tangga Provinsi Jawa Timur 2016*, BPS Provinsi Jawa Timur, Surabaya.
- Cecep (2000), “Faktor–faktor yang Mempengaruhi Konsumsi Susu pada Remaja”. Institut Pertanian Bogor (IPB) , Bogor.
- Chen dan Lin, C. (2005), “An Introduction to Quantile Regression and The QUANTREG Procedure”, *Statistics and Data Analysis*, Paper 213-30.
- Chernozhukov, V. and Hong, H. (2002). “Three-Step Censored Quantile Regression and Extramarital Affairs.” *Journal of the American Statistical Association*, Vol. 155, hal. 872–882.
- Carlin, B. P. dan Chib, S. (1995), “Bayesian Model Choice via Markov Chain Monte Carlo Methods”, *Journal of the Royal Statistical Society Series B (Methodological)*, Vol. 57, No. 3, hal. 473–484.
- Chib, S. (1992), “Bayes Inference in The Tobit Censored Regression Model.” *Journal of Econometrics*, Vol. 51, hal. 79-99.
- Cintiani (2017), *Pemodelan Regresi Kuantil Tersensor Studi Kasus Pengeluaran Rumah Tangga Untuk Konsumsi Rokok*, Tesis, Institut Teknologi Sepuluh Nopember (ITS), Surabaya.
- Doornik, J. A. (2007), *Object-Oriented Matrix Programming Using Ox(3rd.)*, Timberlake Consultants Press and Oxford, London.
- Draper, N. R. dan Smith, H. (1992), *Applied Regression Analysis, Second Edition*, John Wiley and Sons, Inc., New York.
- Fitriani, Ayu (2011), *Analisis Faktor-Faktor yang Mempengaruhi Pengeluaran Rumah Tangga untuk Konsumsi Susu di Jawa Timur dengan Menggunakan Regresi Tobit* ,Skripsi, Institut Teknologi Sepuluh Nopember (ITS), Surabaya.

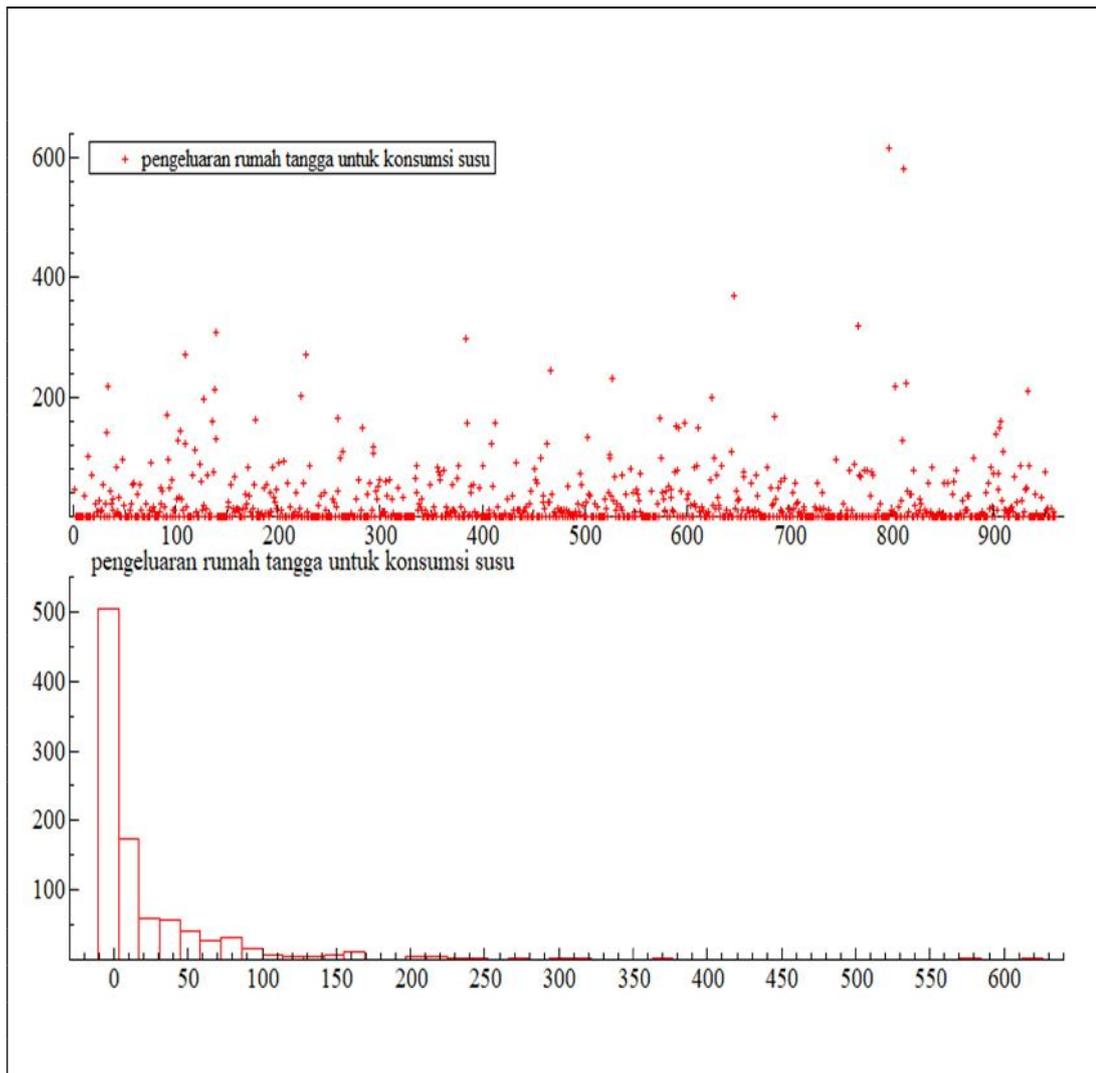
- Furno, M. (2007), "Parameter Instability in Quantile Regression", *Statistical Modelling*, Vol.7, No.4, hal. 345-362.
- Greene, W.H. (2000), *Econometric Analysis*, 4th edition, Prentice Hall, New Jersey.
- Greene, W. H. (2008), *Econometrics Analysis*, 6th edition, Prentice Hall, New Jersey.
- Gujarati, Damodar H. (2004), *Basic Econometrics: Fourth Edition*, The McGraw–Hill Companies, Inc., New York.
- Hatirli, S.A., Ozkan, B., dan Aktas, A.R. (2004), "Factors Affecting Fluid Milk Purchasing Sources in Turkey", *Food Quality and Preference*, Vol. 15, No.6, hal. 509–515.
- Hao, L. dan Naiman, D., Q. (2007), *Quantile Regression*, Sage Publication Inc., United State of America.
- Kass, R. E. dan Raftery A. E. (1995), "Bayes Factors", *Journal of The American Statistical Association*. Vol. 90, hal. 773-795.
- King, R., Morgan, B., Gimenez, O., dan Brooks, S. (2009), *Bayesian Analysis for Population Ecology*, Chapman and Hall/CRC, New York.
- Kobayashi, Genya (2017), "Bayesian Endogenous Tobit Quantile Regression", *Bayesian Analysis*, Vol. 12, No.1, hal. 161-191.
- Koenker, R. dan Basset, G. (1978), "Regression Quantiles", *Econometrica*. Vol. 46, hal. 33-50.
- Koenker, R. dan Hallock, K. (2001), "Quantile Regression", *Journal of Econometric Perspectives*, Vol. 15, hal. 143-156.
- Koenker, R. dan Machado, J. A. F. (1999), "Goodness of Fit and Related Inference Processes for Quantile Regression", *Journal of the American Statistical Association*, Vol. 94, hal. 1296-1310.
- Koop, G. (2003), *Bayesian Econometrics*, Jhon Wiley&Sons, Chichester, England.
- Kozumi, H. dan Kobayashi, G. (2011), "Gibbs Sampling Methods for Bayesian Quantile Regression", *Journal of Statistical Simulation and Computation*, Vol. 81, hal. 1565–1578.

- Lusiana, E. D. (2015), *Pemodelan Regresi Tobit Kuantil Bayesian Pada Pengeluaran Rumah Tangga Untuk Konsumsi Susu*, Tesis, Institut Teknologi Sepuluh Nopember (ITS), Surabaya.
- Long, J. Scott (1997), *Regression Models for Categorical and Limited Dependent Variables*, Sage Publications, Thousand Oaks, CA.
- Nelson (1998), *Ilmu Kesehatan Anak*, Penerbit Buku Kedokteran, Jakarta.
- Ogasawara, K. and Kobayashi, G. (2015), “The Impact of Social Workers on Infant Mortality in Inter-war Tokyo: Bayesian Dynamic Panel Quantile Regression with Endogenous Variables”, *Clometrica*, Vol. 9, hal. 97–130.
- Powell, J. L. (1986), “Censored regression quantiles”, *Journal of Econometrics*, Vol. 32, hal. 143–155.
- Purnomo, T. C. (2008), “Perbandingan Model Regresi Linier Klasik dan Tobit Bivariat Studi Kasus pada Pengeluaran Rumah Tangga Untuk Konsumsi Daging dan Susu”. Institut Teknologi Sepuluh Nopember (ITS), Surabaya.
- Raharja, P. dan Manurung, M. (2005), *Teori Ekonomi Makro*, Fakultas Ekonomi Universitas Indonesia, Jakarta.
- Raudenbush, S. W. dan Bryk, A. S.(2002), *Hierarchical Linear Models: Applications and Data Analysis Methods, Second Edition (Vol.1)*, Sage, London.
- Riduwan dan Kuncoro, E. A. (2011), *Cara Memakai dan Menggunakan Analisis Jalur (Path Analysis)*, Alfabeta, Bandung.
- Salimo, Harsono (1994), *Ilmu Kesehatan Anak: Pediatri Sosial bagi Anak dan Remaja*, UNS Press, Surakarta.
- Sethuraman, J., (1994), “A Constructive Definition of Dirichlet priors.” *Statistica Sinica*, Vol.4, hal. 639–650.
- Siswanta, Lilik (2011), “Analisis faktor-faktor yang Memengaruhi Pendapatan Pengrajin Genteng Studi Kasus pada Industri Kerajinan Genteng di Ceper, Klaten”, *Jurnal Akmenika UPY*, Vol. 7.

- Smith, R. J. dan Blundell, R. W. (1986), “An Exogeneity Test for A Simultaneous Equation Tobit Model with An Application to Labor Supply”, *Econometrica*, Vol. 56, hal. 679–685.
- Suhardi, I. Y. dan Llewelyn. R. (2001), “Penggunaan Model Regresi Tobit untuk Menganalisa Faktor-Faktor yang Berpengaruh terhadap Kepuasan Konsumen untuk Jasa Pengangkutan Barang”. *Jurnal Manajemen & Kewirausahaan*, Vol.3, No.2, hal. 106-112.
- Taddy, A. M. dan Kottas, A. (2010), “A Bayesian Nonparametric Approach to Inference for Quantile Regression”, *Journal of the American Statistical Association*, Vol. 28, hal. 357–369.
- Tobin, J. (1958), “Estimation of Relationship for Limited Dependent Variable”, *Jurnal Econometrica*, Vol.26, No.1, hal. 24-36.
- Todaro, Michael (2000), *Pembangunan Ekonomi di Dunia Ketiga*, Edisi Ketujuh Jilid I, Erlangga, Jakarta.
- Yu, K. dan Moyeed, R. A. (2001), “Bayesian Quantile Regression”, *Statistics and Probability Letters*, Vo. 54, hal. 437–447.
- Yu, K. dan Stander, J. (2007), “Bayesian Analysis of A Tobit Quantile Regression Model”, *Journal of Econometrics*, Vol. 137, hal. 260–276.
- Zain, I. dan Suhartono (1997), *Model Regresi Tobit dan Aplikasinya*, Laporan Penelitian LPPM, Institut Teknologi Sepuluh Nopember (ITS), Surabaya.

LAMPIRAN-LAMPIRAN

Lampiran 1. Plot Variabel Respon Pengeluaran Rumah Tangga untuk Konsumsi Susu



Lampiran 2. Output Uji Endogenitas dan Eksogenitas untuk Regresi Tobit dengan *Software Stata12*

```
. ivtobit Y X11 X12 X2 X3 X4 X5 X6 (D = W), ll twostep
```

Two-step tobit with endogenous regressors

Number of obs	=	960
Wald chi2(8)	=	115.73
Prob > chi2	=	0.0000

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
D	.1563714	.0718124	2.18	0.029	.0156217	.297121
X11	30.92454	10.93516	2.83	0.005	9.492026	52.35705
X12	18.75615	25.95194	0.72	0.470	-32.10872	69.62101
X2	2.231969	.7375857	3.03	0.002	.786328	3.677611
X3	-36.22061	23.0036	-1.57	0.115	-81.30683	8.865611
X4	1.19683	.2677505	4.47	0.000	.6720487	1.721611
X5	-.3159939	.1687099	-1.87	0.061	-.6466592	.0146713
X6	-.0822105	.3159233	-0.26	0.795	-.7014088	.5369877
_cons	-95.98423	36.33951	-2.64	0.008	-167.2084	-24.7601

Instrumented: D
Instruments: X11 X12 X2 X3 X4 X5 X6 W

Wald test of exogeneity: chi2(1) = 6.61 Prob > chi2 = 0.0101

Obs. summary: 497 left-censored observations at Y<=0
463 uncensored observations
0 right-censored observations

```
. reg D X11 X12 X2 X3 X4 X5 X6 W
```

Source	SS	df	MS	Number of obs = 960		
Model	903841126	8	112980141	F(8, 951)	=	360.19
Residual	298301744	951	313671.655	Prob > F	=	0.0000
Total	1.2021e+09	959	1253537.93	R-squared	=	0.7519
				Adj R-squared	=	0.7498
				Root MSE	=	560.06

	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
X11	68.04061	42.36313	1.61	0.109	-15.09539	151.1766
X12	283.9342	64.51099	4.40	0.000	157.3339	410.5346
X2	-9.080187	1.633842	-5.56	0.000	-12.28654	-5.873834
X3	298.9868	17.91361	16.69	0.000	263.832	334.1416
X4	-1.214278	1.297066	-0.94	0.349	-3.759721	1.331165
X5	2.354301	.0659111	35.72	0.000	2.224953	2.483649
X6	-4.899343	1.28026	-3.83	0.000	-7.411804	-2.386882
W	3.087245	1.025893	3.01	0.003	1.073969	5.100521
_cons	-328.892	118.4969	-2.78	0.006	-561.4376	-96.34643

Lampiran 3. Uji Multikolinieritas dengan *Software Stata12*

```
. corr D X11 X12 X2 X3 X4 X5 X6 W
(obs=960)
```

	D	X11	X12	X2	X3	X4	X5	X6
D	1.0000							
X11	-0.0430	1.0000						
X12	0.4277	-0.4394	1.0000					
X2	-0.5101	0.0207	-0.3823	1.0000				
X3	0.2761	0.0045	0.0572	0.0641	1.0000			
X4	0.0471	0.1050	0.0192	0.0267	0.2676	1.0000		
X5	0.7411	-0.0417	0.3817	-0.5281	-0.1858	-0.0747	1.0000	
X6	-0.1286	0.1001	-0.0317	0.0385	-0.5597	-0.1061	0.1962	1.0000
W	0.0926	0.1495	0.0792	0.0263	0.0591	0.1460	0.0592	0.4194


```
. vif
```

Variable	VIF	1/VIF
X6	2.16	0.462175
X3	1.81	0.551124
X12	1.68	0.595529
X5	1.60	0.626600
X2	1.54	0.648458
W	1.50	0.666177
X11	1.37	0.728579
X4	1.11	0.899761
Mean VIF	1.60	

Lampiran 4. *Syntax* ALDP dengan *Software* OxMetrics v6

```
#ifndef _INCLUDED_ALDP_
#define _INCLUDED_ALDP_
class ALDP
{
    // data
    decl m_vY;
    decl m_mX;
    decl m_vD;
    decl m_mZ;
    decl m_vCens;
    decl m_asXLAB;
    decl m_asZLAB;

    // quantile
    decl m_dProb;
    decl m_dThetaP;
    decl m_dTau2P;

    // dimension
    decl m_iN;
    decl m_cX;
    decl m_cZ;
    decl m_iCens;

    // parameters
    decl m_vBeta;
    decl m_dSigma;
    decl m_vG;
    decl m_vYs;
    decl m_vXb;

    decl m_vGamma;
    decl m_vH;
    decl m_dAlpha;
    decl m_vZg;

    decl m_vPhi;
    decl m_vIndPhi;
    decl m_vV;
    decl m_vU;
    decl m_vK;
    decl m_vPi;
    decl m_dA;
    decl m_iKs;
```

Lampiran 4. *Syntax* ALDP dengan *Software OxMetrics v6* (lanjutan)

```
decl m_iK;

decl m_dThetaA;
decl m_dTau2A;

// prior
decl m_vBeta0;
decl m_mInvB0;
decl m_vSigma0;
decl m_vGamma0;
decl m_mInvG0;
decl m_vPhi0;
decl m_vA0;

// MCMC
decl m_iBurn;
decl m_iDraw;
decl m_dStepA;
decl m_dAcceptA;
decl m_mPostBeta;
decl m_mPostGamma;
decl m_mPostPara;

// member functions
ALDP( const dP );
LoadData( const sFile1, const sFile2 );
SetPrior( const vBeta0, const dEta0, const vSigma0, const vGamma0, const
vPhi0, const vA0 );
SetMCMC( const iH );
RunMCMC( const vOpt, const vStep );
PostSum();
PostDraw( const sFlag );

traceMCMC( const iTer, const iStep );

genV();
genU();
genA();
genK();

genYs();
genBeta();
genSigma();
genG();
```

Lampiran 4. *Syntax* ALDP dengan *Software OxMetrics v6* (lanjutan)

```
genGamma();
genPhi();
genH();
genAlpha();

logPostAL( const vD, const vZg, const vPhi, const dAlpha );
setTT();
}

#endif

#include <oxstd.h>
#include <oxprob.h>
#include <oxfloat.h>
#include <oxdraw.h>
#import <maximize>
#include "subroutine.ox"

static const decl EPS = 1.0e-5;
static const decl DMAX = 7.0e+2;

ALDP::ALDP( const dP )
{
    // set the parameters
    m_dProb = dP;
    m_dThetaP = ( 1.0 - 2.0*m_dProb ) / ( m_dProb - m_dProb^2.0 );
    m_dTau2P = 2.0 / ( m_dProb - m_dProb^2.0 );

    // set format
    format( 1024 );
    format( "%#10.3f" );

    // set seed
    ranseed( 13733 );

    // print
    println( DLINE, DLINE );
    println( "Bayesian Endogenous Tobit Quantile Regression: ALDP" );
    println( "%80s", time() ~ "/" ~ date() );
    println( DLINE, DLINE );
}

ALDP::LoadData( const sFile1, const sFile2 )
{
```

Lampiran 4. *Syntax* ALDP dengan *Software OxMetrics v6* (lanjutan)

```
decl i, j, data1, data2, ydat, xdat, ddat, zdat, nc=7;

title( "LOAD DATA" );

// data for second stage
data1 = loadmat( sFile1, &m_asXLAB );
m_asXLAB = {"const"} ~ m_asXLAB[1:] ~ {"eta"};

m_iN = rows( data1 );
m_vY = data1[][0];
m_mX = 1 ~ data1[][1:] ~ zeros( m_iN, 1 );
m_cX = columns( m_mX );

m_vCens = selectifr( range(0,m_iN-1)', m_vY .== 0 );
m_iCens = rows( m_vCens );

// data for first stage
data2 = loadmat( sFile2, &m_asZLAB );
m_asZLAB = {"const"} ~ m_asZLAB[1:];

m_vD = data2[][0];
m_mZ = 1 ~ data2[][1:];
m_cZ = columns( m_mZ );

// print
println( "number of observations =", "%10d", m_iN );
println( "number of regressors   =", "%10d", m_cX );
println( "censoring rate         =", "%10.3f", m_iCens / m_iN );
println( "probability             =", "%10.3f", m_dProb );
println( LINE, LINE );
statlab();
statistics( m_vY, "y" );
println( DOT, DOT );
for ( i = 0; i < m_cX; i++ ) {
    statistics( m_mX[][i], m_asXLAB[i] );
}
println( DOT, DOT );
statistics( m_vD, "d" );
println( DOT, DOT );
for ( i = 0; i < m_cZ; i++ ) {
    statistics( m_mZ[][i], m_asZLAB[i] );
}
println( LINE, LINE );
println( NL );
}
```

Lampiran 4. *Syntax* ALDP dengan *Software* OxMetrics v6 (lanjutan)

```
ALDP::SetPrior( const vBeta0, const dEta0, const vSigma0, const vGamma0,
const vPhi0, const vA0 )
{
    title( "PRIOR" );

    // beta
    m_vBeta0 = ones( m_cX, 1 ) * vBeta0[0];
    m_mInvB0 = unit( m_cX ) ./ vBeta0[1];
    m_mInvB0[m_cX-1][m_cX-1] = 1.0 / dEta0;

    // sigma
    m_vSigma0 = vSigma0;

    // gamma
    m_vGamma0 = ones( m_cZ, 1 ) * vGamma0[0];
    m_mInvG0 = unit( m_cZ ) ./ vGamma0[1];

    // base measure, precision of DP
    m_vPhi0 = vPhi0;
    m_vA0 = vA0;

    // print
    println( "beta ~ N(", vBeta0[0], " ", vBeta0[1], " )" );
    println( "eta ~ N(", vBeta0[0], " ", dEta0, " )" );
    println( "sigma ~ IG(", vSigma0[0], " ", vSigma0[1], " )" );
    println( "gamma ~ N(", vGamma0[0], " ", vGamma0[1], " )" );
    println( "G = IG(", vPhi0[0], " ", vPhi0[1], " )" );
    println( NL );
}

ALDP::SetMCMC( const iH )
{
    decl i, h, k, p, ps;

    title( "INITIAL VALUES" );

    // set the initial values
    // beta, sigma, g, ys
    m_vBeta = zeros( m_cX, 1 );
    m_dSigma = 1.0;
    m_vXb = m_mX * m_vBeta;
    m_vG = 0.1 * ones( m_iN, 1 );
    m_vYs = m_vY;
```

Lampiran 4. *Syntax* ALDP dengan *Software OxMetrics v6* (lanjutan)

```
// gamma, h, alpha
m_vGamma = zeros( m_cZ, 1 );
m_vZg    = m_mZ * m_vGamma;
m_vH     = 0.1 * ones( m_iN, 1 );
m_dAlpha = 0.5;
setTT();

// phi, a, weights, allocation
m_vPhi    = 1.0 ./ rangamma( 1, iH, m_vPhi0[0], m_vPhi0[1] );
m_vIndPhi = zeros( m_iN, 1 );
m_dA      = 1.0;
m_vV      = ranbeta( 1, iH, 1, m_dAlpha );
m_vPi     = zeros( 1, iH );
for ( h = 0; h < iH; h++ ) {
    m_vPi[h] = h == 0 ? m_vV[h] : m_vV[h] * prodr( 1.0 - m_vV[:h-1] );
}
m_vK = zeros( m_iN, 1 );
m_vU = zeros( m_iN, 1 );
for ( i = 0; i < m_iN; i++ ) {
    k = vecindex( ranmultinomial( 1, m_vPi ), 1 );
    m_vK[i] = k;
    m_vU[i] = m_vPi[k] * ranu( 1, 1 );

    m_vIndPhi[i] = m_vPhi[k];
}
m_iK = columns( unique( m_vK ) );

p = 1.0 - min( m_vU );
h = 0;
ps = 0;
while ( ps < p ) {
    ps = sumr( m_vPi[:h] );
    ++h;
}
m_iKs = h;

// print
println( "%-10s", "beta =", m_vBeta[0] );
println( "%-10s", "sigma =", m_dSigma );
println( LINE, LINE );
println( "%-10s", "p    =", m_dProb );
println( NL );
}
```

Lampiran 4. *Syntax* ALDP dengan *Software* OxMetrics v6 (lanjutan)

```
ALDP::RunMCMC( const vOpt, const vStep )
{
    decl i, it, step;

    title( "MCMC" );

    // set iterations
    m_iBurn = vOpt[0];
    m_iDraw = vOpt[1];
    m_dStepA = vStep[0];

    m_dAcceptA = 0;
    step      = int( 0.1*m_iDraw );

    println( "number of draws =", "%10d", m_iDraw );
    println( "burn-in period =", "%10d", m_iBurn );
    print( NL );

    // set parameters
    m_mPostBeta = zeros( m_cX, m_iDraw );
    m_mPostGamma = zeros( m_cZ, m_iDraw );
    m_mPostPara = zeros( 3, m_iDraw );
    for ( it = 0; it < m_iDraw; it++ ) {
        genV();
        genU();
        genPhi();
        genK();
        genA();
        genH();
        genAlpha();
        genGamma();
        genYs();
        genG();
        genBeta();
        genSigma();

        traceMCMC( it, step );

        m_mPostBeta[][it] = m_vBeta;
        m_mPostGamma[][it] = m_vGamma;
        m_mPostPara[][it] = m_dSigma | m_dAlpha | m_dA;
    }
}
```

Lampiran 4. *Syntax* ALDP dengan *Software OxMetrics v6* (lanjutan)

```
// print
println( LINE, LINE );
println( "acceptance rate" );
println( " alpha =", "%10.3f", 100 * m_dAcceptA / m_iDraw );
println( NL );
}

ALDP::PostSum()
{
    decl i, j, pmntBeta, pmntGamma, pmntPara;
    decl asCol, asBeta, asPara, bm;
    title( "POSTERIOR SUMMARY" );

    bm = ( m_iDraw - m_iBurn ) * 0.1;
    asCol = {"mean", "std", "95% CI", "IF"};
    asPara = {"sigma", "alpha", "a"};

    // posterior summary
    postsummary( m_mPostBeta[][m_iBurn:], bm, m_asXLAB );
    postsummary( m_mPostPara[][m_iBurn:], bm, asPara );
    println( NL );
}

ALDP::PostDraw( const sFlag )
{
    decl i, j, fbeta, fgamma, fother, lag;
    decl asBeta, asGamma, asPara;

    title( "GRAPHICS" );
    // file names
    fbeta = sprintf( "beta-", sFlag, "-", "%2d", 100*m_dProb, ".pdf" );
    fgamma = sprintf( "gamma-", sFlag, "-", "%2d", 100*m_dProb, ".pdf" );
    fother = sprintf( "other-", sFlag, "-", "%2d", 100*m_dProb, ".pdf" );
    asPara = {"$\sigma$", "$\alpha$", "$a$"};
    lag = min( 100, m_iDraw - m_iBurn );
    postdraw( m_mPostBeta, m_iBurn, lag, m_asXLAB, fbeta );
    postdraw( m_mPostGamma, m_iBurn, lag, m_asZLAB, fgamma );
    postdraw( m_mPostPara, m_iBurn, lag, asPara, fother );

    // print
    println( "beta    --> ", fbeta );
    println( "gamma    --> ", fgamma );
    println( "other    --> ", fother );
    println( NL );
}
```

Lampiran 4. *Syntax* ALDP dengan *Software* OxMetrics v6 (lanjutan)

```
ALDP::traceMCMC( const iter, const nstep )
{
  if (iter==0) {
    print( "%20s", "beta[0]", "%10s", "delta", "%10s", "eta" );
    println( "%10s", "gamma[1]", "%10s", "alpha" );
    println( LINE, LINE );
  }
  if ( !fmod(iter, nstep) || iter==m_iBurn ) {
    print( "it=", "%7d", iter );
    print( "%10.3f", m_vBeta[0], "%10.3f", m_vBeta[m_cX-2], "%10.3f",
m_vBeta[m_cX-1] );
    println( "%10.3f", m_vGamma[1], "%10.3f", m_dAlpha );
  }
}

ALDP::genBeta()
{
  decl mX, mB, vb;
  mX = m_mX ./ ( m_dTau2P*m_dSigma*m_vG );
  mB = mX' * m_mX + m_mInvB0;
  vb = mX' * ( m_vYs - m_dThetaP * m_vG ) + m_mInvB0 * m_vBeta0;
  mB = invertsym( mB );
  vb = mB * vb;

  m_vBeta = vb + choleski( mB ) * rann( m_cX, 1 );

  m_vXb = m_mX * m_vBeta;
}

ALDP::genG()
{
  decl nu, aa, bb, i, g;
  nu = 0.5;
  bb = sqrt( m_dThetaP^2 / ( m_dTau2P * m_dSigma ) + 2.0 / m_dSigma );
  for ( i = 0; i < m_iN; i++ ) {
    aa = sqrt( ( m_vYs[i] - m_vXb[i] )^2 / ( m_dTau2P * m_dSigma ) );

    g = rangig( 1, 1, nu, aa, bb );
    if ( g < EPS || isnan( g ) ) {
      continue;
    }
    m_vG[i] = g;
  }
}
```

Lampiran 4. *Syntax* ALDP dengan *Software OxMetrics v6* (lanjutan)

```

ALDP::genSigma()
{
  decl e2, n1, s1;

  e2 = ( m_vYs - m_vXb - m_dThetaP*m_vG ).^2 ./ ( 2.0 * m_dTau2P * m_vG
);

  n1 = 0.5 * m_vSigma0[0] + 1.5 * m_iN;
  s1 = 0.5 * m_vSigma0[1] + sumc( e2 ) + sumc( m_vG );

  m_dSigma = 1.0 / rangamma( 1, 1, n1, s1 );
}

ALDP::genYs()
{
  decl i, j, ys;

  for ( i = 0; i < m_iCens; i++ ) {
    j = m_vCens[i];
    ys = Rantnr( m_vXb[j] + m_dThetaP*m_vG[j],
m_dTau2P*m_dSigma*m_vG[j], 0 );
    if ( !missing(ys) ) {
      m_vYs[j] = ys;
    }
  }
}

ALDP::genGamma()
{
  decl vBeta, dEta, mX, vE, mZ1, mZ2, mG, vg;

  vBeta = m_vBeta[:m_cX-2];
  dEta = m_vBeta[m_cX-1];
  mX = m_mX[:m_cX-2];
  vE = dEta * ( m_vYs - mX * vBeta - dEta * m_vD - m_dThetaP * m_vG );

  mZ1 = m_mZ ./ ( m_dTau2P * m_dSigma * m_vG );
  mZ2 = m_mZ ./ ( m_dTau2A * m_vIndPhi .* m_vH );

  mG = invertsym( mZ1' * ( dEta^2 * m_mZ ) + mZ2' * m_mZ + m_mInvG0 );
  vg = mG * ( -mZ1' * vE + mZ2' * ( m_vD - m_dThetaA * m_vH ) +
m_mInvG0 * m_vGamma0 );
}

```

Lampiran 4. *Syntax* ALDP dengan *Software OxMetrics v6* (lanjutan)

```

m_vGamma = vg + choleski( mG ) * rann( m_cZ, 1 );
m_vZg    = m_mZ * m_vGamma;

m_mX[][m_cX-1] = m_vD - m_vZg;
m_vXb          = m_mX * m_vBeta;
}

ALDP::genPhi()
{
  decl e2, n1, s1, ind, h, nj;

  for ( h = 0; h < m_iKs; h++ ) {
    ind = vecindex( m_vK, h );
    nj  = rows( ind );
    n1  = m_vPhi0[0];
    s1  = m_vPhi0[1];
    if ( nj > 0 ) {
      e2 = ( m_vD[ind] - m_vZg[ind] - m_dThetaA*m_vH[ind] ).^2 ./ ( 2.0 *
m_dTau2A * m_vH[ind] );
      n1 = m_vPhi0[0] + 1.5 * nj;
      s1 = m_vPhi0[1] + sumc( e2 ) + sumc( m_vH[ind] );
    }
    m_vPhi[h]    = 1.0 / rangamma( 1, 1, n1, s1 );
    if ( nj > 0 ) {
      m_vIndPhi[ind] = m_vPhi[h];
    }
  }
}

ALDP::genH()
{
  decl nu, aa, bb, i, h;

  nu = 0.5;
  for ( i = 0; i < m_iN; i++ ) {
    aa = sqrt( ( m_vD[i] - m_vZg[i] )^2 / ( m_dTau2A * m_vIndPhi[i] ) );
    bb = sqrt( m_dThetaA^2 / ( m_dTau2A * m_vIndPhi[i] ) + 2.0 /
m_vIndPhi[i] );

    h = rangig( 1, 1, nu, aa, bb );
    if ( h < EPS || isnan(h) ) {
      continue;
    }
  }
}

```

Lampiran 4. *Syntax* ALDP dengan *Software OxMetrics v6* (lanjutan)

```
    m_vH[i] = h;
  }
}

ALDP::genAlpha()
{
  decl alpha_o, alpha_n, lnp_o, lnp_n, a;

  alpha_o = m_dAlpha;
  alpha_n = Rantnlr( alpha_o, m_dStepA^2, 0, 1 );

  lnp_o = logPostAL( m_vD, m_vZg, m_vIndPhi, alpha_o );
  lnp_n = logPostAL( m_vD, m_vZg, m_vIndPhi, alpha_n );

  a = exp( lnp_n - lnp_o );
  if ( a > ranu(1,1) ) {
    m_dAlpha = alpha_n;
    ++m_dAcceptA;
  }
  setTT();
}

ALDP::setTT()
{
  m_dThetaA = ( 1.0 - 2.0*m_dAlpha ) / ( m_dAlpha - m_dAlpha^2.0 );
  m_dTau2A = 2.0 / ( m_dAlpha - m_dAlpha^2.0 );
}

ALDP::logPostAL( const vD, const vZg, const vPhi, const dAlpha )
{
  decl s = vPhi;
  decl e = ( vD - vZg ) ./ s;
  decl aa = dAlpha * ( 1.0 - dAlpha ) ./ s;
  decl f = log( aa ) - 0.5 * ( fabs( e ) + ( 2.0 * dAlpha - 1.0 ) * e );

  return sumc( f );
}

ALDP::genK()
{
  decl i, f, p, A, k;
```

Lampiran 4. *Syntax* ALDP dengan *Software OxMetrics v6* (lanjutan)

```

for ( i = 0; i < m_iN; i++ ) {
  A = m_vPi[:m_iKs-1] .> m_vU[i];
  f = exp( logPostAL( m_vD[i], m_vZg[i], m_vPhi[:m_iKs-1], m_dAlpha ) );
  p = f .* A;
  p = p ./ sumr( p );
  k = vecindex( ranmultinomial( 1, p ), 1 );

  m_vK[i] = k;
  m_vIndPhi[i] = m_vPhi[k];
}
m_iK = columns( unique( m_vK ) );
}

ALDP::genV()
{
  decl h, m, nj, a1, b1;

  // v
  m = 0;
  for ( h = 0; h < m_iKs; h++ ) {
    nj = sumc( m_vK .== h );
    m += nj;

    a1 = nj + 1;
    b1 = m_iN - m + m_dA;

    m_vV[h] = ranbeta( 1, 1, a1, b1 );
    m_vPi[h] = h == 0 ? m_vV[h] : m_vV[h] * prodr( 1.0 - m_vV[:h-1] );
  }
}

ALDP::genU()
{
  decl i, k, p, ps, h;

  for ( i = 0; i < m_iN; i++ ) {
    k = m_vK[i];
    m_vU[i] = m_vPi[k] * ranu( 1, 1 );
  }
  p = 1.0 - min( m_vU );
  ps = 0;
  h = 0;
  while ( ps < p ) {

```

Lampiran 4. *Syntax* ALDP dengan *Software OxMetrics v6* (lanjutan)

```
    if ( h >= m_iKs ) {
        m_vV[h] = ranbeta( 1, 1, 1, m_dA );
        m_vPi[h] = h == 0 ? m_vV[h] : m_vV[h] * prodr( 1.0 - m_vV[:h-1] );
    }
    ps = sumr( m_vPi[:h] );
    ++h;
}
m_iKs = h;
}

ALDP::genA()
{
    decl eta, k, p, a;

    eta = ranbeta( 1, 1, m_dA + 1, m_iN );
    k = ( m_vA0[0] + m_iK - 1 ) / ( m_iN * ( m_vA0[1] - log( eta ) ) );
    p = k / ( 1 + k );
    if ( p > ranu(1,1) ) {
        a = rangamma( 1, 1, m_vA0[0] + m_iK, m_vA0[1] - log( eta ) );
    }
    else {
        a = rangamma( 1, 1, m_vA0[0] + m_iK - 1, m_vA0[1] - log( eta ) );
    }
    if ( !isnan( a ) ) {
        m_dA = a;
    }
}
}
```

Lampiran 5. *Syntax* Subroutine dengan *Software OxMetrics v6*

```
#include <oxstd.h>
#include <oxprob.h>
#include <oxfloat.h>
#include <oxdraw.h>

// constant variables
const decl SEED = 13733;
const decl DLINE = "=====";
const decl LINE = "-----";
const decl DOT = ".....";
const decl NL = "\n";
const decl QUANT = <0.025, 0.5, 0.975>;
const decl NSTAT = 7;
const decl CSTAT = {"mean", "std", "min", "5%", "50%", "95%", "max"};

static title( const sT )
{
    println( DLINE, DLINE );
    println( sT );
    println( DLINE, DLINE );
}

// basic statistics
static statlab()
{
    println( LINE, LINE );
    println( "summary statistics:" );
    print( "%10s", "" );
    for (decl i = 0; i < NSTAT; i++) {
        print( "%10s", CSTAT[i] );
    }
    println( "\n", LINE, LINE );
}

static statistics( const vX, const sLab )
{
    decl i, xdat;
    xdat = zeros( 1, NSTAT );
    xdat[0] = meanc( vX );
    xdat[1] = sqrt( varc( vX ) );
    xdat[2] = min( vX );
    xdat[3:5] = quantilec( vX, QUANT );
}
```

Lampiran 5. *Syntax* Subroutine dengan *Software OxMetrics v6* (lanjutan)

```
xdat[6] = max( vX );

print( "%-10s", sLab );
    for ( i = 0; i < NSTAT; i++ ) {
        print( xdat[i] );
    }
print( "\n" );
}

static fTvar( const mX, const dBm )
{
    decl msp = periodogram( mX', dBm, 2, 1 ) / columns(mX);

    return M_2PI*( msp[0][ ] );
}

// function for calculating the posterior summary
static postsummary( const mX, const dBm, const sLab )
{
    decl i, j;
    decl npar = rows( mX );
    decl xdat = zeros( npar, 5 );
    decl asCol = {"mean", "std", "95% CI", "IF"};

    xdat[][0] = meanr( mX );
    xdat[][1] = sqrt( varr( mX ) );
    xdat[][2] = quantiler( mX, 0.025 );
    xdat[][3] = quantiler( mX, 0.975 );
    xdat[][4] = fTvar( mX, dBm ) ./ varr( mX );

    // print
    print( "%10s", "" );
    for ( i = 0; i < 4; i++ ) {
        if ( i == 2 ) {
            print( "%17s", asCol[i], "%6s", "" );
        }
        else {
            print( "%10s", asCol[i] );
        }
    }
    println( "\n", LINE, LINE );
    for ( i = 0; i < npar; i++ ) {
        print( "%-10s", sLab[i] );
    }
}
```

Lampiran 5. *Syntax Subroutine dengan Software OxMetrics v6 (lanjutan)*

```
    for ( j = 0; j < 2; j++ ) {
        print( xdat[i][j] );
    }
    print( "%5s", "(", "%8.3f", xdat[i][2], ", " );
    print( "%8.3f", xdat[i][3], ")" );
    println( xdat[i][4] );
}
println( LINE, LINE );
}

static postdraw( const mX, const iBurn, const iLag, const sLab, const sFile )
{
    decl i, j;

    for ( i = j = 0; i < rows(mX); ++i, j+=3 ) {
        DrawMatrix(    j, mX[i][:],    sLab[i], 1, 1 );
        DrawCorrelogram( j+1, mX[i][iBurn:], sLab[i], iLag );
        DrawDensity(    j+2, mX[i][iBurn:], sLab[i], 0, 1, 0 );
    }
    DrawAdjust( ADJ_AREAMATRIX, rows(mX), 3 );
    SaveDrawWindow( sFile );
    CloseDrawWindow();
}

// truncated normal s.t. x < R
static Rantnr( const vMu, const vS2, const vRight )
{
    decl s = sqrt( vS2 );
    decl u = ranu( sizer( vMu ), 1 );
    decl pr = probn( (vRight - vMu) ./ s );
    decl pl = 0.0;
    decl x = vMu + s .* quann( u .* (pr - pl) + pl );

    return x;
}

// truncated normal s.t. L < x < R
//
static Rantnr( const vMu, const vS2, const vLeft, const vRight )
{
    decl s = sqrt( vS2 );
    decl u = ranu( sizer( vMu ), 1 );
    decl pl = probn( (vLeft - vMu) ./ s );
    decl pr = probn( (vRight - vMu) ./ s );
```

Lampiran 5. *Syntax* Subroutine dengan *Software OxMetrics v6* (lanjutan)

```
decl x = vMu + s .* quann( u .* (pr - pl) + pl );  
  
return x;  
}  
  
// log density of multivariate normal distribution  
static logdensMVN( const vX, const vMu, const mSigma )  
{  
  decl p = sizer( vX );  
  decl det = determinant( mSigma );  
  decl tmp = -0.5*p*log( M_2PI ) - 0.5*log( det );  
  decl sse = -0.5*( vX - vMu )' invertsym( mSigma ) * ( vX - vMu );  
  
  return double( tmp + sse );  
}
```

Lampiran 6. *Syntax* Main Program TKB *Endogenous* dengan *Software Oxmetrics v6*

```
#include <oxstd.h>
#include <oxprob.h>
#include <oxfloat.h>
#include <oxdraw.h>
#include "aldp.ox"

main()
{
    // quantile
    decl p = 0.05;

    // prior
    decl vBeta0 = <0, 500>;
    decl dEta0 = 5.0;
    decl vGamma0 = <0, 500>;
    decl vPhi0 = <2, 0.5>;
    decl vSigma0 = <0.1, 0.1>;
    decl vA0 = <2, 2>;

    // MCMC
    decl vOpt = <10000, 30000>; // burn-in, draw
    decl vStep = <0.1>; // step size for alpha
    decl obj, time, i;

    time = timer();
    obj = new ALDP( p );
    obj -> LoadData( "data_yx.csv", "data_dz.csv" ); // second stage, first stage
data
    obj -> SetPrior( vBeta0, dEta0, vSigma0, vGamma0, vPhi0, vA0 );
    obj -> SetMCMC( 960 );
    obj -> RunMCMC( vOpt, vStep );
    obj -> PostSum();
    obj -> PostDraw( "aldp" );

    delete obj;
    // time
    println( LINE, LINE );
    println( "Execution time: ", "%62s", timespan( time ) );
    println( "Program finished: ", "%62s", date() );
    println( LINE, LINE );
}
```

Lampiran 7. Output TKB Endogenous pada Masing-Masing Kuantil

1. Output pada Kuantil ke 0,025

number of observations	= 960						
number of regressors	= 10						
censoring rate	= 0.518						
probability	= 0.025						

summary statistics:							
	mean	std	min	5%	50%	95%	max

y	23.943	51.554	0.000	0.000	0.000	159.025	617.000

const	1.000	0.000	1.000	1.000	1.000	1.000	1.000
x11	0.510	0.500	0.000	0.000	1.000	1.000	1.000
x12	0.156	0.363	0.000	0.000	0.000	1.000	1.000
x2	47.870	13.739	7.520	17.458	49.370	71.317	84.180
x3	3.659	1.359	1.000	1.000	4.000	6.000	11.000
x4	8.474	14.692	0.000	0.000	0.000	50.000	75.000
x5	359.983	346.456	53.700	83.537	269.985	1182.481	6479.190
x6	27.403	20.768	0.000	0.000	25.000	100.000	100.000
d	1229.810	1119.032	73.990	240.346	946.655	3940.186	11268.160
eta	0.000	0.000	0.000	0.000	0.000	0.000	0.000

d	1229.810	1119.032	73.990	240.346	946.655	3940.186	11268.160

const	1.000	0.000	1.000	1.000	1.000	1.000	1.000
x11	0.510	0.500	0.000	0.000	1.000	1.000	1.000
x12	0.156	0.363	0.000	0.000	0.000	1.000	1.000
x2	47.870	13.739	7.520	17.458	49.370	71.317	84.180
x3	3.659	1.359	1.000	1.000	4.000	6.000	11.000
x4	8.474	14.692	0.000	0.000	0.000	50.000	75.000
x5	359.983	346.456	53.700	83.537	269.985	1182.481	6479.190
x6	27.403	20.768	0.000	0.000	25.000	100.000	100.000
w	37.967	21.588	0.000	0.000	42.000	82.050	98.000
=====							
PRIOR							
beta ~ N(0.000 , 500.000)							
eta ~ N(0.000 , 5.000)							
sigma ~ IG(0.100, 0.100)							
gamma ~ N(0.000 , 500.000)							
G = IG(2.000 , 0.500)							
=====							
INITIAL VALUES							
beta = 0.000							
sigma = 1.000							

Lampiran 7. Output TKB *Endogenous* pada Masing-Masing Kuantil (lanjutan)

p	=	0.025				
=====						
MCMC						
=====						
number of draws	=	30000				
burn-in period	=	10000				
	beta[0]	delta	eta	gamma[1]	alpha	

it=	0	-1.011	0.000	0.000	-13.842	0.500
it=	3000	-0.753	0.295	-0.297	6.633	0.664
it=	6000	-9.227	0.281	-0.282	0.004	0.626
it=	9000	-26.862	0.148	-0.147	0.002	0.638
it=	12000	-29.785	0.097	-0.097	0.001	0.631
it=	15000	-6.590	0.355	-0.356	0.002	0.634
it=	18000	-29.808	0.315	-0.318	-0.004	0.650
it=	21000	-40.218	0.243	-0.244	0.001	0.656
it=	24000	-25.034	0.410	-0.409	-0.001	0.630
it=	27000	-30.917	0.311	-0.312	0.007	0.609

acceptance rate						
alpha	=	13.050				
=====						
POSTERIOR SUMMARY						
=====						
	mean	std	95% CI	IF		

const	-25.079	17.730	(-59.144, 10.065)	27.857		
x11	21.150	4.103	(13.917, 30.216)	402.234		
x12	23.984	4.857	(14.923, 34.010)	248.796		
x2	0.008	0.117	(-0.210, 0.245)	130.781		
x3	-29.724	14.799	(-59.067, -1.092)	30.172		
x4	0.562	0.060	(0.432, 0.669)	161.966		
x5	-1.229	0.558	(-2.345, -0.153)	41.070		
x6	3.413	1.689	(0.167, 6.804)	47.985		
d	0.309	0.140	(0.039, 0.588)	41.209		
eta	-0.309	0.140	(-0.588, -0.040)	41.427		

	mean	std	95% CI	IF		

sigma	1.247	0.057	(1.141, 1.366)	301.452		
alpha	0.630	0.014	(0.601, 0.655)	14.478		
a	0.547	0.277	(0.148, 1.211)	33.211		

Lampiran 7. Output TKB *Endogenous* pada Masing-Masing Kuantil (lanjutan)

	mean	std	95% CI	IF
const	-15.070	19.561	(-53.041, 23.021)	12.592
x11	20.664	4.079	(13.066, 29.148)	174.615
x12	23.801	4.793	(14.802, 33.775)	114.129
x2	0.051	0.121	(-0.201, 0.288)	124.293
x3	-25.762	11.790	(-49.109, -2.806)	22.332
x4	0.565	0.066	(0.429, 0.692)	70.659
x5	-1.199	0.508	(-2.213, -0.235)	65.174
x6	2.338	1.149	(0.199, 4.619)	109.395
d	0.302	0.127	(0.061, 0.556)	64.738
eta	-0.301	0.127	(-0.555, -0.060)	64.661
	mean	std	95% CI	IF
sigma	2.489	0.120	(2.263, 2.734)	91.618
alpha	0.644	0.014	(0.614, 0.671)	23.253
a	0.480	0.247	(0.129, 1.080)	6.044

3. Output pada kuantil ke 0,25

number of observations	=	960
number of regressors	=	10
censoring rate	=	0.518
probability	=	0.250
=====		
PRIOR		
=====		
beta	~ N(0.000 , 500.000)
eta	~ N(0.000 , 5.000)
sigma	~ IG(0.100 , 0.100)
gamma	~ N(0.000 , 500.000)
G	= IG(2.000 , 0.500)
=====		
INITIAL VALUES		
=====		
beta	=	0.000
sigma	=	1.000
=====		
p	=	0.250
=====		
MCMC		

Lampiran 7. Output TKB Endogenous pada Masing-Masing Kuantil (lanjutan)

=====					
number of draws = 30000					
burn-in period = 10000					
	beta[0]	delta	eta	gamma[1]	alpha

it=	0	0.336	0.001	-0.001	-13.842 0.500
it=	3000	-51.251	0.274	-0.265	0.003 0.624
it=	6000	-55.451	0.040	-0.031	-0.003 0.621
it=	9000	-41.872	0.213	-0.205	-0.003 0.621
it=	12000	-30.717	0.342	-0.336	0.005 0.637
it=	15000	-18.066	0.380	-0.369	0.004 0.633
it=	18000	-20.300	0.405	-0.393	0.002 0.633
it=	21000	-19.411	0.505	-0.494	0.001 0.661
it=	24000	1.886	0.613	-0.605	0.001 0.651
it=	27000	-2.553	0.579	-0.570	0.005 0.647

acceptance rate					
alpha = 12.653					
=====					
POSTERIOR SUMMARY					

	mean	std	95% CI	IF	

const	-24.564	18.234	(-60.728, 11.497)	5.138	
x11	19.228	4.289	(10.956, 27.628)	28.391	
x12	21.415	5.848	(10.008, 32.911)	12.362	
x2	0.325	0.140	(0.054, 0.602)	18.236	
x3	-32.636	14.058	(-60.284, -4.875)	2.663	
x4	0.659	0.094	(0.470, 0.839)	13.170	
x5	-1.337	0.540	(-2.397, -0.270)	3.352	
x6	3.496	1.534	(0.456, 6.517)	4.167	
d	0.343	0.135	(0.076, 0.608)	3.390	
eta	-0.334	0.135	(-0.599, -0.066)	3.395	

	mean	std	95% CI	IF	

sigma	12.151	0.553	(11.111, 13.283)	16.412	
alpha	0.637	0.014	(0.608, 0.663)	26.857	
a	0.675	0.304	(0.218, 1.409)	6.413	

Lampiran 7. Output TKB *Endogenous* pada Masing-Masing Kuantil (lanjutan)

4. *Output* pada kuantil ke 0,50

number of observations =	960
number of regressors =	10
censoring rate =	0.518
probability =	0.500
=====	
PRIOR	
=====	
beta ~ N(0.000 , 500.000)
eta ~ N(0.000 , 5.000)
sigma ~ IG(0.100 , 0.100)
gamma ~ N(0.000 , 500.000)
G = IG(2.000 , 0.500)
=====	
INITIAL VALUES	
=====	
beta =	0.000
sigma =	1.000
=====	
p =	0.500
=====	
MCMC	
=====	
number of draws =	30000
burn-in period =	10000
=====	
	beta[0] delta eta gamma[1] alpha

it=	0 -0.465 0.002 -0.001 -13.842 0.500
it=	3000 -33.016 0.235 -0.229 -0.001 0.639
it=	6000 -33.690 0.207 -0.186 -0.001 0.657
it=	9000 -3.164 0.320 -0.297 0.007 0.643
it=	12000 -19.744 0.271 -0.257 0.000 0.633
it=	15000 -41.975 0.223 -0.199 0.004 0.643
it=	18000 -47.272 0.277 -0.256 0.003 0.643
it=	21000 -28.777 0.295 -0.283 0.004 0.636
it=	24000 -15.084 0.363 -0.346 -0.000 0.625
it=	27000 1.046 0.492 -0.471 0.002 0.652

acceptance rate	
alpha =	12.443
=====	

Lampiran 7. Output TKB Endogenous pada Masing-Masing Kuantil (lanjutan)

POSTERIOR SUMMARY				
	mean	std	95% CI	IF
const	-20.914	19.180	(-57.750, 16.940)	80.536
x11	19.197	4.097	(11.353, 27.251)	9.069
x12	30.159	5.991	(18.443, 41.776)	8.531
x2	0.456	0.174	(0.116, 0.799)	24.241
x3	-27.935	14.060	(-56.529, -0.869)	23.137
x4	0.714	0.118	(0.485, 0.945)	5.641
x5	-1.204	0.579	(-2.412, -0.126)	65.128
x6	2.908	1.625	(0.008, 6.384)	134.427
d	0.316	0.145	(0.046, 0.618)	64.501
eta	-0.299	0.146	(-0.602, -0.028)	63.471
	mean	std	95% CI	IF
sigma	22.090	1.001	(20.199, 24.111)	9.995
alpha	0.637	0.015	(0.609, 0.664)	69.663
a	0.477	0.247	(0.125, 1.072)	14.508

5. Output pada kuantil ke 0,75

number of observations = 960
number of regressors = 10
censoring rate = 0.518
probability = 0.750

PRIOR

beta ~ N(0.000 , 500.000)
eta ~ N(0.000 , 5.000)
sigma ~ IG(0.100 , 0.100)
gamma ~ N(0.000 , 500.000)
G = IG(2.000 , 0.500)

INITIAL VALUES

beta = 0.000
sigma = 1.000

p = 0.750

Lampiran 7. Output TKB *Endogenous* pada Masing-Masing Kuantil (lanjutan)

MCMC					
number of draws = 30000					
burn-in period = 10000					
	beta[0]	delta	eta	gamma[1]	alpha
it= 0	-1.902	0.002	-0.000	-13.842	0.500
it= 3000	-50.856	0.060	-0.041	-0.002	0.616
it= 6000	10.231	0.343	-0.304	-0.000	0.622
it= 9000	-44.071	0.105	-0.082	0.004	0.667
it= 12000	-17.820	0.277	-0.257	0.002	0.641
it= 15000	-6.349	0.361	-0.331	0.000	0.659
it= 18000	11.793	0.236	-0.213	0.003	0.651
it= 21000	-6.023	0.239	-0.223	0.005	0.640
it= 24000	-1.910	0.344	-0.318	-0.001	0.644
it= 27000	-1.780	0.426	-0.403	-0.001	0.629
acceptance rate					
alpha = 13.093					
POSTERIOR SUMMARY					
	mean	std	95% CI	IF	
const	-14.709	19.421	(-52.688, 23.629)	0.613	
x11	21.107	4.472	(12.300, 29.979)	2.977	
x12	38.242	7.446	(23.645, 52.559)	4.627	
x2	0.641	0.192	(0.268, 1.021)	3.438	
x3	-24.153	12.834	(-49.208, 0.900)	1.871	
x4	0.852	0.137	(0.588, 1.126)	4.796	
x5	-0.887	0.473	(-1.804, 0.039)	2.351	
x6	2.009	1.304	(-0.537, 4.545)	2.354	
d	0.250	0.119	(0.018, 0.482)	2.415	
eta	-0.222	0.120	(-0.454, 0.012)	2.509	
	mean	std	95% CI	IF	
sigma	25.673	1.191	(23.425, 28.113)	8.762	
alpha	0.637	0.015	(0.607, 0.667)	27.887	
a	0.712	0.314	(0.238, 1.447)	9.167	

Lampiran 7. Output TKB Endogenous pada Masing-Masing Kuantil (lanjutan)

6. Output pada kuantil ke 0,95

number of observations =	960					
number of regressors =	10					
censoring rate =	0.518					
probability =	0.950					

=====						
PRIOR						
=====						
beta ~ N(0.000 , 500.000)					
eta ~ N(0.000 , 5.000)					
sigma ~ IG(0.100 , 0.100)					
gamma ~ N(0.000 , 500.000)					
G = IG(2.000 , 0.500)					
=====						
INITIAL VALUES						
=====						
beta =	0.000					
sigma =	1.000					

p =	0.950					
=====						
MCMC						
=====						
number of draws =	30000					
burn-in period =	10000					
beta[0]	delta	eta	gamma[1]	alpha		

it=	0	0.218	-0.002	0.002	-13.842	0.500
it=	3000	-39.675	0.430	-0.355	0.003	0.631
it=	6000	-0.847	0.483	-0.427	0.003	0.668
it=	9000	-9.646	0.500	-0.425	0.003	0.626
it=	12000	-31.970	0.432	-0.344	-0.000	0.659
it=	15000	-11.870	0.432	-0.357	0.004	0.621
it=	18000	-9.864	0.353	-0.260	-0.001	0.629
it=	21000	-10.734	0.436	-0.375	0.005	0.655
it=	24000	-16.410	0.374	-0.284	0.000	0.638
it=	27000	-18.855	0.430	-0.354	0.000	0.654

acceptance rate						
alpha =	12.967					
=====						

Lampiran 7. Output TKB *Endogenous* pada Masing-Masing Kuantil (lanjutan)

POSTERIOR SUMMARY				
	mean	std	95% CI	IF
const	-24.214	18.817	(-60.607, 13.188)	0.960
x11	23.480	6.067	(11.528, 35.480)	9.831
x12	41.065	12.494	(18.569, 67.017)	28.032
x2	1.204	0.284	(0.643, 1.764)	4.951
x3	-32.212	15.418	(-62.570, -2.041)	3.206
x4	1.181	0.243	(0.700, 1.679)	6.796
x5	-1.333	0.618	(-2.550, -0.130)	3.478
x6	3.350	1.779	(-0.110, 6.844)	3.742
d	0.405	0.155	(0.103, 0.709)	3.942
eta	-0.331	0.157	(-0.639, -0.027)	3.543
	mean	std	95% CI	IF
sigma	12.893	0.601	(11.763, 14.090)	34.180
alpha	0.634	0.014	(0.606, 0.662)	27.351
a	0.565	0.279	(0.157, 1.221)	9.196

7. Output pada kuantil ke 0,975

number of observations =	960
number of regressors =	10
censoring rate =	0.518
probability =	0.975
PRIOR	
beta ~ N(0.000 , 500.000)
eta ~ N(0.000 , 5.000)
sigma ~ IG(0.100 , 0.100)
gamma ~ N(0.000 , 500.000)
G = IG(2.000 , 0.500)
INITIAL VALUES	
beta =	0.000
sigma =	1.000
p =	0.975
MCMC	

Lampiran 7. Output TKB Endogenous pada Masing-Masing Kuantil (lanjutan)

number of draws = 30000						
burn-in period = 10000						
	beta[0]	delta	eta	gamma[1]	alpha	
it=	0	-6.943	-0.008	0.006	-13.842	0.500
it=	3000	-59.964	0.267	-0.192	-0.001	0.640
it=	6000	-14.326	0.601	-0.542	-0.000	0.646
it=	9000	-12.560	0.379	-0.286	0.002	0.658
it=	12000	-19.796	0.635	-0.585	0.004	0.623
it=	15000	-50.773	0.195	-0.101	0.002	0.638
it=	18000	-20.963	0.649	-0.586	0.002	0.636
it=	21000	-0.736	0.683	-0.609	0.002	0.630
it=	24000	-35.345	0.449	-0.369	0.007	0.616
it=	27000	-45.961	0.394	-0.319	0.004	0.626
acceptance rate						
alpha = 12.867						
POSTERIOR SUMMARY						
	mean	std	95% CI	IF		
const	-29.215	19.216	(-66.448, 8.827)	2.390		
x11	34.714	7.286	(19.799, 48.772)	29.112		
x12	66.082	11.832	(40.088, 86.390)	16.706		
x2	1.658	0.321	(1.019, 2.295)	17.560		
x3	-38.940	15.955	(-70.087, -7.641)	3.283		
x4	1.111	0.263	(0.630, 1.647)	20.216		
x5	-1.565	0.646	(-2.842, -0.291)	3.999		
x6	3.972	1.882	(0.280, 7.706)	4.727		
d	0.468	0.162	(0.150, 0.789)	3.834		
eta	-0.396	0.165	(-0.725, -0.072)	4.069		
	mean	std	95% CI	IF		
sigma	8.203	0.385	(7.484, 8.993)	33.975		
alpha	0.634	0.015	(0.605, 0.663)	38.886		
a	0.498	0.255	(0.133, 1.098)	13.331		

Lampiran 7. Output TKB Endogenous pada Masing-Masing Kuantil (lanjutan)

8. Output pada Kuantil ke 0,99

number of observations =	960
number of regressors =	10
censoring rate =	0.518
probability =	0.990
=====	
PRIOR	
=====	
beta ~ N(0.000 , 500.000)
eta ~ N(0.000 , 5.000)
sigma ~ IG(0.100 , 0.100)
gamma ~ N(0.000 , 500.000)
G = IG(2.000 , 0.500)
=====	
INITIAL VALUES	
=====	
beta =	0.000
sigma =	1.000
p =	0.990
=====	
MCMC	
=====	
number of draws =	30000
burn-in period =	10000
beta[0]	delta eta gamma[1] alpha

it=	0 -10.187 0.011 -0.006 -13.842 0.500
it=	3000 -51.569 0.783 -0.693 11.477 0.025
it=	6000 -37.846 0.718 -0.617 11.734 0.461
it=	9000 -40.947 0.799 -0.690 0.003 0.618
it=	12000 -51.361 0.950 -0.844 0.005 0.611
it=	15000 -63.762 0.807 -0.730 0.002 0.638
it=	18000 -64.447 0.925 -0.832 0.003 0.640
it=	21000 -65.992 0.850 -0.757 0.008 0.628
it=	24000 -72.647 1.060 -0.959 0.001 0.619
it=	27000 -19.507 1.082 -1.000 -0.002 0.646

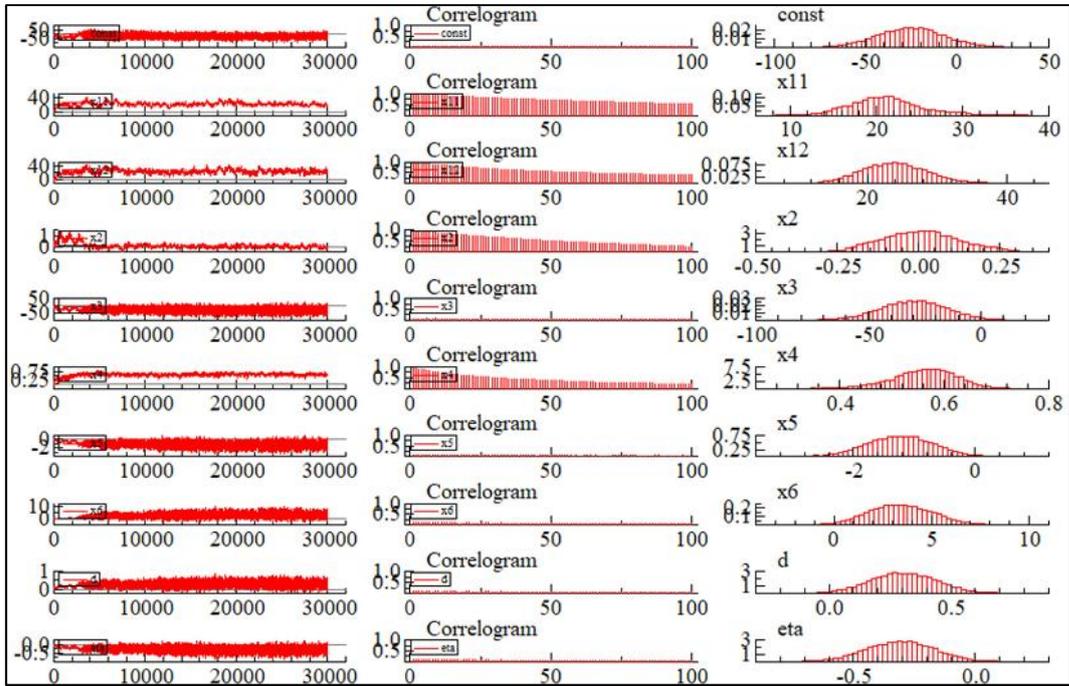
acceptance rate	
alpha =	10.887
=====	

**Lampiran 7. Output TKB Endogenous pada Masing-Masing Kuantil
(lanjutan)**

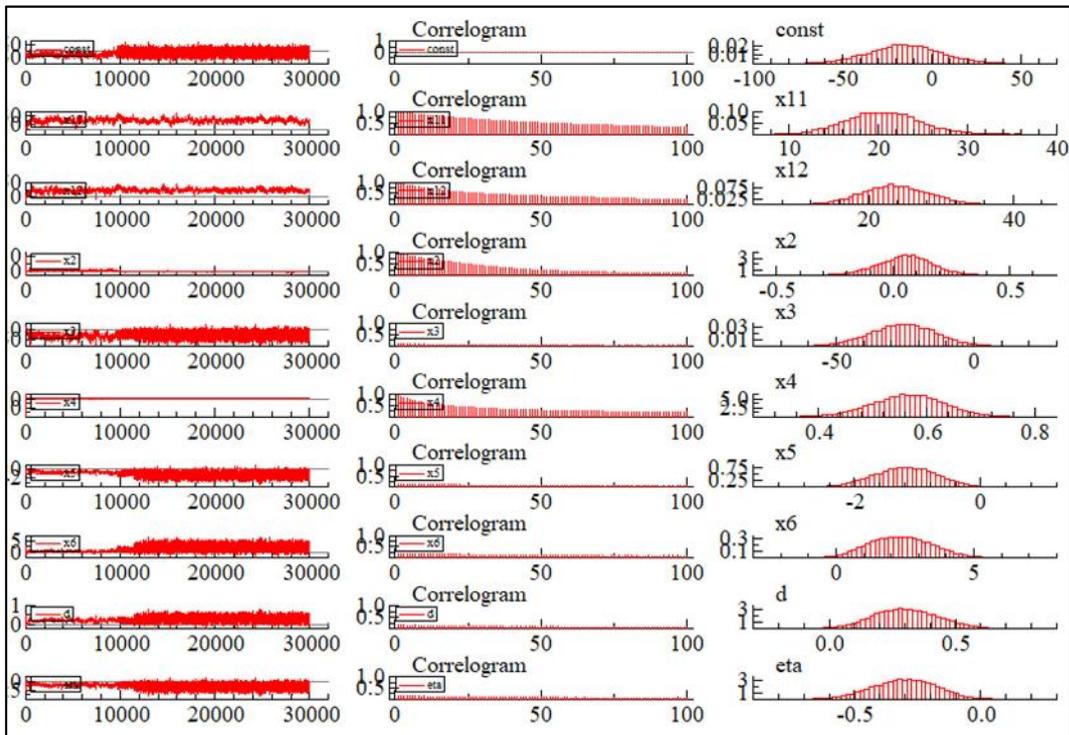
POSTERIOR SUMMARY				
	mean	std	95% CI	IF
const	-49.487	20.272	(-89.023, -10.098)	16.412
x11	31.006	12.499	(5.522, 54.711)	76.099
x12	45.311	14.051	(16.441, 71.742)	28.922
x2	3.565	0.395	(2.784, 4.337)	33.736
x3	-80.949	15.383	(-110.756, -50.067)	18.485
x4	0.600	0.298	(-0.016, 1.159)	31.649
x5	-3.132	0.613	(-4.305, -1.874)	30.869
x6	8.263	1.759	(4.683, 11.611)	54.438
d	0.893	0.156	(0.573, 1.191)	31.544
eta	-0.801	0.157	(-1.100, -0.478)	30.845
	mean	std	95% CI	IF
sigma	4.483	0.206	(4.105, 4.910)	73.966
alpha	0.634	0.014	(0.604, 0.661)	20.452
a	0.566	0.279	(0.153, 1.227)	32.414

Lampiran 8. Trace plot, Corelogram, dan Density parameter Beta

1. Pada Kuantil ke 0,025

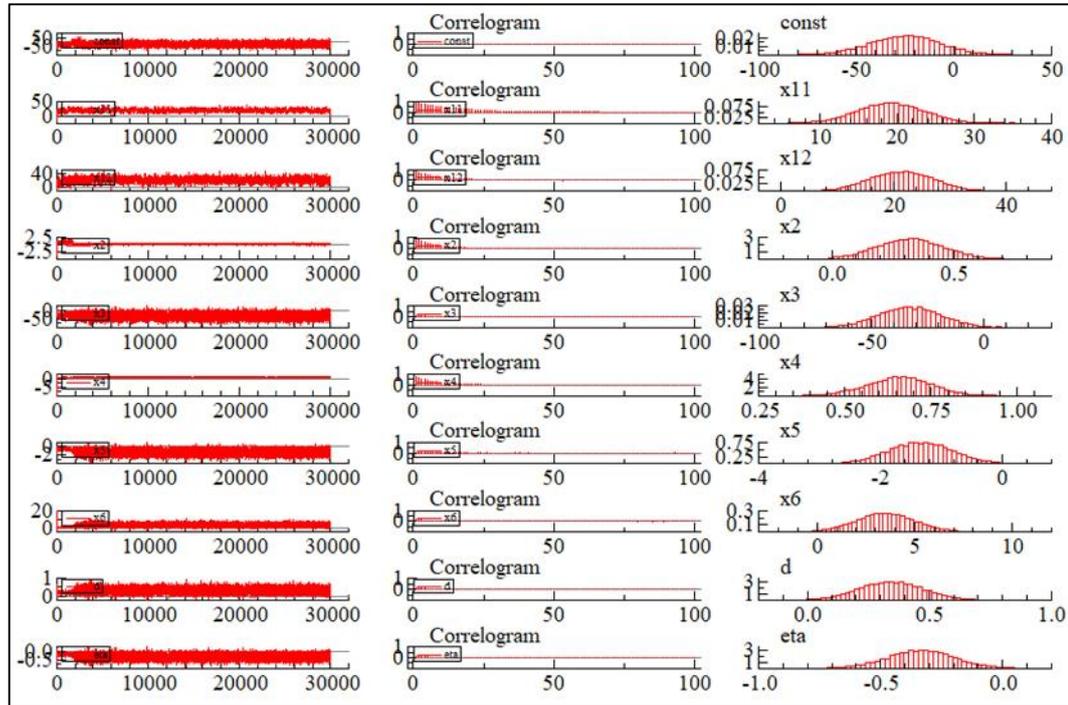


2. Pada Kuantil ke 0,05

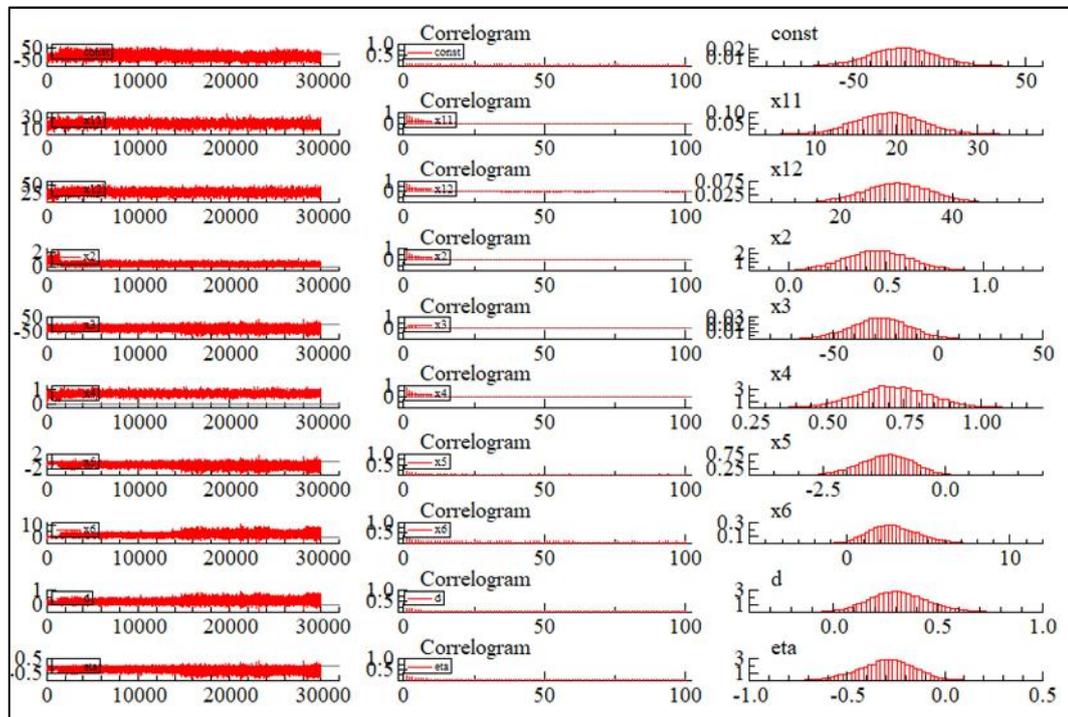


Lampiran 8. Trace plot, Corelogram, dan Density parameter Beta (lanjutan)

3. Pada Kuantil ke 0,25

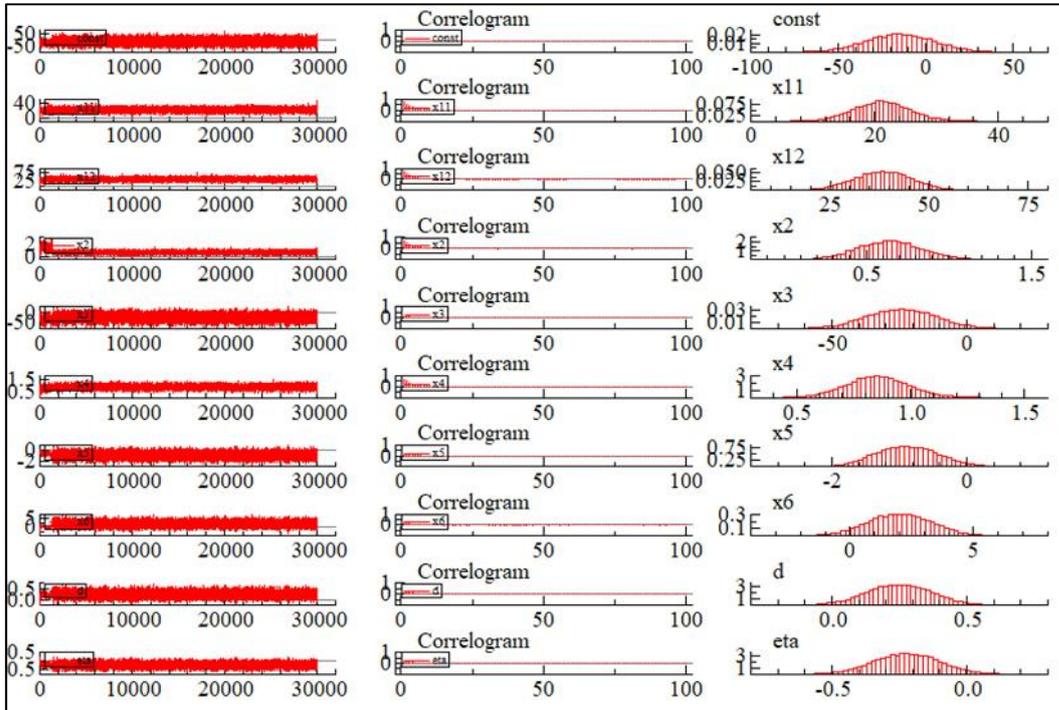


4. Pada Kuantil ke 0,50

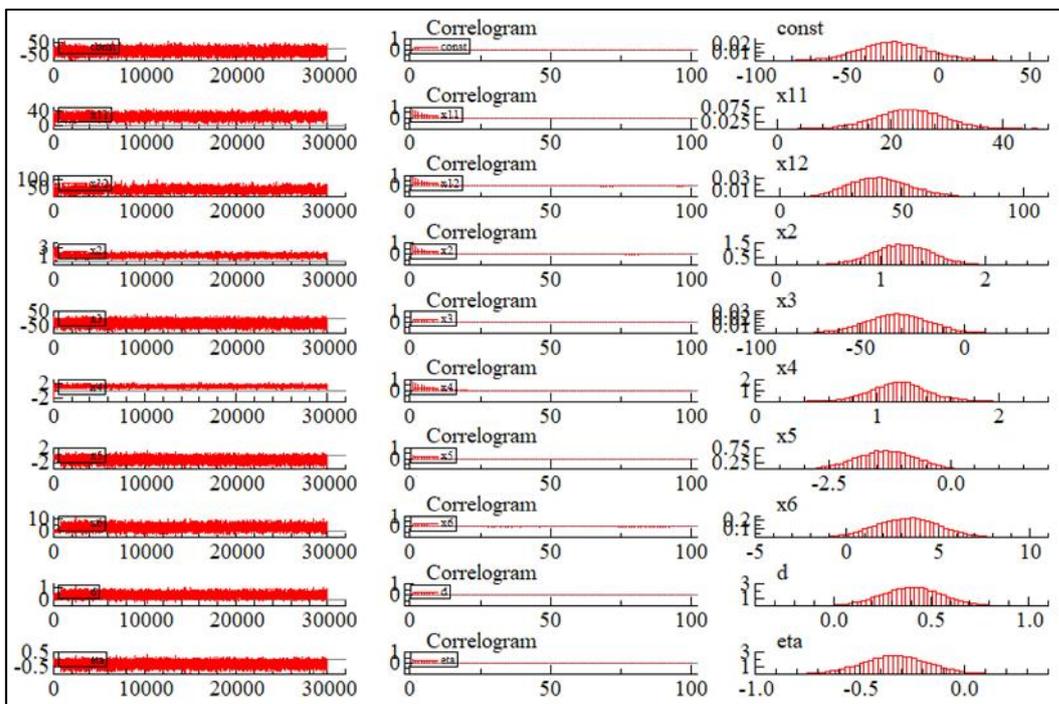


Lampiran 8. Trace plot, Corelogram, dan Density parameter Beta (lanjutan)

5. Pada Kuantil ke 0,75

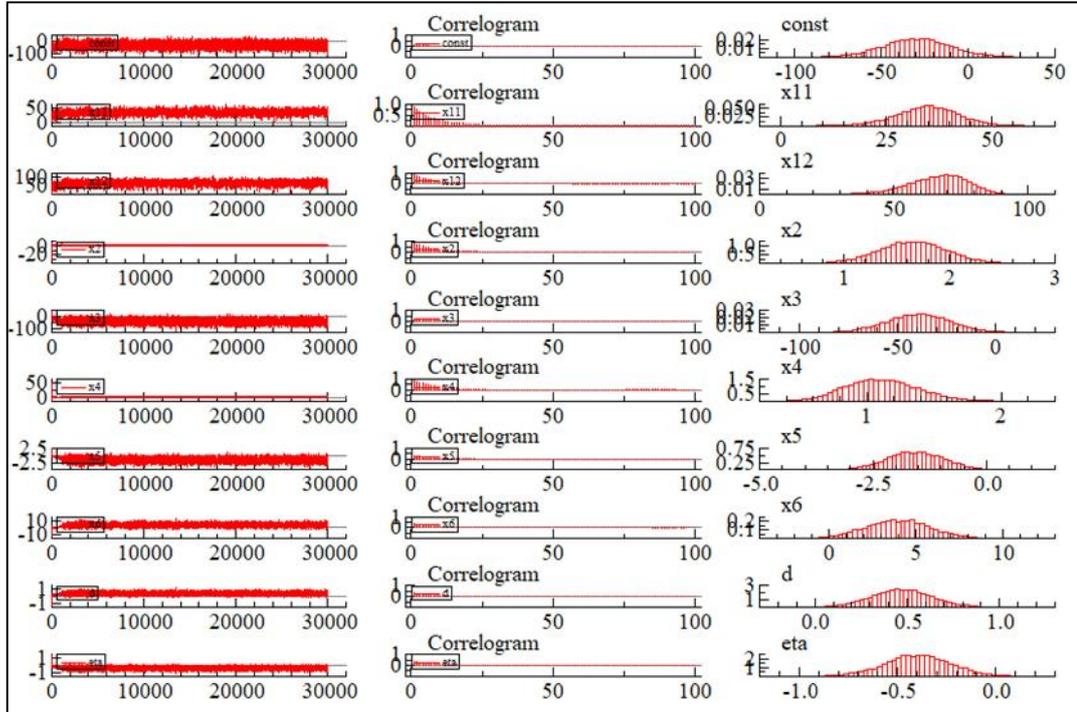


6. Pada Kuantil ke 0,95

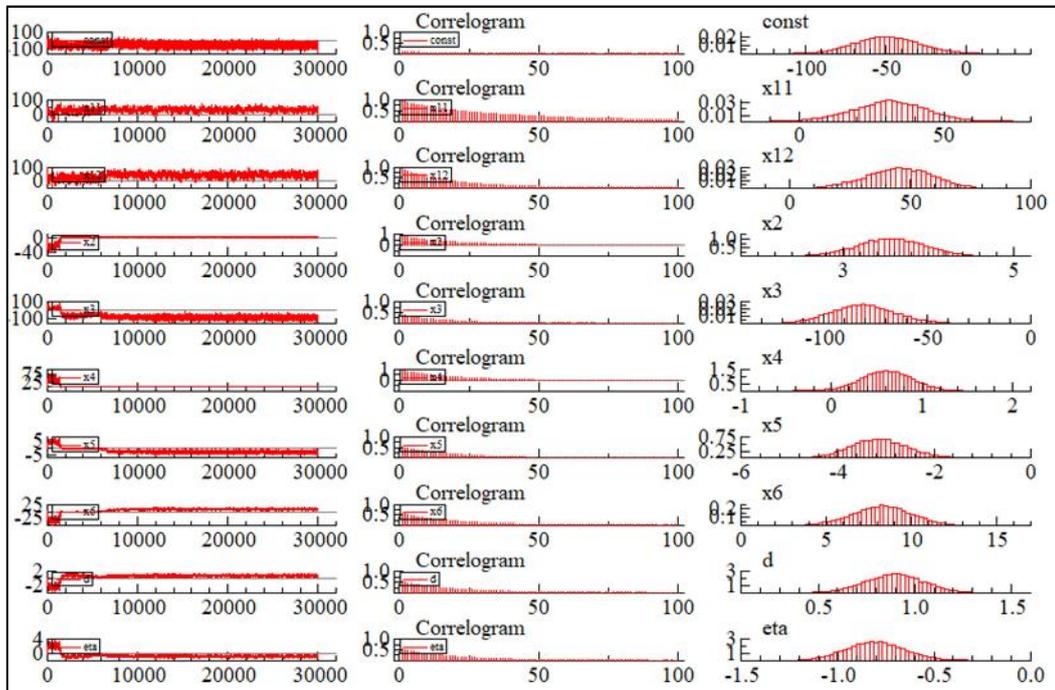


Lampiran 8. Trace plot, Corelogram, dan Density parameter Beta (lanjutan)

7. Pada Kuantil ke 0,975



8. Pada Kuantil ke 0,99



“Halaman ini sengaja dikosongkan”

BIOGRAFI PENULIS



Sartika Ayu Wulandari lahir di Blitar, Jawa Timur pada tanggal 11 November 1986, merupakan putri dari pasangan Bapak Anwari dan Ibu Sadar Susiani. Pendidikan formal yang pernah ditempuh oleh Penulis adalah SD Negeri Bagelenan 02 (1992-1998), SLTP Negeri 1 Srengat (1998-2001), SMU Negeri 1 Srengat (2001-2002), SMU Negeri 1 Waru (2002-2004), Sekolah Tinggi Ilmu Statistik (STIS) Jakarta (2004-2008). Setelah menyelesaikan program studi D-IV STIS, penulis bekerja di Badan Pusat Statistik (BPS) Kota Dumai, Provinsi Riau. Sejak September 2014, penulis dimutasi ke BPS Provinsi Jawa Timur. Pada pertengahan tahun 2016, penulis mendapatkan kesempatan dan kepercayaan oleh Pusat Pendidikan dan Pelatihan (Pusdiklat) Badan Pusat Statistik (BPS) untuk melanjutkan pendidikan Magister (S2) di Program Pascasarjana Jurusan Statistika, Fakultas Matematika, Komputasi, dan Sains Data (MKSD), Institut Teknologi Sepuluh Nopember (ITS) Surabaya. Pembaca yang ingin memberikan kritik, saran dan berdiskusi mengenai penelitian ini dapat menghubungi penulis melalui email sartika.ayu@bps.go.id.