PONTIFÍCIA UNIVERSIDADE CATÓLICA
DO RIO DE JANEIRO

Danilo Caiano Delgado

Moment-based estimation of nonlinear models

Dissertação de Mestrado

Thesis presented to the Postgraduate Program in Economics of the Departamento de Economia,PUC-Rio as partial fulfillment for the degree of Mestre em Economia.

Advisor: Prof. Marcelo Cunha Medeiros

Rio de Janeiro, September 2012

Pontifícia Universidade Católica
do Rio de Janeiro

Danilo Caiano Delgado

Moment-based estimation of nonlinear models

Thesis presented to the Postgraduate Program in Economics of the Departamento de Economia, PUC-Rio as partial fulfilment of the requirements for the degree of Mestre em Economia. Approved by the following commission.

Prof. Marcelo Cunha Medeiros
Advisor
Departamento de Economia - PUC-Rio

Prof. Álvaro Veiga
Departamento de Engenharia Elétrica - PUC-Rio

Prof. Cristiano Fernandes
Departamento de Engenharia Elétrica - PUC-Rio

Prof. Monica Herz
Coordinator of the Centro de
Ciências Sociais - PUC-Rio

Rio de Janeiro, September 28th 2012

Danilo Caiano Delgado

Undergraduate degree in Economics from Universidade Federal do Rio de Janeiro in 2009 and Master's degree in Economics from PUC-Rio in 2012.

Bibliographic data

Delgado, Danilo Caiano

Moment-based estimation of nonlinear models/ Danilo Caiano Delgado; advisor:Marcelo Medeiros;– 2012.

41 f : il. ; 30 cm

Dissertação (Mestrado em Economia)–Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2012.

Incluí referências bibliográficas.

1. Economia – Teses. 2. Modelos não-lineares 3. Estimador com função controle 4. Estimador não-linear de máxima verossimilhança de informação limitada I. Medeiros, Marcelo. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Economia. III. Título.

CDD:330

Resumo

Delgado, Danilo Caiano; Medeiros, Marcelo Cunha. **Estimação de modelos não-lineares baseados em condições de momento.** Rio de Janeiro, 2013. 41 p. Dissertação de Mestrado - Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro.

O objetivo desta dissertação é comparar através de um estudo de simulação diferentes estimadores de modelos não-lineares. Nós consideramos neste trabalho o estimador não-linear de mínimos quadrados em dois estágios (NL2SLS), o estimador não-linear de máxima verossimilhança de informação limitada (LIML) e o estimador com função controle (CF). Os resultados mostram que os estimadores CF e LIML possuem em geral uma performance superior ao do NL2SLS para os modelos selecionados. O trabalho considera uma aplicação de uma Curva de Phillips não-linear para a Economia Brasileira.

Palavras-chave

modelos não-lineares; estimador com função controle; estimador não-linear de máxima verossimilhança de informação limitada (LIML); estimador não-linear de mínimos quadrados em dois estágios (NL2SLS)

Abstract

Delgado, Danilo Caiano; Medeiros, Marcelo Cunha (Advisor). **Moment-based estimation of nonlinear models.** Rio de Janeiro, 2013. 41 p. Dissertação de Mestrado - Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro.

The aim of this dissertation is to compare, in a simulation study, different nonlinear estimators for selected models. We consider the two-stage nonlinear least-squares (NL2SLS), the nonlinear limited information maximum likelihood (LIML), and the control function (CF) estimator. Our results show that usually either CF or LIML estimators perform better than the NL2SLS estimator for the selected models. In an application with real data, we consider the estimation a nonlinear Phillips Curve for Brazilian economy.

Keywords

nonlinear models; control function estimator; nonlinear limited information maximum likelihood estimator (LIML); nonlinear two-stage least squares estimator (NL2SLS)

Sumário

List of figures

List of tables

1 Introduction

Economists often use nonlinear models because they can improve results obtained with linear models. A common method used on the estimation of nonlinear models is the nonlinear two-stage least squares (NL2SLS). In this method, among the several conditions imposed to ensure the parameter estimation, the rank condition is a key one. It states that instrumental variables must be correlated with the gradient vector of the nonlinear function evaluated at the true value of the parameters. In the linear framework, this means that instruments must be correlated with endogenous variables. However, in the nonlinear setting the instruments may have low correlation with gradient vector for some function classes. Then, the rank condition for the nonlinear framework will not be accomplished.

In order to improve correlation between instruments and the gradient vector, Takeshi (1974) derived the optimal instrument for nonlinear instrumental variables estimator. Takeshi (1975) proofed that the nonlinear limited information maximum likelihood (LIML) estimator is asymptotically more efficient than NL2SLS estimator when the endogenous variable is a linear functions of the exogenous variables. Takeshi (1974), Chamberlain (1987), Newey (1990) approximated the optimal instrument through nonparametric regression utilizing constructed instruments such as polynomials.

In this paper we compare the performance between the NL2SLS estimator, the nonlinear limited information maximum likelihood (LIML) estimator and the control function (CF) estimator for specific families of nonlinear functions. Simulated data have been generated to measure the performance of the mentioned estimators. The first example has generated data from an exponential function that is usually found Count Data models. The second example has generated data for the logistic function that is usually found in Smooth Transition Regression Models. The third example uses the same model as in the previous example except that the endogenous variable now is a nonlinear function of the exogenous variable. In the presented examples we compare bias, standard deviation, skewness and curtosis of the estimators. Simulation results show that CF estimator has better results in most of the settings for all settings, followed by the LIML estimator. The results of the developed estimators have presented for simulated data that theirs performance do not differ too much from the NL2SLS estimator. In application with real data, all the estimators have provided results that are close to the underlined economic theory.

Section 2 presents the Generalized Method of Moments (GMM) and estimation methods for nonlinear models and section 3 presents the Control Function Estimator. Section 4 shows results for simulated data and section 5 shows results for an application for a nonlinear Phillips Curve with Brazilian data. Appendix presents histograms for simulated data of Section 4.

2 GMM and nonlinear models estimation

This section presents the Generalized Method of Moments (GMM) used to obtain consistent estimators for nonlinear models. As in Takeshi (1974), consider a data generating process in which

(i) $\{u\}_{t=1}^{T}, T > 0$, is a sequence of random variables such that $E(u_t) = 0\ \forall t$, $E(u_t^2) = \sigma_0^2 < \infty\ \forall t$, $E(u_t u_s) = 0\ \forall t \neq s$;

(ii) $g(\boldsymbol{x}_t; \psi_0)$ is a nonlinear function of covariates $\boldsymbol{x}_t \in \mathbb{R}^{q_x}$ indexed by the true parameter $\psi_0 \in \Psi \subset \mathbb{R}^K$;

(iii) $\{y\}_{t=1}^{T}, T > 0$, is generated by the nonlinear model $y_t = g(\boldsymbol{x}_t; \psi_0) + u_t$;

(iv) $\boldsymbol{x}_t$ is an endogenous variable such that $E(u_t | \boldsymbol{x}_t) \neq 0$.

Define $\dot{g}(\boldsymbol{x}_t; \psi_0) = \frac{\partial g(\boldsymbol{x}_t; \psi)}{\partial \psi}|_{\psi=\psi_0}$. According to this model, we have that $E(y_t | \boldsymbol{x}_t) \neq g(\boldsymbol{x}_t; \psi_0)$ and, consequently, $E(x_t \dot{g}(\boldsymbol{x}_t; \psi_0)) \neq 0$. In this case, due to endogeneity of $\boldsymbol{x}_t$, the standard nonlinear least squares estimator for ψ_0 is inconsistent.

The endogeneity problem can be dealt with by using the Generalized Method of Moments (GMM). Let $\boldsymbol{w}_t \in \mathbb{R}^{q_w}$ be a vector of instrumental variables and $\boldsymbol{z}_t = \boldsymbol{z}_t(\boldsymbol{w}_t) \in \mathbb{R}^{q_z}, q_z \geq K$. $\boldsymbol{z}_t(\boldsymbol{w}_t): \mathbb{R}^{q_w} \to \mathbb{R}^{q_z}$ is a function of $\boldsymbol{w}_t$ such that q_z moment conditions are given by $E(u_t \boldsymbol{z}_t) = \boldsymbol{0}$.

Define

$Y_t = (y_t, x_t', z_t')$, $h(Y_t, \psi) = \frac{1}{T}\sum_{t=1}^{T} z_t [y_t - g(\boldsymbol{x}_t; \psi)] = \frac{1}{T}\sum_{t=1}^{T} z_t [u_t]$ and $\Omega = E(u_t^2 \boldsymbol{z}_t \boldsymbol{z}_t') = \sigma_0^2 E(\boldsymbol{z}_t \boldsymbol{z}_t')$. Let $\widehat{\Omega}$ be a consistent estimator for Ω. Then, the GMM estimator is given by

$$\hat{\psi}_{GMM} = argmin_{\psi \in \Psi}\left[h'(Y_t; \psi)\widehat{\Omega}^{-1} h(Y_t; \psi)\right].$$

Now define $\boldsymbol{y} = (y_1, \dots, y_T)'$, $\boldsymbol{X} = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_T)'$, $g(\boldsymbol{X}, \psi) = [g(\boldsymbol{x}_1; \psi), \dots, g(\boldsymbol{x}_T; \psi)]'$, and $\boldsymbol{Z} = (\boldsymbol{z}_1, \dots, \boldsymbol{z}_T)'$. As in Takeshi (1974), by taking σ_0^2 as constant and $\frac{1}{T}\sum_{t=1}^{T} \boldsymbol{z}_t \boldsymbol{z}_t'$ as a consistent estimator for $E(\boldsymbol{z}_t \boldsymbol{z}_t')$, the nonlinear instrumental variables estimator is

$$\hat{\psi}_{GMM} = argmin_{\psi \in \Psi} \left\{ \frac{1}{T} \sum_{t=1}^{T} z_t [y_t - g(\boldsymbol{x}_t; \psi)]' \right\}' \left\{ \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{z}_t \boldsymbol{z}_t' \right\}^{-1} \left\{ \frac{1}{T} \sum_{t=1}^{T} z_t [y_t - g(\boldsymbol{x}_t; \psi)]' \right\}$$

$$= argmin_{\psi \in \Psi} \frac{1}{T} [\boldsymbol{y} - \boldsymbol{g}(\boldsymbol{X}; \psi)]' \boldsymbol{Z} \{\boldsymbol{Z}'\boldsymbol{Z}\}^{-1} \boldsymbol{Z}' [\boldsymbol{y} - \boldsymbol{g}(\boldsymbol{X}; \psi)]$$

When $g(\boldsymbol{x}_t; \psi)$ is nonlinear in both parameters and variables and the instruments are assumed to be fixed in repeated samples, Takeshi (1974) proofs consistency and asymptotic normality of the GMM estimator for IID data. The GMM estimator is also efficient when $g(\boldsymbol{x}_t; \psi)$ is nonlinear only in the parameters.

An important condition to ensure the identification of ψ is to comply with the first order condition of the optimization problem, which states that $\text{plim}_{t \to \infty} \boldsymbol{Z}' \dot{\boldsymbol{g}}(\boldsymbol{X}; \psi_0)$ has full rank. Therefore, the instrumental variables must be correlated with the gradient vector of the nonlinear function. Thus, an instrument that is highly correlated with the endogenous variables in the linear setting may be a weak instrument in a nonlinear framework. Takeshi (1975) shows that the optimal instrument is given by $\boldsymbol{E}(\dot{g}(\boldsymbol{x}_t; \psi_0)|\boldsymbol{w}_t)$.

Now consider the following linear framework $\boldsymbol{x}_t = \Theta_0 \boldsymbol{w}_t + \boldsymbol{v}_t$, $E(\boldsymbol{v}_t|\boldsymbol{w}_t) = \boldsymbol{0}$, where $\{v_t\}_{t=1}^{T}$ is a sequence of I.I.D. random variables with zero mean and correlated with u_t. Define $\boldsymbol{W} = (\boldsymbol{w}_1, \dots, \boldsymbol{w}_T)'$, $\hat{\boldsymbol{V}} = \boldsymbol{X} - \boldsymbol{W}(\boldsymbol{W}'\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{X}$ and $\boldsymbol{I}$ as the identity matrix of order T. [Takeshi1975] shows that

$$\psi_{MIV} = argmin_{\psi\in\Psi}\frac{1}{T}[\boldsymbol{y}-\boldsymbol{g}(\boldsymbol{X};\psi)]'\boldsymbol{Z}\{\boldsymbol{Z}'\boldsymbol{Z}\}^{-1}\boldsymbol{Z}'[\boldsymbol{y}-\boldsymbol{g}(\boldsymbol{X};\psi)]$$

is more efficient than the GMM estimator.

The GMM optimization problem requires that the instrumental variables must be correlated with the gradient vector in order to ensure that $\text{plim}_{t\to\infty}\boldsymbol{Z}'\dot{\boldsymbol{g}}(\boldsymbol{X};\psi_0)$ has full rank. Nevertheless, the optimal instrument depends on the true value of the parameter ψ_0, that is unknown. The usual procedure is to estimate the parameter through nonlinear two-stage least squares. The first stage obtains a consistent estimator $\hat{\psi}$ for ψ and the second stage uses $E\left(\dot{g}\left(\boldsymbol{x}_t;\hat{\psi}\right)\middle|\boldsymbol{w}_t\right)$ as instrument. However, in the case we have weak instruments, the rank condition may not be satisfied and the resulted estimator might be inconsistent. The next section presents the Control Function estimator and how it is used to obtain consistent estimation without rely on the rank condition.

3 The Control Function Estimation

A different approach to estimate nonlinear models is given by the use of the Control Function (CF) estimator. As in Imbens and Wooldridge (2007), consider the linear model

$y_1 = \mathbf{z}_1 \boldsymbol{\delta}_1 + \alpha_1 y_2 + u_1$

Where $\mathbf{z_1}$ is a $1xL_1$ strict sub vector of $\mathbf{z}$, a $1xL$ vector of exogenous variables that includes a constant. The $\boldsymbol{L}$ zero covariance conditions are given by

$$E(\mathbf{z}'u_1) = \mathbf{0}\ .$$

We write the reduced form with an error term as

$$y_2 = \mathbf{z}\boldsymbol{\pi}_2 + v_2\ ,$$

$$E(\mathbf{z}'v_2) = 0\ .$$

y_2 is endogenous if u_1 is correlated with v_2. Write the linear projection of u_1 on v_2, in error form, as

$$u_1 = \rho_1 v_2 + e_1\ ,$$

where $\rho_1 = E(v_2 u_1)/E(v_2^2)$. Then, $E(v_2 e_1) = 0$, and $E(\mathbf{z}'e_1) = \mathbf{0}$ due to uncorrelation of z with both u_1 and v_2. Then, we have

$$y_1 = \mathbf{z}_1 \boldsymbol{\delta}_1 + \alpha_1 y_2 + u_1 = \mathbf{z}_1 \boldsymbol{\delta}_1 + \alpha_1 y_2 + \rho_1 v_2 + e_1\ .$$

OLS regression of y_1 on $\mathbf{z}_1$, y_2 and v_2 estimates consistently α_1 and δ_1 because e_1 is uncorrelated with z_1, y_1 and v_2 by construction. However, v_2 is not observed. The suggested approach is to estimate $\boldsymbol{\pi}_2$ in a first moment by running OLS regression on the reduced form equation. This first stage will provide $\hat{v}_2 = y_2 - \mathbf{z}\hat{\boldsymbol{\pi}}_2$, that will replace v_2 in the structural equation:

$$y_1 = \mathbf{z}_1 \boldsymbol{\delta}_1 + \alpha_1 y_2 + \rho_1 \hat{v}_2 + e_1 + \rho_1 \mathbf{z}(\hat{\boldsymbol{\pi}}_2 - \boldsymbol{\pi}_2)\ .$$

OLS estimation by running a regression of y_1 on z_1, y_2 and $\hat{v}_2$ is called control function estimation, in the sense that inclusion of $\hat{v}_2$ in the structural equation controls for the endogeneity of y_2. The error term depends on the sampling error from $(\hat{\pi}_2 - \pi_2)$, unless $\rho_1 = 0$ (no endogeneity). Some algebra shows that function control estimates are identical to 2SLS estimates. When correctly specified, the control function (CF) estimator is more efficient than the IV estimator. However, its estimation depends on assumptions of linear relation between u_t and v_t, what makes CF estimator less robust than the IV estimator, once it may not be consistent for different functional forms.

Now consider the following nonlinear model in which $y_t = g(x_t;\psi_0) + u_t$ and $x_t = f(z_t) + v_t$. Assume that $E(u_t|x_t) \neq 0$ and that the relation between u_t and v_t is linear, i.e. , $u_t = \alpha v_t + e_t$. We also assume also that $E(e_t|z_t) = 0$ and $(e_t|x_t) = 0$, that is more restrictive than the previous correlation hypothesis. Then we have

$$y_t = g(x_t;\psi_0) + \alpha v_t + e_t$$

$$= g(x_t;\psi_0) + \alpha(x_t - f(z_t)) + e_t.$$

This model can be estimated through simple nonlinear least-squares as well. The first step consists on the construction of a residual vector $\hat{v}$ from the nonparametric regression of x_t on z_t. The second step runs the regression of y_t on $g(x_t;\psi_0)$ and $\hat{v}$. As before, CF estimator is more efficient than the IV estimator when correctly specified, but less robust to different functional forms.

The next session shows simulation results for LIML, NL2SLS and CF estimators for different data generating processes in order to evaluate theirs performances for different nonlinear models.

4 Simulation

In this section, we use simulated data to evaluate the performance of the indicators for selected non-linear models. Simulated data was generated by using the exponential and logistic functions, that are widely used in the economic models as Count Data models and Smooth Transition Regression (STR) models, respectively. We present results for NL2SLS, LIML, and CF estimators (estimations of the control function parameter are omitted for simplicity). Histograms are provided in Appendix We make 1000 simulations with sample sizes N=100, N=250 and N=500. We constrain the parameters of the optimization problem on intervals of length R=8. The first example is given by the following data generating process:

$$y_i = \exp(\beta_1 + \beta_2 x_{1i} + \beta_3 x_{2i}) + u_i$$

$$x_{1i} = w_i + e_i$$

$$u_i = v_i + e_i$$

In this example, w_i, u_i,x_{2i} and v_i are standard normal *iid* random variables. x_{1i} is an endogenous variable and w_i is an exogenous observed variable. The true parameter values are $\beta_1 = 2$, $\beta_2 = -1$, $\beta_3 = -0.5$. Results are shown on Table 1.

TABLE 1 – Estimation Results of Example1

Parameter	Sample Size	Estimator	Bias	Standard Deviation	Skewness	Kurtosis
BETA 1	100	LIML	-0.4287	1.6921	-3.3140	10.8972
		NL2SLS	-0.3841	1.8838	-3.6799	11.9317
		CF	-0.2513	1.4215	-4.3331	20.0472
	250	LIML	-0.3216	1.3761	-3.6113	14.7775
		NL2SLS	-0.1259	1.3926	-5.4208	27.8049
		CF	-0.1045	0.9790	-5.3180	36.8406
	500	LIML	-0.1966	1.0665	-4.2078	23.0809
		NL2SLS	-0.0287	1.1154	-6.9402	46.8110
		CF	-0.0559	0.9913	-5.6619	39.2679
BETA 2	100	LIML	-0.0698	0.6219	1.3624	19.4416
		NL2SLS	-0.0280	0.6759	0.4782	19.1284
		CF	-0.0894	0.4515	-3.7518	17.1932
	250	LIML	-0.0842	0.4410	0.6248	20.8432
		NL2SLS	0.0018	0.4553	0.7502	36.9238
		CF	-0.0457	0.2730	-3.5570	21.6015
	500	LIML	-0.0637	0.3566	1.6965	35.9536
		NL2SLS	0.0025	0.3283	1.2245	65.1663
		CF	-0.0285	0.2526	-4.1127	25.8734
BETA 3	100	LIML	-0.1083	0.6512	-3.9627	41.9855
		NL2SLS	-0.1137	0.8022	-3.3253	31.1256
		CF	-0.0455	0.3748	-3.6093	31.1749
	250	LIML	-0.0549	0.4242	-1.1224	46.0618
		NL2SLS	-0.0418	0.4596	-5.0496	69.1375
		CF	-0.0273	0.2846	-4.0060	30.1992
	500	LIML	-0.0432	0.3152	-7.2574	79.5337
		NL2SLS	-0.0331	0.3653	-9.2114	94.3324
		CF	-0.0078	0.2571	0.6042	43.7625

Results from Table 1 show that CF estimator performs better than NL2SLS and LIML estimators for all parameters and all sample sizes when standard deviation is observed. CF estimator also shows good performance on bias for parameters β_1 and β_3. LIML estimator has in most configurations lower standard deviation than NL2SLS estimator, but it shows usually higher bias. The results for standard deviation are according expectations, given that the CF estimator is more efficient than the other estimators. On Skewness, LIML estimator presents in most settings the lowest skewness measures in absolute values while FC estimator usually presents the higher measures. In most settings, the skewness values are negative. Kurtosis values vary among different settings, without showing a clear pattern among different settings.

In the next example, we simulate data for the logistic function. The simulated data ha the following generating process:

$$y_i = \beta_1 + \beta_2 x_i + (\beta_3 + \beta_4 x_i)\frac{1}{1+\exp(-\gamma(x_i - c))} + u_i$$

$$x_i = w_i + e_i$$

$$u_i = w_i + v_i$$

In this example, w_i, u_i and v_i are standard normal *iid* random variables. x_i is the endogenous variable and w_i is an exogenous observed variable. The true parameter values are $\beta_1 = -0.2$, , $\beta_2 = 1.4$, $\beta_3 = 0.6$, $\beta_4 = -2.3$, $\gamma = 0.10$, $c = -2$. Each coordinate is a real interval with length $R = 12$ centralized in the true parameter value. Results are shown on Table 2.

TABLE 2 – Estimation Results of Example 2

Parameter	Sample Size	Estimator	Bias	Standard Deviation	Skewness	Curtosis
BETA 1	100	LIML	0.0547	0.3935	-4.6685	74.1905
		NL2SLS	0.0173	1.2410	-2.0484	35.2310
		CF	-0.0065	0.1633	-3.8999	41.4142
	250	LIML	0.0813	0.1882	0.5200	3.8301
		NL2SLS	-0.0724	1.1717	-6.4681	51.0804
		CF	0.0221	0.0737	-0.0116	0.0628
	500	LIML	0.0572	0.1220	0.0102	1.1245
		NL2SLS	-0.1198	0.9356	-6.6804	54.2414
		CF	0.0213	0.0509	0.0233	0.0392
BETA 2	100	LIML	0.0710	0.3074	-0.5326	2.1532
		NL2SLS	0.0664	0.6767	0.1401	22.3509
		CF	-0.0061	0.1874	-0.9391	2.9145
	250	LIML	0.0725	0.1837	-0.5073	2.7746
		NL2SLS	0.0108	0.5039	-4.3014	30.9611
		CF	0.0151	0.1023	-0.2612	0.3338
	500	LIML	0.0570	0.1193	-0.4182	1.5387
		NL2SLS	-0.0242	0.3933	-5.1687	42.3855
		CF	0.0211	0.0697	-0.0003	-0.0817
BETA 3	100	LIML	-0.0603	6.9763	-0.2306	-0.7602
		NL2SLS	-0.0118	8.2506	-0.0553	-1.1953
		CF	2.8607	3.9409	-0.7023	1.6471
	250	LIML	0.0406	5.2844	-0.4565	0.3627
		NL2SLS	0.6276	6.9347	-0.1694	-0.5846
		CF	2.7364	2.2395	-0.2070	1.3943
	500	LIML	0.8264	4.0105	-0.5621	1.9981
		NL2SLS	1.7073	5.5606	-0.2560	0.2223
		CF	2.6687	1.5686	-0.2180	0.1117
BETA 4	100	LIML	0.0341	2.6524	0.2261	-0.0584
		NL2SLS	0.1950	3.0869	0.1787	-0.4619
		CF	-0.9987	1.4406	0.4215	1.9739
	250	LIML	-0.0298	1.9751	0.3975	0.8144
		NL2SLS	-0.0346	2.4714	0.1414	-0.0404
		CF	-0.9409	0.8094	-0.0992	1.9126
	500	LIML	-0.2741	1.4639	0.5450	2.2567
		NL2SLS	-0.3741	1.9320	0.3595	0.8724
		CF	-0.8972	0.5523	0.0578	0.2211
GAMMA	100	LIML	28.0135	48.1874	1.2298	-0.2753
		NL2SLS	70.0454	56.9214	-0.3318	-1.8067
		CF	23.2848	46.1554	1.4868	0.4017
	250	LIML	10.2422	33.3717	2.7430	6.0586
		NL2SLS	66.4734	58.3630	-0.2284	-1.8807
		CF	4.1552	22.7637	4.3923	18.9138
	500	LIML	3.9727	21.7184	4.4426	19.9778
		NL2SLS	64.6422	58.3556	-0.1730	-1.8958
		CF	-0.3235	9.4305	9.7049	112.0940
C	100	LIML	0.2997	0.3086	2.2638	13.8849
		NL2SLS	0.5769	0.6285	2.0932	6.0902
		CF	0.2521	0.2500	0.9479	1.6021
	250	LIML	0.2333	0.2028	2.4290	12.9883
		NL2SLS	0.5300	0.5530	2.2025	6.5778
		CF	0.1972	0.1428	0.5322	1.0059
	500	LIML	0.2080	0.1367	2.0720	11.5628
		NL2SLS	0.4708	0.4801	2.3090	7.8329
		CF	0.1900	0.1026	0.0479	-0.4559

Results from Table 2 show that CF estimator has again a better performance on standard deviation than NL2SLS and LIML estimators for all parameters and sample sizes. CF estimator shows good performance on bias for parameters β_1, β_2, γ, and c . LIML estimator has better performance on bias for parameters β_3 and β_4 than CF and NL2SLS estimators. The results for standard deviation are according expectations, given that the CF estimator is more efficient than the others. This result holds from the previous example due to the linear relation between the endogenous variable and the instrumental variable. The next example changes this linear relationship. Therefore, we should expect that CF estimator will not be the more efficient anymore. Performance on skewness differ among estimators depending on the estimated parameter. NL2SLS has better results for gamma and CF has better results for c. For other parameters, there is not an estimator that is consistent closer to zero than the others. On kurtosis, there is no estimator that is consistently closer to zero than other estimators. It only happens on the estimation of the parameter c, in which FC estimator shows results closer to zero for all sample sizes.

In the third example, 1000 simulated data have been generated with the same specifications as the previous example, except in the data generating process of the endogenous variable. In this example x_I is a quadratic function of the exogenous variable w_i:

$$y_i = \beta_1 + \beta_2 x_i + (\beta_3 + \beta_4 x_i)\frac{1}{1 + \exp(-\gamma(x_i - c))} + u_i$$

$$x_i = -2 + 4w_i^2 + e_i$$

$$u_i = w_i + v_i$$

Simulation results are presented on Table 3:

TABLE 3 – Estimation Results of Example 3

Parameter	Sample Size	Estimator	Bias	Standard Deviation	Skewness	Curtosis
BETA 1	100	LIML	0.3192	1.3920	0.9280	11.1951
		NL2SLS	0.4787	3.3778	-0.2695	3.9673
		CF	0.4620	0.4057	-5.9099	110.1470
	250	LIML	0.3146	1.1494	1.2199	19.5811
		NL2SLS	0.4262	4.1899	-0.3713	2.1552
		CF	0.4340	0.2139	0.5118	0.2970
	500	LIML	0.1704	0.7037	3.4992	21.1529
		NL2SLS	0.2287	4.9284	-0.4562	1.1509
		CF	0.3990	0.1485	0.4925	0.7266
BETA 2	100	LIML	0.2731	1.1421	1.1509	9.7344
		NL2SLS	0.4207	2.2282	0.1232	2.5026
		CF	0.6061	1.5012	1.8964	42.9453
	250	LIML	0.2758	0.9367	2.5995	17.0629
		NL2SLS	0.5149	2.7022	0.1816	2.0787
		CF	0.4680	1.1482	-1.0367	70.2909
	500	LIML	0.1437	0.5919	3.4469	19.7187
		NL2SLS	0.6203	2.9106	0.1225	1.0238
		CF	0.4829	1.3803	2.1850	45.3437
BETA 3	100	LIML	-0.6854	2.5421	-0.8842	6.3821
		NL2SLS	-0.8770	4.7317	0.1104	1.5145
		CF	-0.1971	0.7941	1.1515	16.5548
	250	LIML	-0.5733	1.9576	-2.0341	13.6920
		NL2SLS	-0.7401	5.2045	0.3080	1.0235
		CF	-0.2117	0.4489	0.0494	0.7688
	500	LIML	-0.3110	1.2620	-3.6649	21.2593
		NL2SLS	-0.4731	5.8241	0.3045	0.3829
		CF	-0.2475	0.2949	-0.1751	0.8656
BETA 4	100	LIML	-0.2478	1.0812	-1.2292	10.5253
		NL2SLS	-0.3945	2.1528	-0.1468	2.7572
		CF	-0.5104	0.1947	-0.3300	1.0763
	250	LIML	-0.2607	0.8944	-2.5864	17.3566
		NL2SLS	-0.4972	2.6511	-0.2101	2.2255
		CF	-0.4851	0.1254	-0.6105	0.5080
	500	LIML	-0.1359	0.5636	-3.3865	19.3691
		NL2SLS	-0.6071	2.8628	-0.1400	1.1050
		CF	-0.4622	0.0844	-0.3495	0.5054
GAMMA	100	LIML	32.7788	47.8411	0.9720	-0.7483
		NL2SLS	53.1008	57.3207	0.2256	-1.8674
		CF	32.0854	51.5647	0.9352	-0.9275
	250	LIML	18.6800	38.9737	1.9443	2.2271
		NL2SLS	73.3277	56.9283	-0.4611	-1.7160
		CF	12.8033	39.4727	2.1934	3.1185
	500	LIML	8.3095	25.7861	3.5859	12.2833
		NL2SLS	78.6949	55.0424	-0.6614	-1.4755
		CF	0.2037	20.2142	5.1424	26.7201
C	100	LIML	0.0667	0.3737	2.8003	12.2197
		NL2SLS	0.3650	0.8644	1.5611	2.3504
		CF	0.4142	0.6525	1.6285	1.7806
	250	LIML	0.0269	0.2314	7.4027	93.5865
		NL2SLS	0.6844	1.1556	1.4053	1.6060
		CF	0.2508	0.4282	2.6609	6.5832
	500	LIML	0.0063	0.1089	4.1612	43.3998
		NL2SLS	0.9871	1.3393	1.1059	0.4112
		CF	0.1331	0.1657	4.9338	40.7142

Results on Table 3 show now that CF estimator has better performance on standard deviation only for parameter β_1 , β_3 , and β_4 . LIML estimator overcomes the performance of the CF estimator on standard deviation for parameters β_2, γ , and c for almost all simple sizes. LIML estimator presents lower bias for parameters β_1 , β_2 , β_4 , and c while CF estimtor has lower bias for β_3 and γ. Differently from previous examples, CF estimator has not a better performance over the other estimators on standard deviation, particularly the LIML estimator. This may occur because the relationship of the endogenous variable and the instrumental variable is not linear anymore. Results show, however, that either LIML or CF estimators perform better than the NL2SLS estimator. Skewness results for NL2SLS estimator is closer to zero than the other estimators in all settings for all parameters. Kurtosis results for NL2SLS are the closest to zero for all parameters and sample sizes. Its kurtosis remains steady among the different settings, without reaching very large values.

5 Application

In this section we estimate a nonlinear Phillips curve for Brazilian economy using a Smooth Transition Regression (STR) model as in Areosa (2011). The estimated model has the following equation:

$$\pi_t = \beta_o^L + \beta_1^L \pi_{t-1} + \beta_2^L x_{t-1} + \beta_3^L E_t \pi_{t+1} +$$

$$+ (\beta_o^N + \beta_1^N \pi_{t-1} + \beta_2^N x_{t-1} + \beta_3^N E_t \pi_{t+1}) \frac{1}{1+\exp\left(-\gamma\left(\tilde{\sigma}_t^{\pi} - c\right)\right)} .$$

The instruments set is $w_t = [\pi_{t-1}, x_{t-1}, E_{t-1}\pi_t, \tilde{\sigma}_{t-1}^{\pi}, \Delta i_{t-1}]$, where π_t is the inflation rate, x_t is the output gap, $E_t\pi_{t+1}$ is the inflation expectation, $\tilde{\sigma}_t^{\pi}$ is the standard deviation of the inflation expectation and i_t is the interest rate. The set of instruments is obtained from the moment condition derived from the rational expectations hypothesis. Data is obtained from IBGE and Brazilian Central Bank. Inflation rate corresponds to annualized consumers price index (IPCA) and interest rate corresponds to monthly average SELIC rate.

In this application it is assumed that each coordinate of Ψ is centered at (0,0290, 0,3310, 0,0040, 0,8440, 0,0160, -0,2250, 0,8950, 2,3380, 18, 1,0600) with length (8, 8, 8, 8, 8, 8, 8, 8, 80, 8). The parameter α of the control function is centered at 1 with interval length equal to 5. Results are shown in Table 4.

Table 4 shows estimation results for LIML, NL2SLS, and CF estimators.

TABLE 4 – Estimation Results for STR Model

PARAMETER	OLS	LIML	NL2SLS	CF
β_0^L	0.0282	0.057	0.059	0.008
	(11.747)	(0.007)	(11.747)	(11.747)
β_1^L	0.4145	-0.166	-0.203	0.565
	(0.956)	(0.111)	(0.903)	(0.956)
β_2^L	-0.0237	0.002	0.006	-0.302
	(0.477)	(0.084)	(0.429)	(0.477)
β_3^L	0.8803	1.423	1.471	0.293
	(0.209)	(0.587)	(0.207)	(0.209)
β_0^N	0.04	0.09	0.099	0.18
	(4.657)	(0.033)	(3.422)	(2.509)
β_1^N	-0.3854	0.193	0.235	-0.869
	(0.650)	(0.235)	(0.605)	(0.469)
β_2^N	1.685	2.599	2.663	3.397
	(0.180)	(0.459)	(0.155)	(0.088)
β_3^N	1.591	0.201	-0.013	1.933
	(0.191)	(1.013)	(0.184)	(0.129)
γ	17.9891	17.954	58	0.948
	(0.002)	(12.4)	(0)	(0.477)
c	1.4656	1.692	1.745	3.234
	(0.653)	(0.064)	(0.762)	(0.204)
λ				-0.014
				(4.497)

The estimated parameters partially differ from the economic theory presented in Areosa (2011) and the values are similar to the obtained in the paper only for some parameters. Results of Table 4 show that β_1^L is not positive and significant for any estimator while β_3^L is positive and significant for almost all estimators. β_2^N, β_3^N and c, on the other hand, are positive and significant for most estimators. It's interesting to notice that with exception from β_1^L, all estimations from CF estimator are in line with the economic theory. Moreover, for the parameters β_2^N, β_3^C and c, the CF estimator presented the lowest standard errors. However, the estimation of λ is not significant.

6 Conclusion

The aim of this paper is to compare the performance of the two-stage nonlinear least-squares (NL2SLS) estimator, the nonlinear limited information maximum likelihood (LIML) estimator and the Control Function (CF) estimator for specific families of nonlinear functions. For simulated data on examples 1 and 2, the CF estimator performed better than the other estimators for standard deviation in settings with different sample sizes, followed by the LIML estimator. This result was expected due to the linear relation between the endogenous variable and the instrumental variable. On example 3, there was a nonlinear relation between the instrumental variable and the endogenous variable. Therefore, it was not straightforward to determine what estimator has presented the best performance. However, results have shown that both CF estimator and LIML estimator performed better than the NL2SLS estimator. The use of the STR model to estimate a nonlinear Phillips curve for the Brazilian economy showed that estimation results for LIML and CF estimators are in line with the ones previously estimated by Areosa et al. and corroborate results underlined by the economic theory.

7 References

Amemiya, Takeshi. The nonlinear two-stage least-squares estimator. Journal of Econometrics, 2(2):105-110, 1974.

Amemiya, Takeshi. The nonlinear limited-information maximum-likelihood estimator and the modified nonlinear two-stage least-squares estimator. Journal of Econometrics, 3(4):375-386, 1975.

Areosa, W.D., McAleer M., Medeiros, M.C.. Moment-based estimation of smooth transition regression models with endogenous variables. Journal of Econometrics, 165(1):100-111, 2011.

Chamberlain, G.. Asymptotic efficiency in estimation with conditional moment restrictions. Journal of Econometrics, 34(3):305-334, 1987.

Imbens, Guido, and Wooldridge, J.. Control Function and Related Methods. What’s New in Econometrics (lecture for NBER), 2007.

Whitney, W.K.. Efficient instrumental variables estimation of nonlinear models. Econometrica, 58(4):pp. 809-837, 1990.

Appendix

FIGURE 1 – Histograms of Example 1 (β_1)

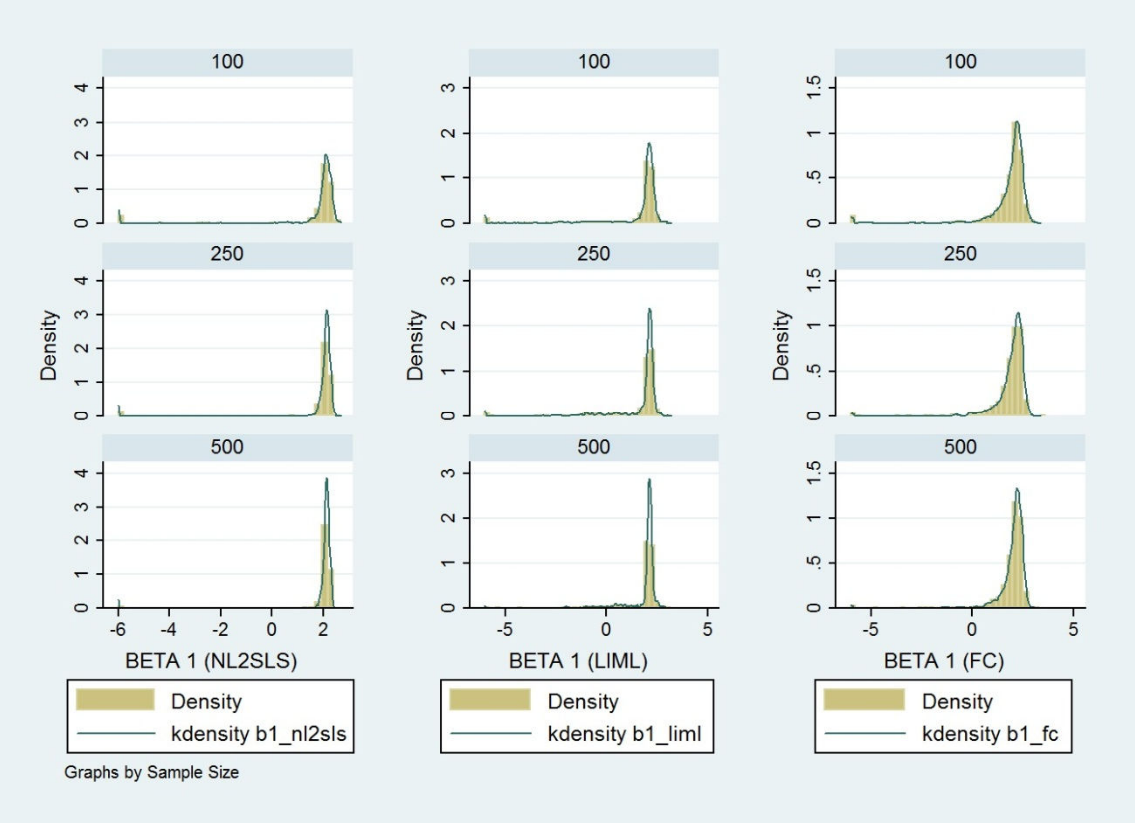

FIGURE 2 – Histograms of Example 1

(β_2)

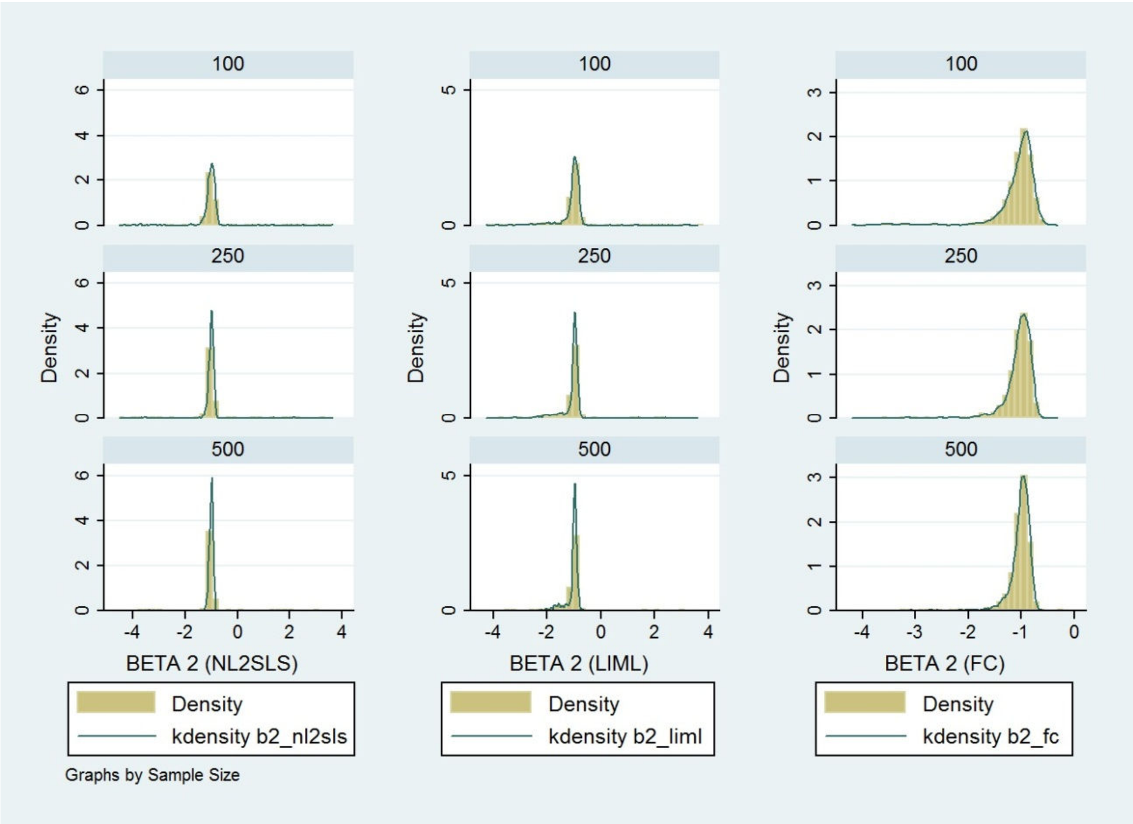

FIGURE 3 – Histograms of Example 1 (β_3)

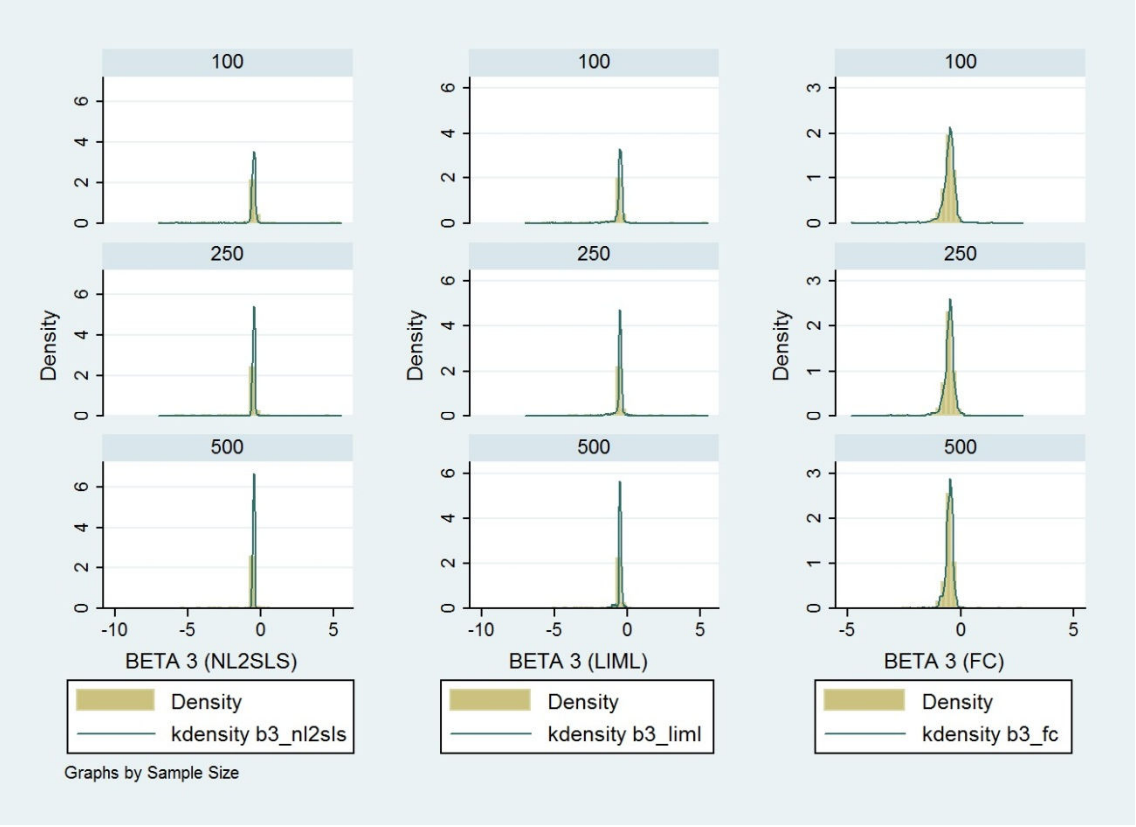

FIGURE 4 – Histograms of Example 2 (β_1)

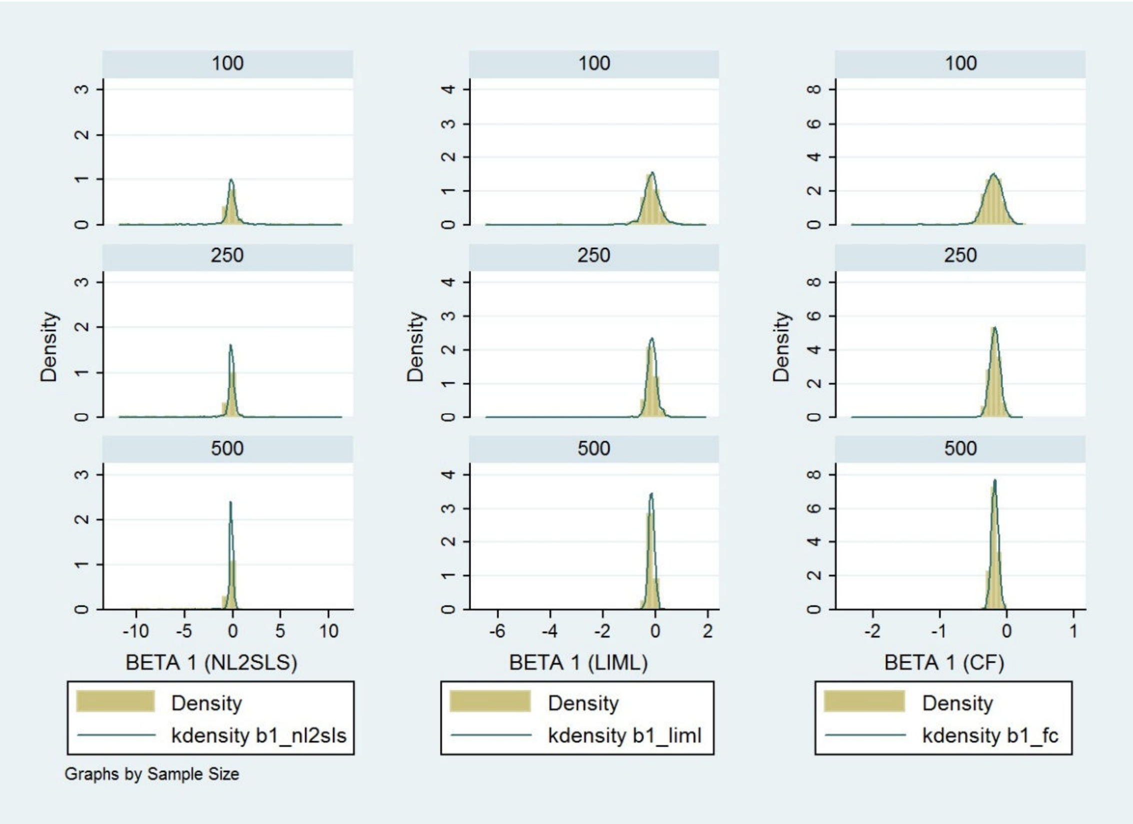

FIGURE 5 – Histograms of Example 2 (β_2)

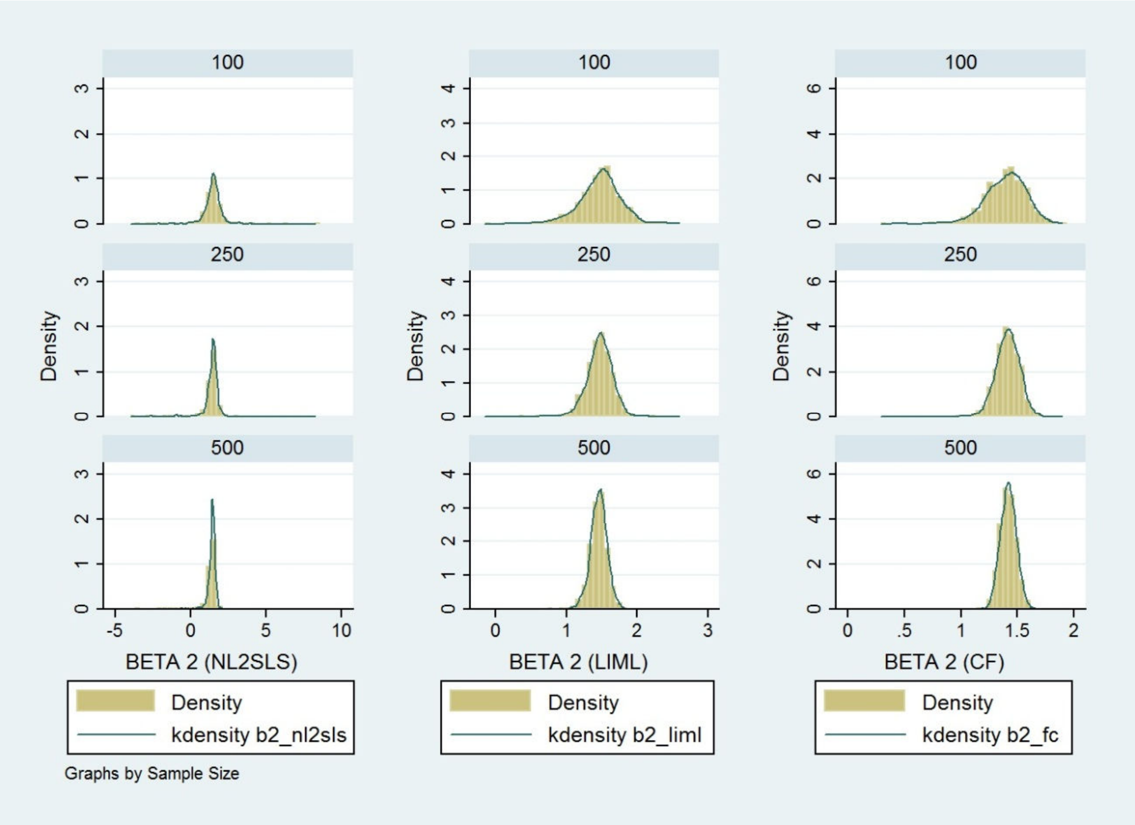

FIGURE 6 – Histograms of Example 2 (β_3)

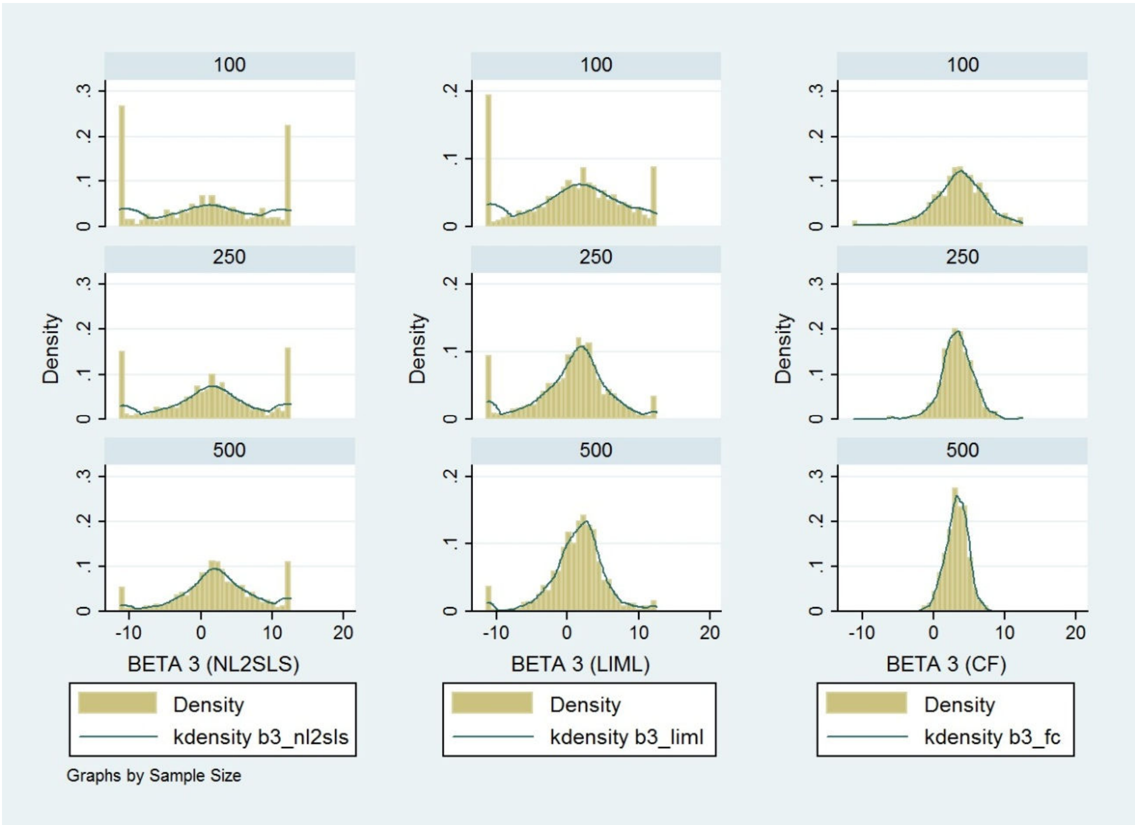

FIGURE 7 – Histograms of Example 2 (β_4)

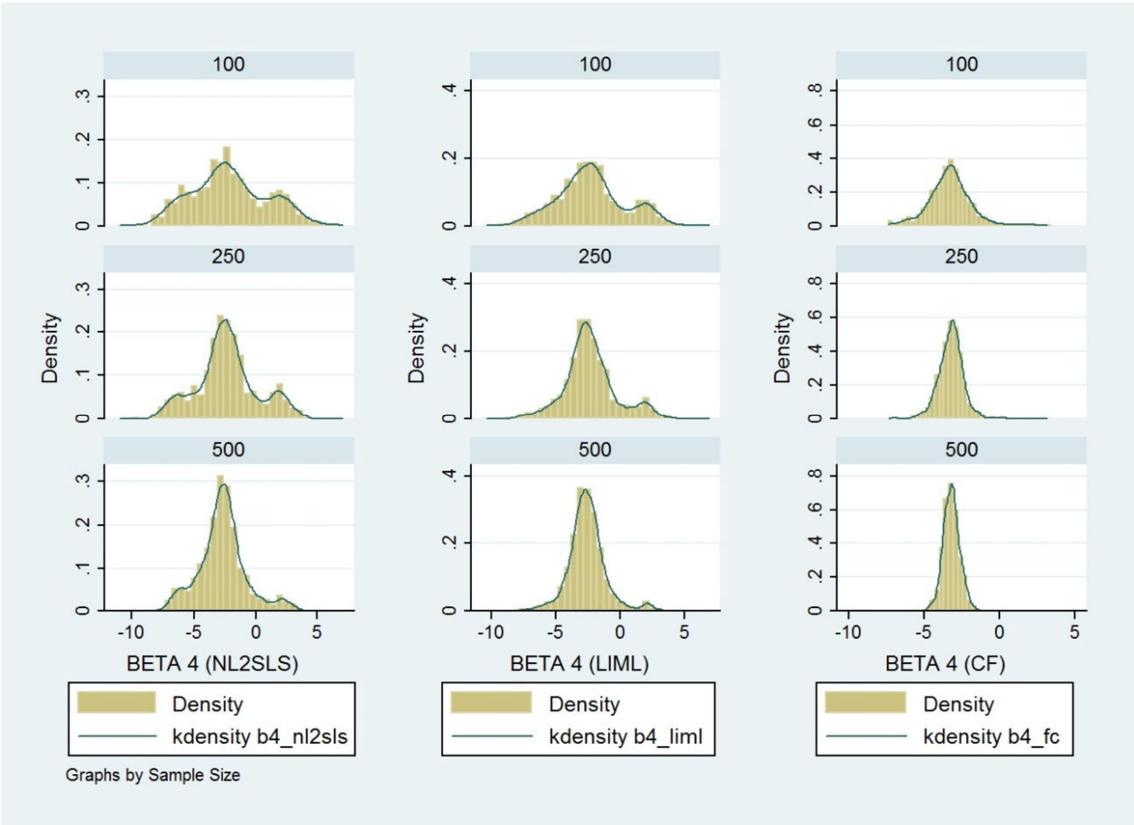

FIGURE 8 – Histograms of Example 2 (γ)

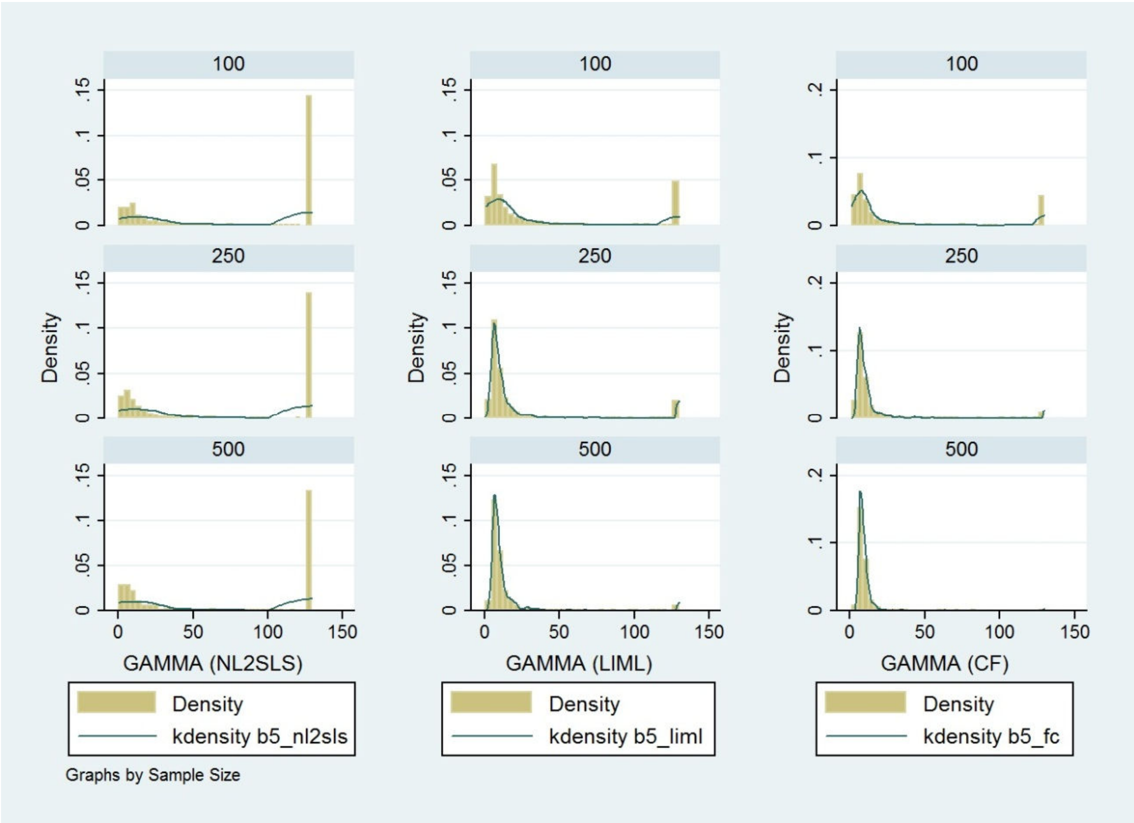

FIGURE 9 – Histograms of Example 2 (c)

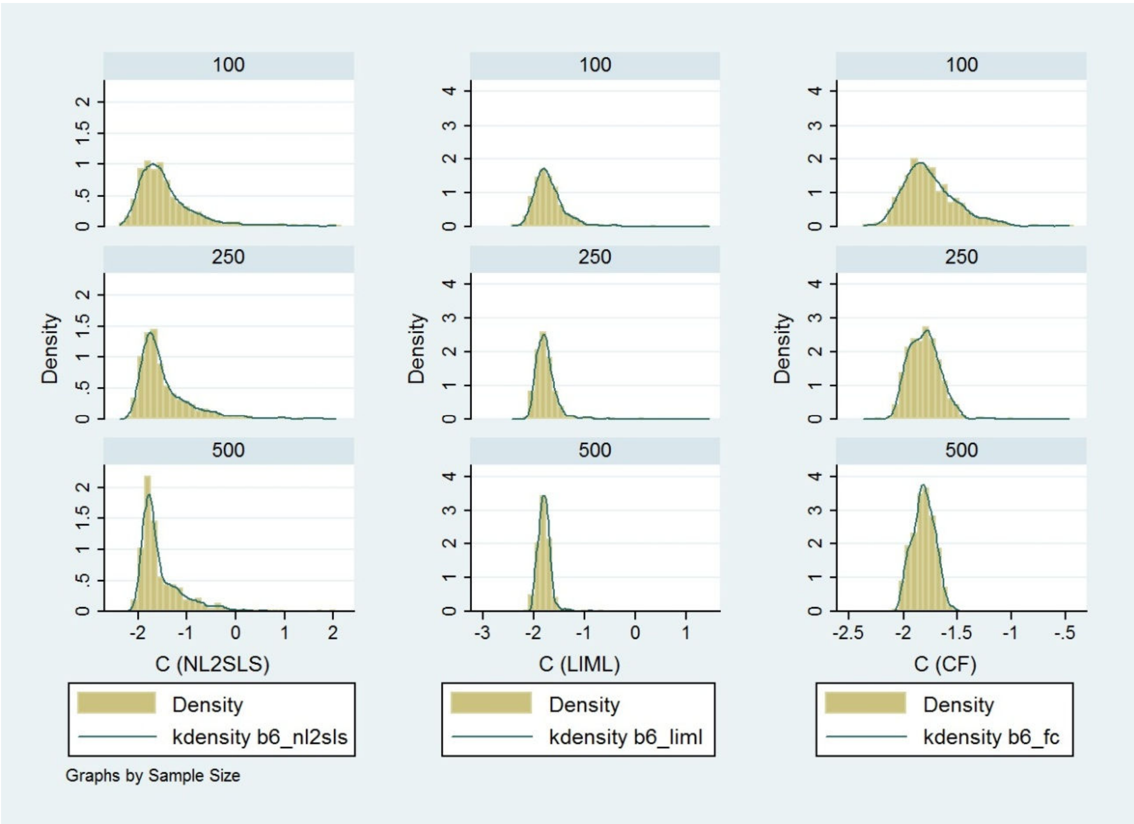

FIGURE 10 – Histograms of Example 3 (β_1)

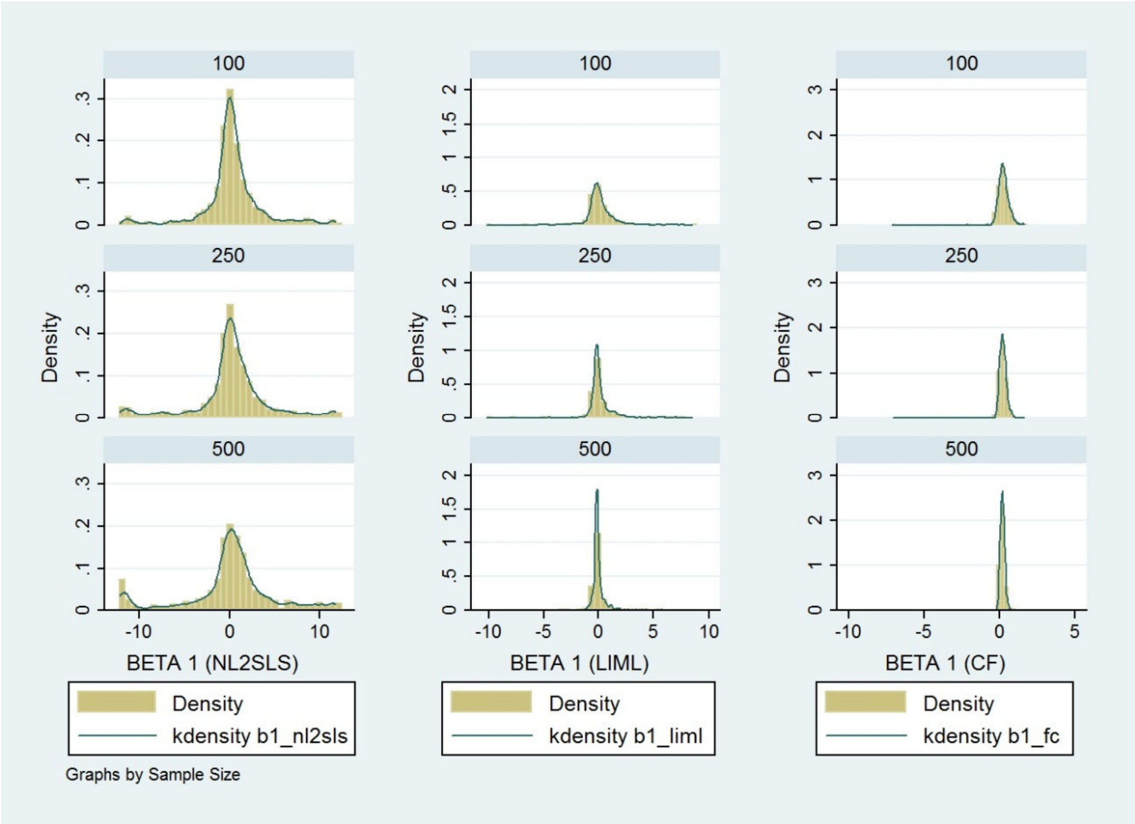

FIGURE 11 – Histograms of Example 3 (β_2)

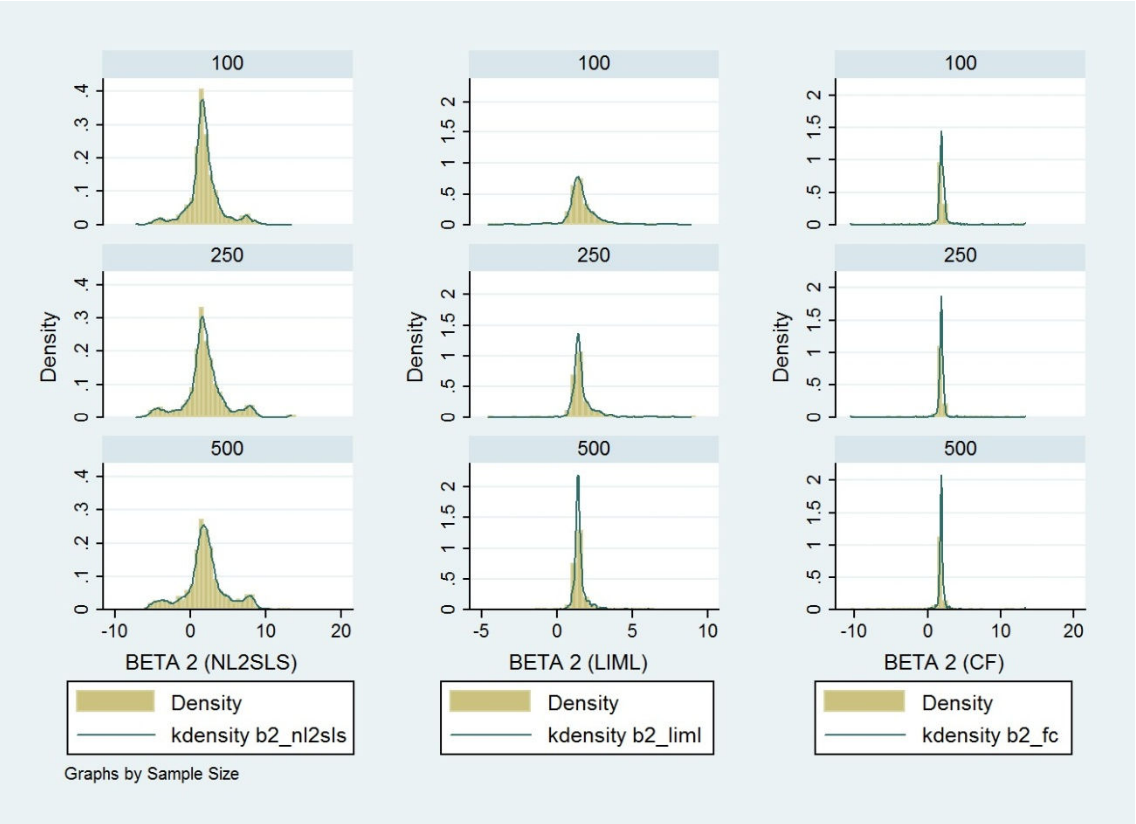

FIGURE 12 – Histograms of Example 3 (β_3)

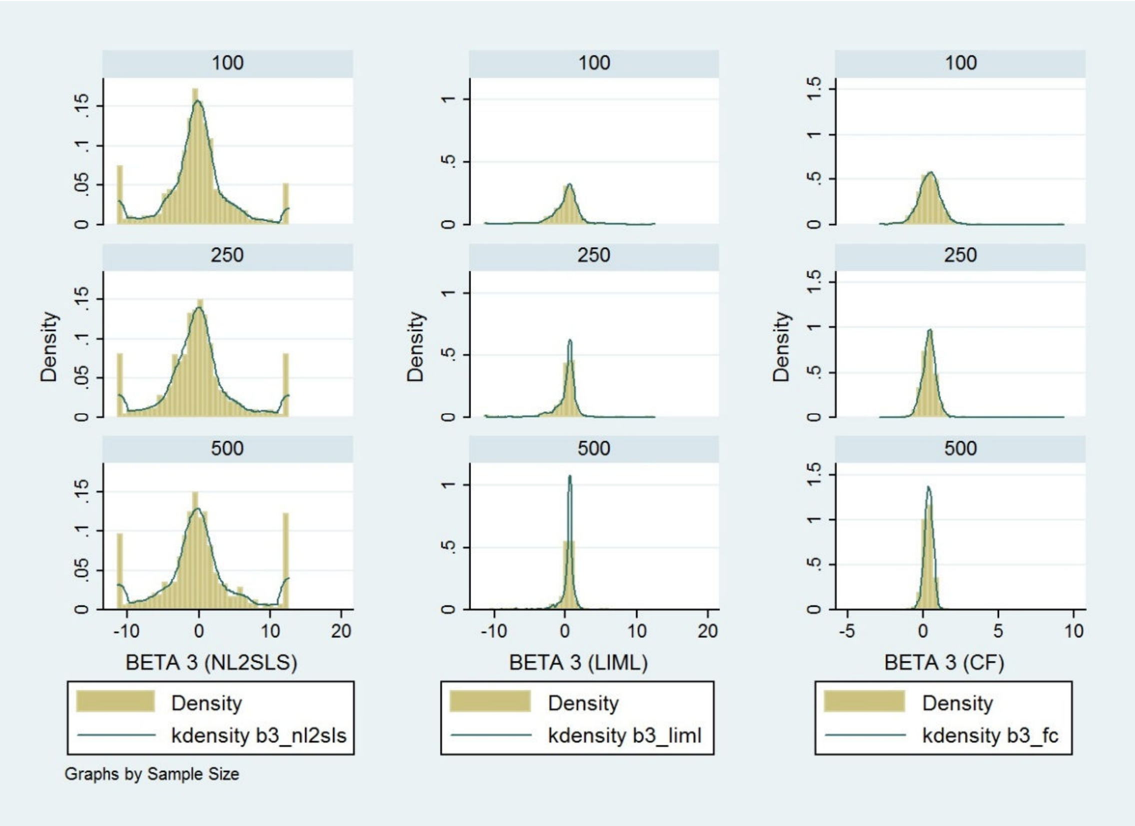

FIGURE 13 – Histograms of Example 4 (β_4)

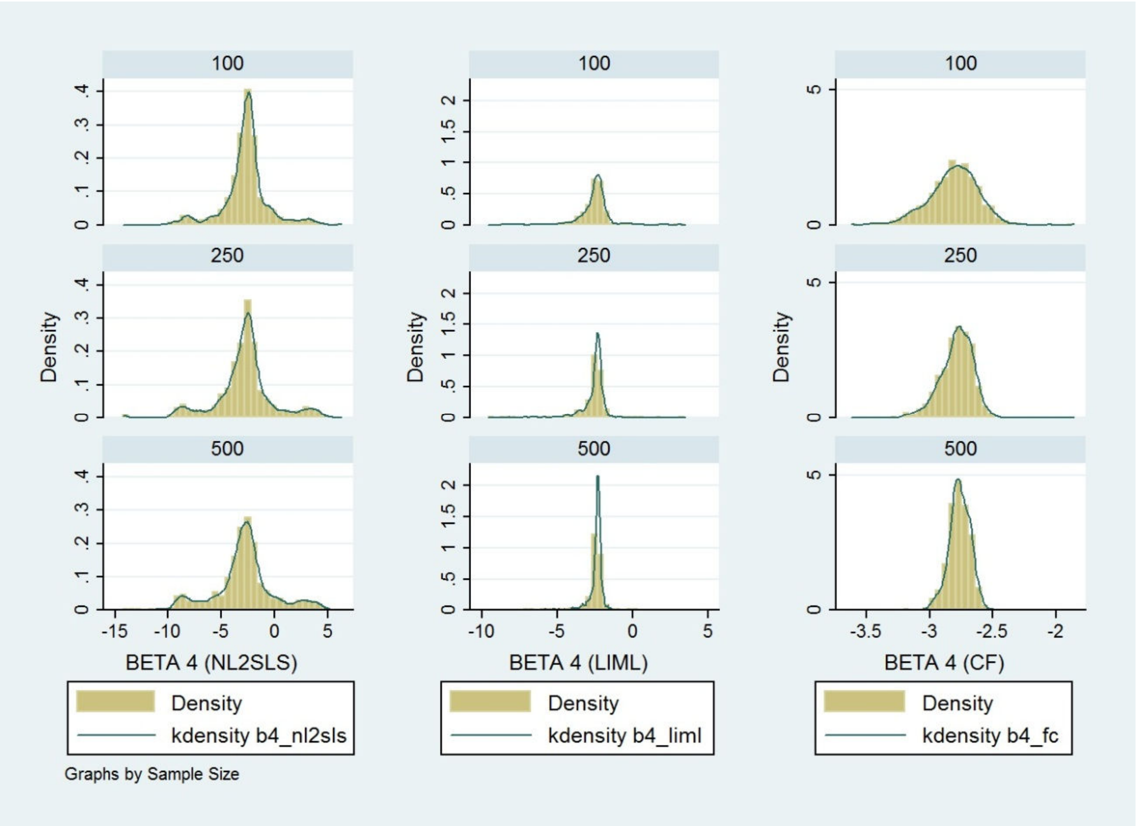

FIGURE 14 – Histograms of Example 3 (γ)

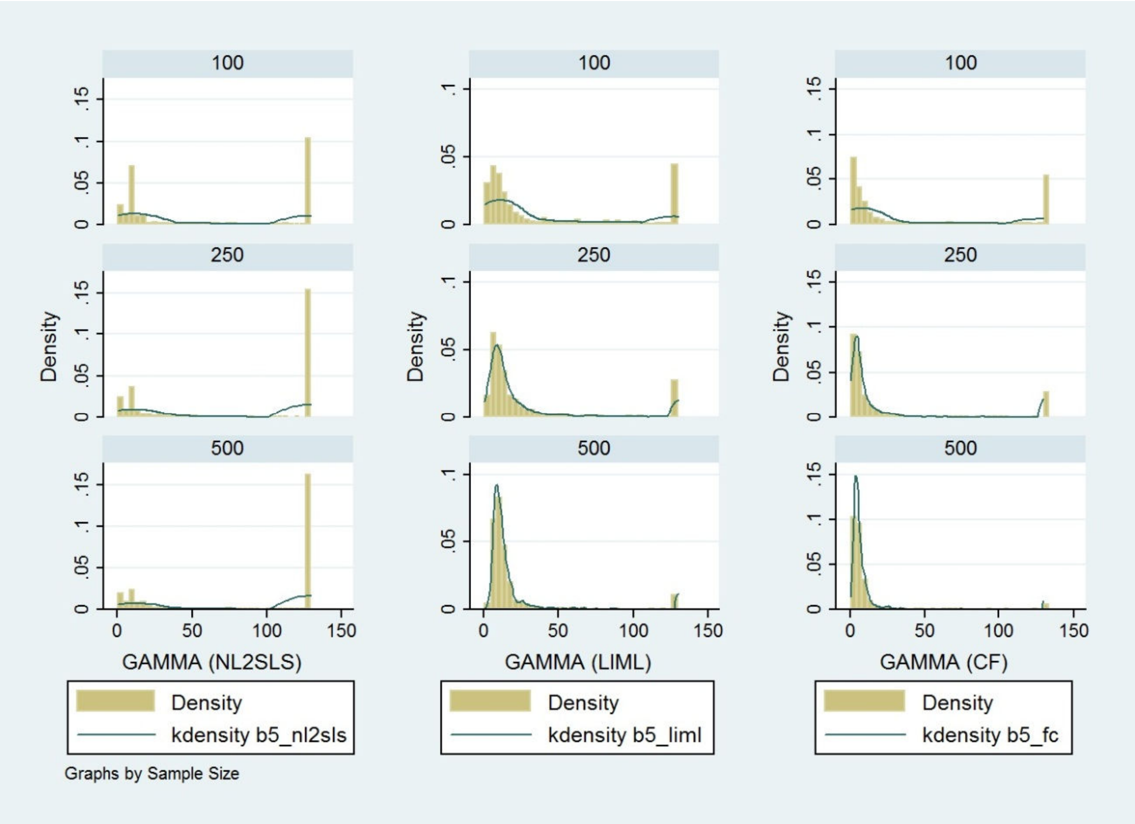

FIGURE 15 – Histograms of Example 3 (c)