

1 Simulation Results

In this section we simulate estimators respectively from Ichimura’s and Klein and Spady’s theoretical models. Then we reflect on properties of these estimates generated via a Monte Carlo experiment with the sample size of 250 and 1000 trials. Results are listed in *Table 1*.

For the generality of comparison, several points are worth noticing in our design of experiment. Firstly, two scenarios are designed regarding to error distributions: standard normal $N(0, 1)$ and joint normal $0.75 \cdot N(-0.5, 1) + 0.25 \cdot N(1.5, 2.5)$. For each scenario, a parametric logit estimator is included as a complement to the two semiparametric single index counterparts. Furthermore, the exogenous vector of variables for all specifications contains two components and both are independently generated from standard normal distributions. Last but not least, our experiment is constructed on the binary choice model, which both enjoys widespread empirical reputation and facilitates us with a efficiency comparison over estimators due to its heteroskedasticity property. Parallel to theoretical conditions for single index, the parameter of the first exogenous component is normalized to 1, while the true value for the second is set to -2 :

$$y_i = I(x_{1i} - 2x_{2i} > \epsilon_i).$$

Our simulation characterizes Ichimura’s nonlinear least squares estimator and Klein and Spady’s maximum likelihood estimator. Again for the convenience of comparison, a fourth order Gaussian kernel function is selected for both cases, in order to achieve the minimal order requirement in Klein and Spady’s model. Meanwhile a trimming function is rigorously defined to guarantee sensible calculation of estimates over each randomly drawn data set.¹ As for bandwidth selection, we preselect two bandwidths respectively for the two models through a likelihood-based cross-validation approach, which is employed in function *npindexbw* of the empirical *NP Package* created by Racine (2014). This optimal bandwidth is investigated by Haerdle, Hall and Ichimura (1991).

In accordance with Ichimura (1993), we chose *grid search* as the method of optimization nested in our estimator simulations. This method performs sufficiently well in finding ‘good’ estimates, as supported by histograms of estimates shown in *Figure 1*. However it may have two considerable deficiencies: on the one hand, as the amount of grids increases, function evaluations grow exponentially; on the other hand, when the true β is unknown, the performance of estimation relies heavily on the way grids are preselected. In our experiment grid search is looped over $(-4, 0)$ with width 0.05. This leads to 81 function evaluations for each data generation, which is already computational costly taking into account an amount of 1000 data generations.

¹Although Ichimura’s and Klein and Spady’s estimators of one individual simulation are extracted from the same trimmed data set, a Klein and Spady ‘floor’ is executed on the latter out of technical reason.

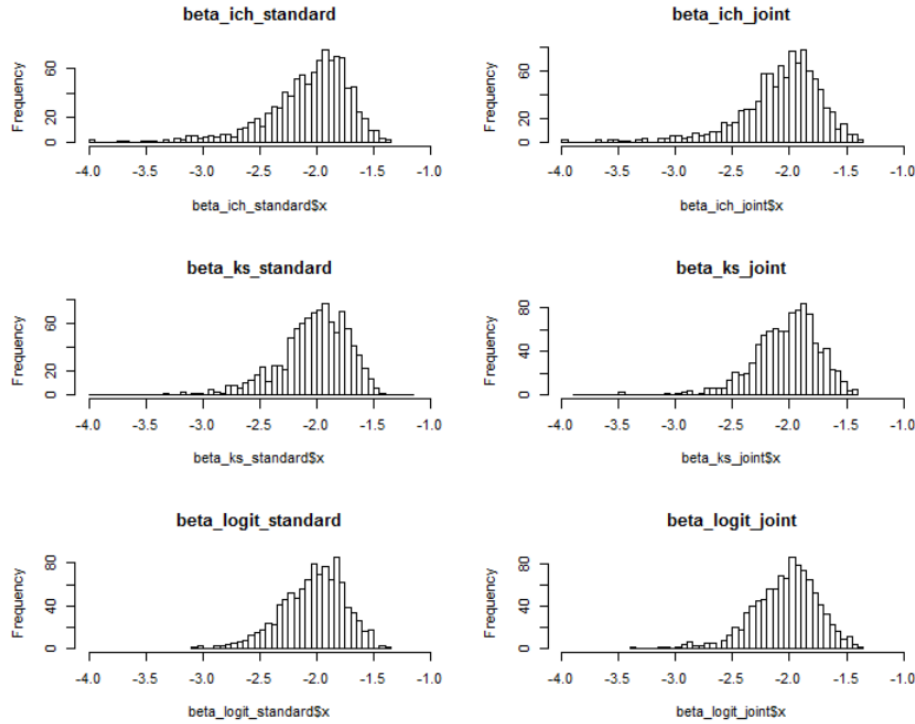
Table 1: Bias and Root Mean Squared Error Comparison

x_1 and x_2 normal - Performance of parametric and semiparametric estimators; a Monte Carlo experiment(250 observations, 1000 trials).

Estimator ¹	Standard Normal		Joint Normal	
	Bias	RMSE	Bias	RMSE
Ichi	-0.045	0.3672	-0.057	0.3665
KS	-0.0095	0.3231	-0.015	0.3125
Logit	-0.030	0.2760	-0.034	0.2813

¹Ichi = Ichimura's method, KS = Klein and Spady's method.

Figure 1: Plot of Estimates



Combining simulation results from the table and the figure it can be distinctly perceived that distributions of all three estimators are skewed to the left, while the biases are bounded below -0.06 . This indicates that the property of asymptotic consistency exists in all three estimators, amongst which the Klein and Spady's estimator converges most rapidly to the true value.

From the perspective of mean squared error there is also no significant difference among estimators in terms of magnitude. Since binary choice models are inherently heteroskedasticity, Ichimura's estimator would require the use of a weight function. However, this is out of the scope of our work, and it can be adapted to simulation in further studies. Nevertheless our results suggest that

this might not be an influencing issue here.

In the end, to guarantee the applicability of our simulation functions, we have employed Racine's *NP Package* on the experiment of standard normal scenario. A comparison between results encouraged us to rely on our original simulation functions. This is listed in the Appendix below.