The Frontier of Estimator in Agriculture Economics: Application for Rice Products in 7 Top Countries in Asia

Chukiat Chaiboonsri¹ and Satawat Wannapan²
¹,² Puey Ungphakorn Excellence Centre of Econometrics
Faculty of Economics, Chiang Mai University, Thailand

¹ chukiat1973@gmail.com ² lionz1988@gmail.com

Abstract. The purpose to econometrically calculate the comparing estimators in parametric estimations is successfully done in this paper. The two powerful estimators such as the James-Stein estimation (JSE) and Maximum Likelihood estimation (MLE) were employed to solve the most efficiently parametric forecasting. Experimentally, the yearly time-series data of rice exporting products during 2006 to 2016 was collected and observed from 7 Asian countries, which are the major exporters in the world, such as Thailand, India, Pakistan, Vietnam, Myanmar, China, and Cambodia. Methodologically, the comparison between the two estimators was based on the linear regression to provide residual terms. Empirically, the results of the estimator comparison implied that the JSE can be substantially replaced the counterpart when doing in the multiple-factors estimation, and error terms estimated were compared for seeking the best estimator to predict trends of rice products. Consequently, it is sensible to introduce the J-S estimator can be one of interesting tools in the modern econometric researches.

1. Introduction
In modern econometrics, the big problems that many quantitative researchers have been trying to computationally solve are the parametric estimation and prediction. Interestingly, to precisely calculate our parameters, is the maximum likelihood estimator really the best estimator for data predictive estimating in the recent moment? Additionally, there has been long time since a hundred years ago that quantitatively econometric papers consider only two dimensions, which refer to one dependent variable related to one independent factor (an unknown-mean variable), others are fixed as constant following assumptions. However, the huge mistake is that we are losing the ability of truth expressions by fixing and assuming as the traditional way. Can we put more dimensions of exogenous variables and make a precise estimation without any assumption simultaneously? This question leads the authors to employ the mathematically experimental study by using the JSE (James-Stein’s estimator) comparing with the maximum likelihood estimator (MLE). Consequently, the empirical result of this paper will be the challenge one that can extend the frontier of econometric estimations and forecasts in the current moment of modern econometrics.

According to the field of agricultural economics, especially rice exporting productions, the trends of this fundamental food have been econometrically predicted for seeking the precise occasions to bargain on the future market and to manage rice supplies in the real market. Obviously, there are many internal and external factors affecting the fluctuations of the trends, and we should not consider them
separately. Consequently, the purpose of the paper is aimed to improve the econometric estimation by comparing two estimators between the J-S estimator and MLE estimator. The hypothesis of the result is to believe that the J-S estimation will be the new tool that can overcome the problem of outliers capturing in the MLE estimation. The paper employs the sampled data of agriculture goods, which are rice products in top 7 countries in Asia. The yearly data is collected during 2006 to 2016. In the term of literature, it is undeniable that mathematical results or even econometrical outcomes are strikingly contrary to generally held belief even though an obviously valid proof is given. James Stein’s estimator, which can be alternatively called “Stein’s idea”, made this issue dramatically in 1961. This is the beginning of the story called “shrinkage estimation”, which deliberate biases are introduce to improve all performance and its modern implementations. Because of the J-S estimator has not been the talk of the town in econometrics yet. This seems the empirical application of J-S testing is still rare in current moment. From literature, for example, the original paper considered into the estimation with quadratic loss by Stein [3], the discussed paper called “Stein’s Paradox in Statistics” by Efron and Morris [2], the paper of Brown [1] regarding the admissibility of invariant estimators in more location parameters, the applied time-series paper with James-Stein estimators by Senda and Taniguchi [5], and the paper of quantum physics using James-Stein estimator by Stander [6].

2. Methodology

2.1 Simulation Method for Estimator Comparison

Before the beginning of real sampled estimations, the comparison between the MLE and JSE is investigated by simulated computations. In this paper, the bootstrap was the computational method relied on the relationship between samples and (unknown) populations by a comparable linkage between the samples at hand and appropriately designed (observable) resamples [7]. Let we compute the bootstrap percentile of the function \( y = f(q_i) \), where \( q_i \) is a motif statistic. The samples, \( q_1, \ldots, q_n \), are assumed to be identically independence distribution (i.i.d) and the function \( P(q_i|I) \), stands for the parametric density of choices [8]. Thus, we can set \( y = f(q_i) \) for each \( I \), the bootstrap percentile is given by

\[
P_{bs}(f,q) = \frac{b_y | y < y |}{n}.
\]

2.2 The James-Stein Estimator

The James–Stein estimator (JSE) is an estimator that provides an estimation of the mean of a multivariate normal distribution, and it dominates the classical Maximum-Likelihood estimator (MLE). The JSE is crucial to stand on their limitations as follows; 1) the JSE is a biased estimator, essentially, it trades bias for risk; 2) the James-Stein estimator is based on a minimax estimator, which is an estimator with the property that its largest risk is no greater than that of any other estimator. If the mean of a multivariate normal distribution is minimax, any estimator that dominates the MLE must also be minimax; 3) the JSE is defined as a minimax, which is an estimator for which no other estimator exists that dominates it. James and Stein showed that the MLE of the mean of a multivariate normal distribution is not an admissible estimator if the dimension exceeds two [4].

According to this paper, the J-S estimator is employed to solve the linear regression problem. The JSE readily extended to deal with a nonidentity covariance matrix, and this leads to two crucial concepts, namely, the “effective dimension” and “shifting the origin”. We conduct \( X \in R^p \) and denote that a normally distributed random vector with mean \( \mu \) as well as positive covariance matrix \( \Omega \) with a normal distribution, \( X \sim N(\mu,\Omega) \). Mathematically, \( P \) is defined as \( P = \left( \Omega^{-1/2} \right) \) and existed since \( \Omega \) is positive and \( PX \sim N(p\mu,1) \). Consequently, the mean equation can be expressed as follows;

\[
\hat{\mu}^{JS} = \left( 1 - \frac{p-2}{X \Omega^{-1} X} \right) X.
\]

(2)
This estimator belongs to the class of spherically symmetric estimator because of \(\hat{\mu} = h(X\Omega^{-1}X)\), where \(h: R \rightarrow R\) which belongs to any Borel-measurable function. Theoretically, if \(tr[\Omega]/\lambda_{\max}[\Omega] \leq 2\) and there is no spherically symmetric estimator that dominates the Maximum Likelihood estimator (MLE), this can be called as the effective dimension which is a real-valued quantity. To aware of the restriction on \(\Omega\) existing and to proof the effectively dimensional condition is crucial for observing the problem for a different angle. Considering a diagonal covariance matrix, \(\Omega = \text{diag}\{\lambda_1,...,\lambda_p\}\). The squared error can be written as \(\|\hat{\mu} - \mu\|^2 = \sum_{j=1}^{p} \hat{\lambda}_j \left( (P \mu_j) - (P \mu_j) \right)^2\), where \(P = \Omega^{-1/2}\). Because the JSE of \(P \mu\) in general does not take a small risk for every individual element, it is evident that if one of the \(\hat{\lambda}_j\) is relatively large, there is no longer possible to compensate for the possibility of introducing a larger MSE into this particular element [4].

The next technique is shifting the origin. It is well known that the risk of the JSE, which decreases as \(\|z\|^2 \rightarrow 0\). Accordingly, the risk of the JSE is decreased by shifting the origin to \(\bar{\mu}\), replacing \(\hat{\mu}\) by \((\hat{\mu} + \bar{\mu})\), and estimating \(\hat{\mu}\). For the parametric estimator by the JSE, the spherically symmetric JSE for the regression parameter is based on the JSE. The differences are that we include the matrix \(D\) in the regression model and the term \(\left(\min[p - 2]^{2}\right)\). Consequently, the James-Stein estimator for \(x\) in the regression is

\[
\hat{x}^{JS} = \left(1 - \sigma^2 \frac{\left(\min[p - 2]^{2}\left[p - 2\right]\right)}{\hat{x}^{ML}^{T} (CRC) \hat{x}^{ML}}\right)^+ \hat{x}^{ML},
\]

where \(R\) is defined as \(R = (DD)^{-1}\). \(C\) is expressed as \(C \in R^{n,p}\), and \((CRC)^{-1}\) is explained as the covariance matrix. \(p\) is the dimension of \(x\), \(n\) the dimension of \(z\), \(\hat{x}^{ML}\) is the effective dimension, which is written as

\[
P^* = \frac{\text{tr}\{CRC\}^{-1}}{\lambda_{\max}\{CRC\}^{-1}}.
\]

If \(n > p\) and \(\sigma^2\) is unknown, \(\sigma^2\) is derived as follows;

\[
\sigma^2 = \frac{\|D^{-1}(z - C \hat{x}^{ML})\|^2}{n - p + 2}.
\]

For the MLE and JSE, they are taken to rely on the following properties:

1. For all \(x \in R^p\), \(J^{JS}(x) \leq J^{ML}(x)\), where \(J^{ML} = E\left[\|\hat{x}^{ML} - x\|^2\right]\) and \(J^{JS} = E\left[\|\hat{x}^{JS} - x\|^2\right]\) are the risk of \(\hat{x}^{ML}\) and \(\hat{x}^{JS}\), respectively.
2. If \(p^* > 2\), \(J^{JS} < J^{ML}\), with an upper bound on \(J^{JS}(x)\) decreasing as \(x\) converging to the origin; \(\|\Omega^{(1/2)}X\|^2 \rightarrow 0\).
3. If \(p^* \leq 2\), \(\hat{x}^{JS} = \hat{x}^{ML}\).

3. Empirical Results

3.1. The Simulation Comparison between MLE Estimator and J-S Estimator

3.1.1 Case 1: changing in observations and factor dimensions. Before the real data goes to the computation, first of all, the simulated approach by the bootstrap sampling is employed to display the difference between the two estimators. Experimentally, 20 and 50 bootstrapping observations are generated with the difference of factor dimensions. For the former calculation, results of the
comparison are shown in the table 1, and it is pictured in the figure 1, which shows the comparison of estimated errors.

Empirically, speaking to the case of 20 samples with 2 factor dimensions, the result evidently expresses that the JSE and MLE represent the similar values of errors. However, the difference is more obvious when the dimensions are increased as multiple factors. The result confirms that the JSE provides the efficient estimator rather than the MLE in the case of 6 factor dimensions. The errors estimated by the JSE are lower than the counterpart, which are 0.4268 and 1.0775, respectively. The estimated outputs are clear when we focus on the figure 1. The mean squared error lose estimated by the JSE is nearly zero when the dimensions are more than six elements. Conversely, the errors provided by the MLE are constant when the dimensions are increased. This states the MLE cannot efficiently capture robustness in the situation contained multiple variables. To proof that the JSE is better than the MLE surely, the second simulation analysis is employed to complete this issue. 50 bootstrapped observations are generated. The results are represented in the table 2.

Table 1. The simulation of 20 observations with 2 and 6 factor dimensions

| Factor dimensions (p) | James-Stein (Square error) | MLE (Square error) |
|-----------------------|----------------------------|-------------------|
| p = 2, theta_norm = 1 | 1.1281 (0.1551)            | 1.1281 (0.1551)   |
| p = 6, theta_norm = 1 | 0.4268 (0.0637)            | 1.0775 (0.0959)   |

Table 2. The simulation of 50 observations with 2 and 6 factor dimensions

| Factor dimensions (p) | James-Stein (Square error) | MLE (Square error) |
|-----------------------|----------------------------|-------------------|
| p = 2, theta_norm = 1 | 1.1281 (0.1551)            | 1.1281 (0.1551)   |
| p = 6, theta_norm = 1 | 0.4268 (0.0637)            | 1.0775 (0.0959)   |

Figure 1. The simulation of 20 observations with 2 and 6 factor dimensions by the JSE and MLE
Evidently, from the table 2, the errors estimated from the J-S estimation are lower than the MLE when the factor dimensions are increased, even though it is not a huge increment in simulated observations. This strongly confirms that the J-S estimator is quite efficient to substitute the MLE, especially in the case of multi-dimensions analyses. Furthermore, the results are crystal clear when we consider into the figure 2.

3.1.2 Case 2: changing only in factor dimensions. This section is to employ 100 bootstrapping observations with differences of factor dimensions during 1 to 16 elements. Empirically, the results numerically detail that error terms estimated from the JSE are obviously lower that the counterpart when the dimensions are raised continuously. This can imply that the JSE is better than the MLE in the case of multiple independent variables, which is more realistic and closed to the true value in parametric estimations. These results are graphically represented as the figure 3.

3.2 Applications for Measuring Rice Exporting in 7 Top Countries in Asia
3.2.1 Data description of rice exporting in 7 top countries in Asia by the James-Stein Estimator. Descriptively, the symmetrical distributions which regard to the divided proportion between the estimated mean and median are pictured as the figure 4. Technically, the median-mean distribution should close to one theoretically, which explains the estimator is perfectly precise, and it can capture outliners completely. Empirically, the symmetric distributions estimated by the JSE are mostly higher...
than the counterpart provided by the MLE, which are the consumptions in Singapore, Malaysia, Indonesia, and Philippines from the J-S estimation, except for Thailand’s consumption.

![Figure 4](Image)

**Figure 4.** The presentation of estimated mean and median from the JSE estimation

3.2.2 **The Comparison measurement between the MLE and J-S estimator.** In the final section, the simulation data of rice exporting indexes in 7 top countries in Asia was processed in the linear regression, which is employed the MLE and JSE to provide error terms. In this case, the estimated residuals were used to indicate the precise estimation. The results displayed in the figure 5 expresses that the J-S estimator can be more fitted for the estimation rather than the MLE, which are 4 of 7 cases of parametric estimations such as Vietnam, Thailand, Pakistan, and India, respectively. On the other hand, the MLE can be fitted for the rice exporting products in Myanmar, China, and Cambodia, respectively.

![Figure 5](Image)

**Figure 5.** The comparison of estimated errors from the JSE and MLE

4. **Conclusion**
The clarification in the comparison of statistical estimators which are the James-Stein estimator (JSE) and Maximum Likelihood estimator (MLE) to apply in the agricultural indicator was successfully finished in this paper. Experimentally, it is obvious that the univariate factor, which is the data of rice exporting products in 7 Asian countries, is the main issue to do the parametric estimation. However, the problem leaded to produce this paper is the precision of the estimation, especially when the dimensions which can be referred to multiple-factor analyses are increased more than two elements. Consequently, the comparison between two estimators will provide the better way for improving the precision of econometrical forecasting.

Empirically, the results of the estimator comparison were clear that the JSE is the better estimator than the counterpart when doing in the multiple-factors estimation. Additionally, the simulation approach called *bootstrapping* was employed to extend the investigation in time-series data robustness. The results were obvious that the JSE can efficiently capture outliers in simulated data with high numbers of factor dimensions rather than the MLE. In the application for agricultural data (rice exporting products in top Asian countries), it is evident that the efficiency of parametric forecasting between the JSE and MLE based on the univariated linear regression can be strongly
substitutable. This implies that the JSE can perfectly replace the MLE and completely overcome the problems of robustness, multicolinearity, even restricted linear assumptions. Accordingly, the JSE is potentially the frontier of econometric estimations and forecasts in the current moment of modern econometrics.

5. References
[1] Brown, L. D. (1966), "On the admissibility of invariant estimators of one or more location parameters", Annals of Mathematical Statistics, 37 (5): 1087–1136, doi:10.1214/aoms/1177699259.
[2] Efron, B. and Morris, C. (1977), “Stein’s Paradox in Statistics.” Scientific American. 236 (5): 119–127.
[3] James, W.; Stein, C. (1961), "Estimation with quadratic loss", Proc. Fourth Berkeley Symp. Math. Statist. Prob., 1, pp. 361–379.
[4] Manton, J. H., Krisnamurthy, V., and Vincent Poor, H. (1998). James–Stein state filtering algorithms. IEEE TRANSACTIONS ON SIGNAL PROCESSING, 46(9), pp. 2431-2447.
[5] Senda, M. and Taniguchi, M. (2006). James–Stein estimators for time series regression models. Journal of Multivariate Analysis, 97(9), pp. 1984-1996.
[6] Stander, M. (2017), Using Stein's estimator to correct the bound on the entropic uncertainty principle for more than two measurements. March 2, 2017. Available at: https://arxiv.org/abs/1702.02440.
[7] Vinod, H.D. (2013). Maximum entropy bootstrap simulations for variance estimation. Available at SSRN: https://ssrn.com/abstract=2295723.
[8] O’Neil, P.K. and Erill, I. (2016). Parametric bootstrapping for biological sequence motifs BMC. Bioinformatics, 17.