



非対称な損失関数に基づく代理変数の統計的性質

車井, 浩子

(Citation)

国民経済雑誌, 191(1):31-39

(Issue Date)

2005-01

(Resource Type)

departmental bulletin paper

(Version)

Version of Record

(JaLCD0I)

<https://doi.org/10.24546/00055975>

(URL)

<https://hdl.handle.net/20.500.14094/00055975>



非対称な損失関数に基づく 代理変数の統計的性質

車 井 浩 子

回帰分析において必要な変数が観測不可能な場合、その変数を省くか、代理変数を用いるかという問題が生じる。本論文では、このような代理変数の問題が生じている場合の推定量の統計的性質について、線形損失関数を用いて、非対称な評価基準に基づくリスク評価を行う。

キーワード 代理変数, MSE, 線形損失関数

1 は じ め に

経済データを用いて回帰分析を行う場合、必要な変数が観測できない場合がある。このような場合、その変数を省いてしまうのか、代わりに代理変数を用いるのかという問題が生じる。観測不可能な変数をどう扱うかによって、得られる推定量の評価が異なることがこれまでの多くの研究から明らかになっている。McCallum (1972), Wickens (1972) では、代理変数を用いる方が変数を省いてしまうよりも推定量のバイアスが小さくなることが指摘されている。しかし、Frost (1979) では、ある条件の下では代理変数を用いることによりバイアスと分散が大きくなってしまふことが明らかにされている。Aigner (1974) では、平均平方誤差 (Mean Squared Error; 以下, MSE) の点においては代理変数を用いるほうが望ましいことが示されている。¹⁾

推定量の評価をする場合、バイアスや分散に注目することも多くあるが、通常は評価基準として2次の損失関数を用いて評価が行われることが多い。しかし、2次の損失関数では過大推定と過小推定を同等に評価してしまうため、これらの重大性が異なる場合には適切な評価基準とはいえない。正の推定誤差と負の推定誤差を非対称的に評価する基準として、これまでに LINEX 損失関数や線形損失関数といった評価基準が提唱されている。²⁾

本論文では、代理変数の問題について、非対称な評価基準である線形損失関数を用いた評価を行う。第2章では本論文で扱うモデルについて説明を行い、第3章では線形損失関数に基づく危険関数を導出する。さらに、第4章では数値計算による結果を考察し、第5章で結

論を述べる。

2 モデルと推定量

以下では次式で表されるモデルを用いる。

$$Y = X\beta + Z\gamma + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2 I_n) \quad (1)$$

Y は観測値の従属変数で $n \times 1$ ベクトル, X は観測値の独立変数で $n \times 1$ ベクトル, Z は観測不可能な独立変数で $n \times 1$ ベクトル, ε は正規分布に従う誤差項で $n \times 1$ ベクトルである。ここで, Z は観測不可能であるため, 次のような 2 つのモデルが考えられる。まず 1 つ目は, Z を省いてしまうモデルであり, 次式で表される。

$$Y = X\xi + u. \quad (2)$$

2 つ目のモデルは Z の代わりに代理変数 P を用いるモデルであり, 次式で表される。

$$Y = X\delta + P\omega + u. \quad (3)$$

なお, 独立変数 X, Z, P はすべて非確率変数であり Y とは独立であると仮定する。

ここで, 表記を簡略化するために (1), (3) 式を次式のように表す。

$$Y = A^* \theta^* + \varepsilon, \quad (4)$$

$$Y = A\theta + \varepsilon, \quad (5)$$

ただし, $A^* = [X \ Z]$, $\theta^* = [\beta \ \gamma]'$, $A = [X \ P]$, $\theta = [\delta \ \omega]'$ である。(5) 式から得られる θ の最小 2 乗推定量 $\hat{\theta}$ は次式で表される。

$$\hat{\theta} = [\hat{\delta} \ \hat{\omega}]' = (A'A)^{-1} A'Y. \quad (6)$$

また, (2) 式から得られる ξ の最小 2 乗推定量を $\hat{\xi}$ で表すと,

$$\hat{\xi} = (X'X)^{-1} X'Y, \quad (7)$$

となる。

3 線形損失関数に基づく危険関数

線形損失関数 (linear loss function) は次式で定義される。

$$L(\hat{\beta}, \beta) = C_1(\hat{\beta} - \beta) \quad \text{if } \hat{\beta} \geq \beta,$$

$$=-C_2(\hat{\beta}-\beta) \quad \text{if } \hat{\beta}<\beta. \quad (8)$$

ただし, $C_1 \geq 0, C_2 \geq 0$

ここで $\hat{\beta}$ は β の推定量を表す。 C_1, C_2 は損失関数の形を決めるスカラーのパラメータである。例えば, 正の推定誤差のほうが負の推定誤差よりも重要だと考えられる場合には, C_1 の値を C_2 よりも大きくとることで推定誤差を非対称的に扱うことができる。

代理変数 Z を用いた場合の δ の最小 2 乗推定量 $\hat{\delta}$ について, 線形損失関数は次式で表される。

$$L(\hat{\delta}, \beta) = C_1(\hat{\delta}-\beta)I[\hat{\delta} \geq \beta] - C_2(\hat{\delta}-\beta)I[\hat{\delta} < \beta]. \quad (9)$$

ここで $I[.]$ は, $[.]$ 内の不等号の関係が満たされれば 1, そうでなければ 0 をとるような定義関数を表す。この損失関数から次式で表される危険関数が導出される。³⁾

$$\begin{aligned} R(\hat{\delta}/\sigma) &= (C_1 + C_2)[P'P/(2\pi D)]^{1/2} \exp(-[D/(2P'P)](\lambda_\gamma/\sigma)^2) \\ &\quad - (\lambda_\gamma/\sigma)(C_1 + C_2)F(-[D/(P'P)]^{1/2}(\lambda_\gamma/\sigma)) + C_1(\lambda_\gamma/\sigma). \end{aligned} \quad (10)$$

ここで,

$$\begin{aligned} D &= (P'P)(X'X) - (P'X)^2 \\ \lambda_\gamma &= [(P'P)X' - (X'P)P']Z\gamma/D, \end{aligned}$$

であり, $F(.)$ は標準正規分布の分布関数を表す。同様に, 変数 Z を省いたモデル(5)式における ξ の最小 2 乗推定量 $\hat{\xi}$ の損失関数, 及び危険関数は次式で表される。

$$L(\hat{\xi}, \beta) = C_1(\hat{\xi}-\beta)I[\hat{\xi} \geq \beta] - C_2(\hat{\xi}-\beta)I[\hat{\xi} < \beta], \quad (11)$$

$$\begin{aligned} R(\hat{\xi}/\sigma) &= (C_1 + C_2)[1/(2\pi X'X)]^{1/2} \exp(-(X'X/2)(\mu_\gamma/\sigma)^2) \\ &\quad - (\mu_\gamma/\sigma)(C_1 + C_2)F(-(X'X)^{1/2}(\mu_\gamma/\sigma)) + C_1(\mu_\gamma/\sigma). \end{aligned} \quad (12)$$

ただし,

$$\mu_\gamma = (X'X)^{-1}X'Z\gamma,$$

である。

ここで, 以下の計算のために, 独立変数 X, Z, P はそれぞれの平均からの偏差であるとし, $X'X = Z'Z = P'P = 1$ と仮定する。すると, 危険関数(10), (12)に含まれるパラメータ $\lambda_\gamma, \mu_\gamma$ を次のように表すことができる。

$$\lambda_\gamma = [(R_{XZ} - R_{PX}R_{PZ})/(1 - R_{PX}^2)]\gamma, \quad (13)$$

$$\mu_\gamma = R_{XZ} \gamma. \quad (14)$$

ここで、 R_{ij} は変数 i, j 間の相関係数を表す。さらに、 P, Z 間の相関係数 R_{PZ} は次のように書き表すことができる。

$$R_{PZ} = ((1 - R_{XZ}^2)(1 - R_{XP}^2))^{1/2} R_{PZ.X} + R_{XZ} R_{XP}. \quad (15)$$

ここで、 $R_{PZ.X}$ は X を固定したときの Z, P 間の偏相関係数である。

4 数値計算による結果

(10), (12) 式の危険関数を用いて、以下では数値計算によるリスク評価を行う。数値計算においては、代理変数を用いた場合のリスクと変数を省いた場合のリスクの相対リスク、 $R(\hat{\delta})/R(\hat{\xi})$ を求める。したがって、得られた値が 1 よりも大きければ、観測不可能な変数 Z に対して代理変数 P を用いるより変数を省くほうが望ましいと判断され、相対リスクが 1 よりも小さければ代理変数を用いるほうが望ましいと判断される。損失関数におけるパラメータ C_1, C_2 に関しては、 $C_1 = 1.0$ とし、 C_2 については 0.1, 1.0, 10 という値を用いた。 $C_2 = 0.1$ とした場合には過大推定を重要視する損失関数となり、 $C_2 = 10$ の場合は過小推定を重要視する損失関数ということになる。危険関数に含まれるその他パラメータに関しては様々な値を用い、相関係数 $R_{XZ}, R_{XP}, R_{PZ.X}$ に用いる値により代理変数の質の影響を見る。以下では、 $C_2 = 0.1$ のケースと $C_2 = 1.0$ のケースの結果を用いてリスク評価を行う。

代理変数が元の変数 Z と相関が低く、良い代理変数ではない場合 ($R_{PZ.X} = 0.1$) の結果が図 1, 2 に示されている。図では、相関係数 R_{XZ}, R_{XP} を、それぞれ XZ, XP と表している。まず、過大推定のほうが重視される $C_2 = 0.1$ のケースにおいて、 X, P 間の相関が高いときは ($R_{XP} = 0.9$)、相対リスクは歪みを持つ形状をとっていることがわかる (図 1)。このとき、 X, Z 間の相関も高ければ ($R_{XZ} = R_{XP} = 0.9$)、 $-5 < \gamma < 3$ 付近では相対リスクの値が大きく、代理変数を用いるよりも変数を省くほうが望ましい。 X, Z 間の相関が低い場合も、変数を省くほうが良いと思われる範囲が $\gamma < 6$ の範囲で広く存在する。 $C_2 = 1.0$ のケースでも同様の傾向が見られるが、代理変数を用いる方が望ましいと判断される範囲が $C_2 = 0.1$ の場合とは異なっており、より広い (図 2)。一方、 X, P 間の相関が低いとき、 X, Z 間の相関も低ければ ($R_{XZ} = R_{XP} = 0.1$)、 $C_2 = 1.0$ のケースにおいては、 $\gamma = 0$ 付近の狭い範囲以外では代理変数を用いるほうが望ましいと思われる。これに対し、 $C_2 = 0.1$ のケースにおいては、 $\gamma < 0$ の範囲では変数を省くほうがよいと思われる範囲が広く存在する。しかし、 X, Z 間の相関が高い $R_{XZ} = 0.9, R_{XP} = 0.1$ の場合は、 $C_2 = 0.1, C_2 = 1.0$ のケースともにリスク差はほとんどない。

次に、代理変数が元の変数 Z と相関が高く、望ましい代理変数である場合 ($R_{PZ.X} = 0.9$)

の結果が図 3, 4 に示されている。まず, X, P 間の相関が高く過大推定のほうが重視されるケース ($C_2=0.1$) においては, $R_{PZ.X}=0.1$ の場合よりもより歪みの強い形状をとる (図 3)。 X, P 間の相関が高い場合, X, Z 間の相関が低ければ ($R_{XZ}=0.1, R_{XP}=0.9$), 非常に広い範囲で変数を省くほうが望ましいと思われる。しかし, X, Z 間の相関も高い場合は ($R_{XZ}=0.9, R_{XP}=0.9$), $\gamma \geq 1, \gamma \leq -8$ の範囲で逆に代理変数を用いるほうがリスクが小さい。過大推定と過小推定が同等に評価される場合は ($C_2=1$), $R_{XP}=0.9, R_{XZ}=0.1$ ならば, 変数を省いてしまう方が代理変数を用いるよりも一様に望ましい (図 4)。 $R_{XZ}=0.1$ の場合は, 変数 Z を省くことによる β の推定への影響はかなり小さいと思われるため, X と相関が高いような代理変数を用いることは望ましくないと解釈でき, このような結果が得られたものと思われる。 X, P 間の相関が低いケースについては, $R_{PZ.X}=0.1$ の場合とほぼ同様の結果が見られた。

なお, 損失関数の形が過小推定を重視するケース ($C_2=1$) においては, いずれの場合も $C_2=0.1$ の場合と, $\gamma=0$ に関して左右対称な結果が得られた。

5 結 論

数値計算の結果から, 代理変数に関するリスク評価が損失関数の形に依存することが明らかになった。過大推定が大きな問題となる場合には, $\gamma \leq 0$ の範囲では代理変数を用いずに変数を省くことが望ましく, γ が正の値をとる場合は広い範囲で代理変数を用いるほうが望ましい。しかし, 過大推定と過小推定を同等に評価する場合には, 代理変数を用いるかどうかは相関係数の値に依存しており, R_{XP} が R_{XZ} に近ければ, 広い範囲において代理変数を用いるほうがリスクが小さく, 代理変数が良い代理変数であるほど ($R_{ZP.X}=0.9$) この傾向は強く見られた。ただし, 過大推定が重視されるケースにおいては代理変数が良い変数であるほどリスクは大きくなり, 代理変数を用いないほうが良いと思われる。

しかし, 変数 P は観測不可能な変数 Z の代理変数であるので, 相関係数 R_{XP} と R_{XZ} はさほど大きく異なる値をとることはないであろう。この点を考慮すると, 過大推定が重視される場合であっても, γ が正の範囲では代理変数を用いるほうが望ましく, 逆に他の範囲では代理変数を用いないほうが良いということが数値計算の結果からわかる。以上の結果から, 代理変数が観測不可能な変数に対して望ましい変数であっても, 用いる場合には注意が必要であると言えるであろう。さらに, (X, Z), (X, P) 間の相関がそれぞれ高い場合には多重共線性が生じている可能性があり, 今後この点を考慮した分析が必要であると思われる。

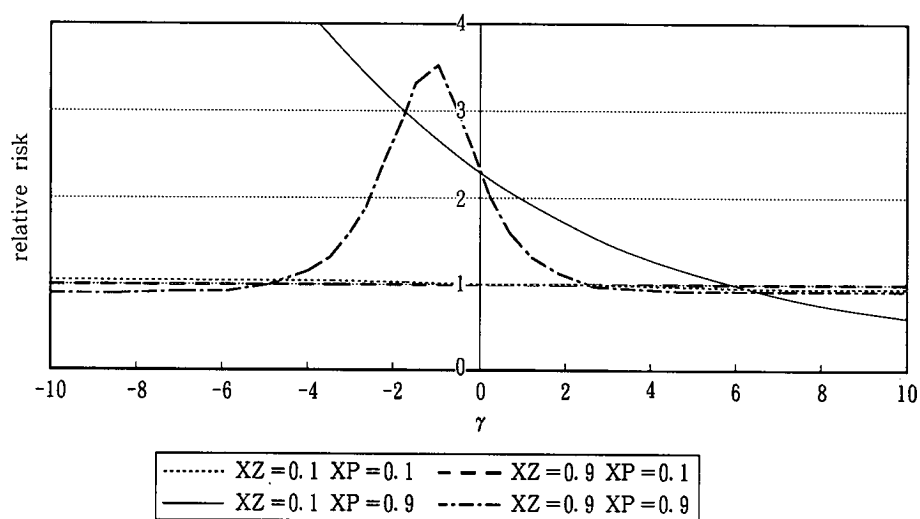
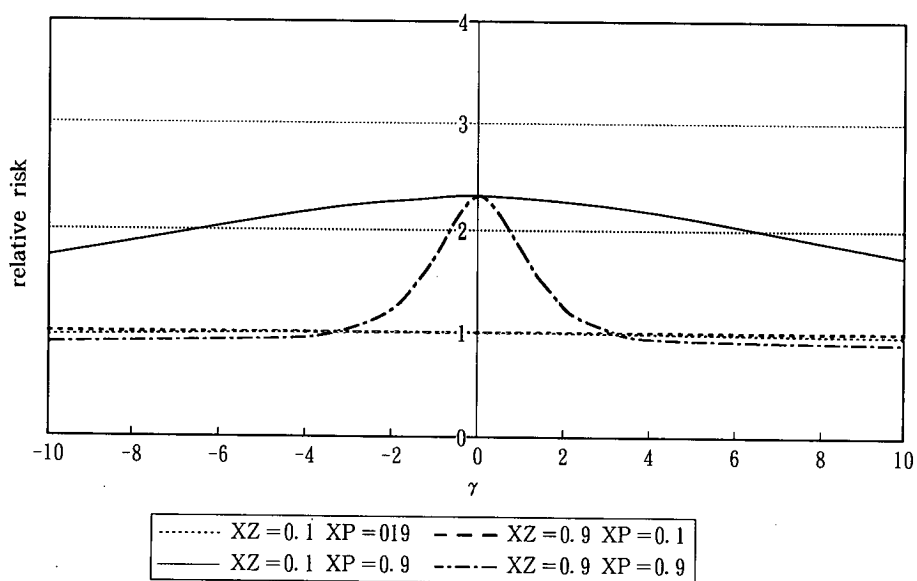
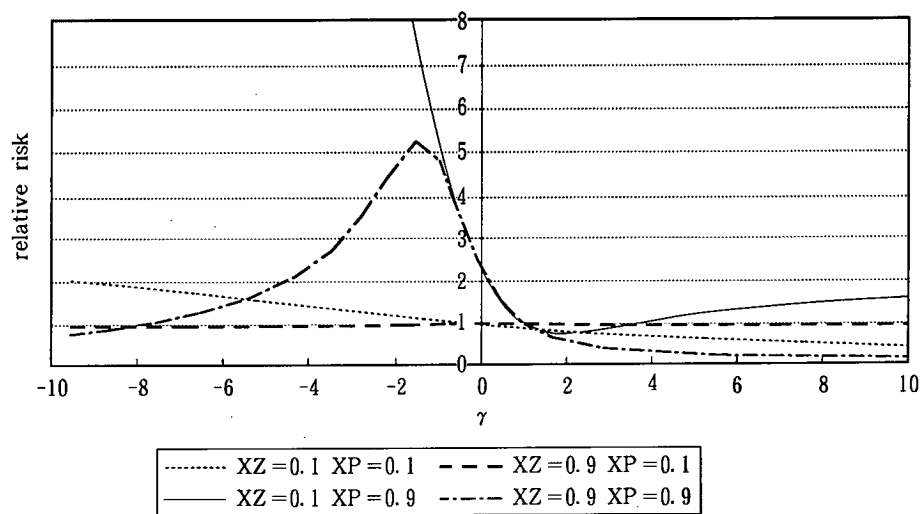
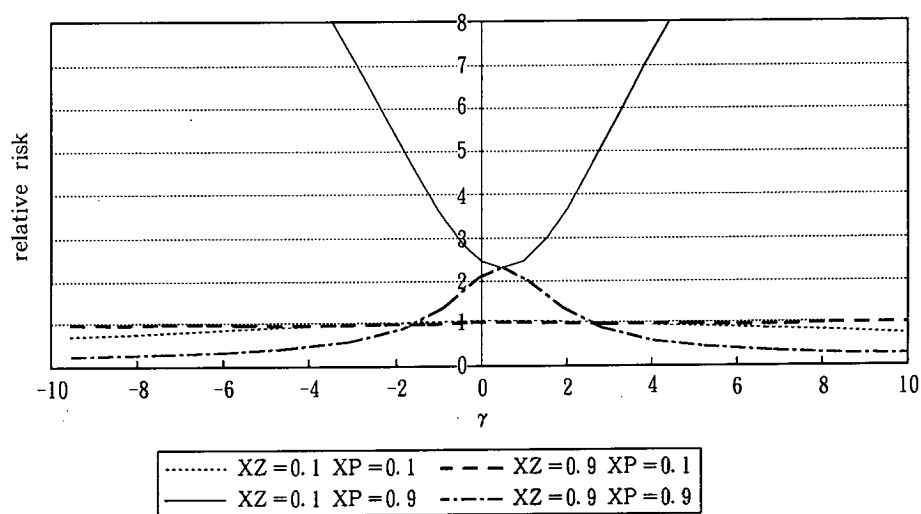
图 1 $C2=0.1$ 图 2 $C2=1$ 

図3 $C2=0.1$ 図4 $C2=1$ 

付 録

(10)式の導出について

まず, $\hat{\delta}$ の危険関数は次式で表される。

$$R(\hat{\delta}) = C_1 \int_{-\infty}^{\beta} (\hat{\delta} - \beta) g(\hat{\delta}) d\hat{\delta} + C_2 \int_{\beta}^{\infty} (\hat{\delta} - \beta) g(\hat{\delta}) d\hat{\delta}, \quad (a)$$

ただし,

$$\begin{aligned} g(\hat{\delta}) &= D^{1/2} [2\pi(P'P)\sigma^2]^{-1/2} \exp(-D(\hat{\delta} - \beta - \lambda_\gamma)^2 / [2\sigma^2(P'P)]), \\ D &= (P'P)(X'X) - (X'P)^2. \end{aligned}$$

ここで, $t = (\hat{\delta} - \beta - \lambda_\gamma) D^{1/2} / [\sigma(P'P)^{1/2}]$ とおくと, まず (a) 式の第 1 項目の積分部分が次のようになる。

$$\begin{aligned} & \int_{-\infty}^{\beta} (\hat{\delta} - \beta) g(\hat{\delta}) d\hat{\delta} \\ &= (2\pi D)^{-1/2} (P'P)^{1/2} \sigma \int_{-\infty}^{-c} t \exp(-t^2/2) dt \\ & \quad + \lambda_\gamma \int_{-\infty}^{-c} (2\pi)^{-1/2} \exp(-t^2/2) dt \\ &= (2\pi D)^{-1/2} (P'P)^{1/2} \sigma \exp(-c^2/2) + \lambda_\gamma F(-c). \end{aligned} \quad (b)$$

ここで, $F(\cdot)$ は標準正規分布の分布関数を表し, また $c = [D/(P'P)]^{1/2} (\lambda_\gamma/\sigma)$ である。同様にして, (a) 式の第 2 項目の積分部分は次のようになる。

$$\begin{aligned} & \int_{\beta}^{\infty} (\hat{\delta} - \beta) g(\hat{\delta}) d\hat{\delta} \\ &= (2\pi D)^{-1/2} (P'P)^{1/2} \sigma \int_{-c}^{\infty} t \exp(-t^2/2) dt \\ & \quad + \lambda_\gamma \int_{-c}^{\infty} (2\pi)^{-1/2} \exp(-t^2/2) dt \\ &= (2\pi D)^{-1/2} (P'P)^{1/2} \sigma \exp(-c^2/2) + \lambda_\gamma (1 - F(-c)). \end{aligned} \quad (c)$$

以上, (b), (c) 式より本文 (10) 式が得られる。

注

- 1) MSE を用いた代理変数の評価については, Ohtani (1981), Srivastava and Madhulika (1990), Kennedy (1992), Trenkler and Toutenburg (1992) においても扱われている。また, 代理変数に関しては Barnow (1976), Ohtani and Hasegawa (1993) などがある。
- 2) LINEX 損失関数, 線形損失関数については Varian (1975), Zellner (1986), Pandey and Rai (1992), Berger (1980), Ohtani (1994, 1995), Wan and Kurumai (1999), Wan (1999) などの研究がある。
- 3) 付録参照。

参 考 文 献

- Aigner, D. J., (1974), "MSE dominance of least squares with errors-of-observation", *Journal of Econometrics*, 2, 365-372.
- Barnow, B. S., (1976), "The use of proxy variable when one or two independent variables are measured with error", *the American Statistician*, 30, 119-121.
- Berger, J. O., (1980), "*Statistical Decision Theory*", New York, Springer-Verlag.
- Frost, P. A., (1979), "Proxy variables and specification bias", *Review of Economics and Statistics*, 61, 323-325.
- Kennedy, P., (1992), "Proxy variables can be useful", *Journal of Quantitative Economics*, 8, 443-445.
- McCallum, B.T., (1972), "Relative asymptotic bias from errors of omission and measurement", *Econometrica*, 40, 757-758.
- Ohtani, K., (1981), "On the use of a proxy variable in prediction: An MSE comparison", *Review of Economics and Statistics*, 63, 627-628.
- Ohtani, K., (1994), "The density function of R^2 and $\bar{R}^2$, and their risk performance under asymmetric loss in misspecified linear regression models", *Economic Modelling*, 11, 463-471.
- Ohtani, K., (1995), "Generalized ridge regression estimators under LINEX loss function", *Statistical Papers*, 36, 99-110.
- Ohtani, K. and H. Hasegawa, (1993), "On small sample properties of R^2 in a linear regression model with multivariate t errors and proxy variables", *Econometric Theory*, 9, 504-515.
- Pandy, B. N. and O. Rai, (1992), "Bayesian estimation of mean and square of mean of normal distribution using LINEX loss function", *Communications in Statistics - Theory and Methods*, 21, 3369-3391.
- Srivastava, V. K. and Madhulika, (1990), "Use of proxy variables in regression analysis", *Journal of Quantitative Economics*, 6, 71-74.
- Trenkler, G. and H. Toutenburg, (1992), "Proxy variables and mean square error dominance in linear regression", *Journal of Quantitative Economics*, 8, 433-442.
- Varian, H. R., (1975), "A Bayesian approach to real estate assesment", in *Studies in Bayesian Econometrics and Statistics in Honor of Leonard J. Savage*, eds. S. E. Fienberg and A. Zellner, Amsterdam: North Holland, 195-208.
- Wan, A. T. K., (1999), "A Note on almost unbiased generalised ridge regression estimator under asymmetric loss", *Journal of Statistical Computation and Simulation* 62, 411-421.
- Wan, A. T. K. and H. Kurumai, (1999), "An iterative feasible minimum mean squared error estimator of the disturbance variance in linear regression under asymmetric loss", *Statistics & Probability Letters*, 45, 253-259.
- Wickens, M. R., (1972), "A note on the use of proxy variables", *Econometrica*, 40, 759-761.
- Zellner, A., (1986), "Bayesian estimation and prediction using asymmetric loss function", *Journal of the American Statistical Association*, 81, 446-451.