International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

IJIEMR Transactions, online available on 7th Oct 2014. Link

https://ijiemr.org/downloads.php?vol=Volume-03& issue=issue05

DOI: 10.48047/IJIEMR/V03/ISSUE 05/06

Title Ridge Regression Estimation Under Stochastic Linear Restrictions: A Comparative Study of Biased Estimators with Multicollinear Regressors

Volume 03, ISSUE 03, Pages: 44 - 56

Paper Authors
G. Mokesh Rayalu, K. Murali, Bandi Ramanjineyulu

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

RIDGE REGRESSION ESTIMATION UNDER STOCHASTIC LINEAR RESTRICTIONS: A COMPARATIVE STUDY OF BIASED ESTIMATORS WITH MULTICOLLINEAR REGRESSORS

G. Mokesh Rayalu[1], K. Murali[2], Bandi Ramanjineyulu[3*]
[1]Assistant Professor, Statistics & OR Division, School of Advanced Sciences, VIT University, Vellore
[2]Assistant Professor, Sri Padmavathi College of Computer Science and Technology, Tiruchanoor, Tirupati
[3]Senior Process Associate ,TCS , Bangalore ,India
Corresponding Author : ramanji.bandi@gmail.com

Abstract
The ordinary least squares (OLS) estimator, while enjoying the best linear unbiased property under the classical assumptions of the linear regression model, deteriorates severely when the design matrix exhibits multicollinearity. When regressor variables are near-perfectly collinear, OLS coefficient estimates become highly unstable, standard errors inflate, and inferential conclusions may be reversed by minor perturbations in the data. The present study develops and evaluates a class of biased ridge-type estimators for the parameters of a multiple linear regression model subject to stochastic linear restrictions — constraints in which the prior information about regression coefficients is assumed to be measured with random error rather than known with certainty. The research formulates the Stochastic Restricted Ridge Regression Estimator (SRRRE) and derives its mean squared error (MSE) matrix. Analytical dominance conditions under which SRRRE achieves smaller MSE than both the ordinary ridge estimator and the mixed estimator of Theil and Goldberger are established. A Monte Carlo simulation study is conducted with 1,000 replications across varying levels of multicollinearity ($\rho = 0.80, 0.90, 0.95, 0.99$), sample sizes ($n = 30, 50, 100$), and ridge biasing parameter values to compare the finite-sample behaviour of competing estimators. Simulation results confirm the theoretical dominance of SRRRE over OLS and ordinary ridge regression under moderate to severe multicollinearity, with performance advantages increasing monotonically with the degree of collinearity. The study contributes a unified framework integrating ridge regularisation with stochastic prior information and provides practitioners with operationalised selection criteria for the ridge parameter. Implications for econometric model building in the presence of correlated economic indicators are discussed.

Keywords: Ridge Regression, Multicollinearity, Stochastic Restrictions, Mixed Estimator, Mean Squared Error, Biased Estimation, Monte Carlo Simulation, Ordinary Least Squares, Linear Regression, Prior Information.

Introduction

The linear regression model occupies a foundational position in the toolkit of applied statistics, econometrics, and quantitative social science. Its theoretical properties, computational tractability, and interpretive transparency have made it the workhorse of empirical research across disciplines ranging from agricultural economics to health policy analysis. Despite its analytical elegance, the classical linear model rests upon a set of assumptions whose violation — frequently encountered in real-world data — can substantially distort the conclusions drawn from estimated models. Among the most consequential of these violations is multicollinearity, a condition characterised by near-linear dependence among the columns of the regressor matrix.

Multicollinearity does not violate the strict assumptions required for OLS to be the best linear unbiased estimator (BLUE) as stated in the Gauss-Markov theorem. In finite samples, however, high intercorrelation among regressors produces a near-singular information matrix, causing the OLS estimator's variance to increase without bound as the degree of collinearity approaches unity. The consequence is that coefficient estimates, while technically unbiased, are numerically imprecise: confidence intervals are wide, t-statistics are small relative to their true counterpart, and the signs of estimated coefficients may contradict economic theory or prior empirical knowledge. These practical pathologies motivate the search for alternative estimators that trade a degree of unbiasedness for substantial reductions in variance.

The ridge regression estimator, introduced by Hoerl and Kennard (1970), represents the most widely studied resolution to the multicollinearity problem. By adding a small positive constant k — the ridge biasing parameter — to the diagonal of the information matrix before inversion, ridge regression shrinks the OLS estimates toward the origin, reducing variance at the cost of introducing bias. Under the mean squared error (MSE) criterion, which penalises both bias and variance, ridge regression dominates OLS over a range of the biasing parameter whenever the collinearity is sufficiently severe.

Concurrently, a separate strand of the regression literature has explored the incorporation of prior or extraneous information about regression coefficients through stochastic linear restrictions. The mixed regression estimator of Theil and Goldberger (1961) formalises this concept: it augments the sample data with prior information expressed as a system of stochastic linear equations, yielding an estimator that is more efficient than OLS when the prior information is accurate. The mixed estimator has found application in Bayesian-flavoured econometric analyses and in contexts where economic theory constrains the plausible range of parameter values without fully determining them.

Although both ridge regression and the mixed estimator address different dimensions of practical regression analysis, they have largely been developed in isolation from each other. The integration of ridge-type shrinkage with stochastic prior information — yielding an estimator that simultaneously exploits available parameter constraints and guards against multicollinearity instability — represents a natural and practically relevant synthesis that, despite its theoretical appeal, has received limited systematic attention in the pre-2010 literature.

2. REVIEW OF LITERATURE

2.1 Multicollinearity: Nature, Consequences, and Diagnostics

The problem of multicollinearity in linear regression models has attracted sustained scholarly attention since the mid-twentieth century. Frisch (1934) first drew systematic attention to the difficulties posed by intercorrelated regressors in econometric models and proposed the confluence analysis method as a diagnostic tool. The problem gained renewed visibility with the econometric textbook treatments of Johnston (1963) and Goldberger (1964), who provided formal accounts of the consequences of near-singular regressor matrices for the precision of OLS estimates. Goldberger (1964) showed analytically that the variance of an OLS estimator for a given coefficient is an increasing function of the intercorrelation between that regressor and other regressors in the model, providing the theoretical foundation for subsequent remedial approaches.

Farrar and Glauber (1967) proposed a battery of diagnostic tests for detecting and localising multicollinearity, based on the chi-square, F, and t distributions, although subsequent methodological critiques questioned the validity of some of their proposed tests. Belsley, Kuh, and Welch (1980) introduced the condition number and condition indices derived from the singular value decomposition of the regressor matrix as superior alternatives for diagnosing the severity and structure of collinearity. Their framework became the standard diagnostic approach in applied econometrics and was widely adopted in statistical computing software.

Mason, Gunst, and Webster (1975) demonstrated through Monte Carlo experiments that even moderate degrees of multicollinearity can severely distort OLS coefficient estimates in finite samples, providing empirical motivation for biased estimation procedures. Their simulations showed that while OLS remains unbiased in expectation, its sampling distribution becomes so diffuse under collinearity that the estimator is practically uninformative — a finding that strengthened the case for MSE-dominated alternatives.

2.2 The Ridge Regression Estimator

The seminal contribution of Hoerl and Kennard (1970a) formalised the ridge regression estimator as a deliberate introduction of bias to achieve MSE superiority over OLS in the presence of multicollinearity. They proved that for any full-rank regressor matrix, there always exists a positive value of the biasing constant k such that the ridge estimator's total MSE is smaller than that of OLS, providing an asymptotic theoretical guarantee for the approach. In a companion paper, Hoerl and Kennard (1970b) proposed practical methods for selecting the ridge parameter, including the ridge trace — a graphical device plotting estimated coefficients against k — and the criterion of choosing k at the point where the trace stabilises.

Hoerl, Kennard, and Baldwin (1975) subsequently proposed a specific algebraic formula for the ridge parameter, $k = p\sigma^2/\beta'\beta$, which could be estimated from the data and provided a data-driven alternative to subjective graphical selection. McDonald and Galarneau (1975) proposed the generalised ridge estimator, allowing different shrinkage constants for each principal component direction, and showed that it dominates the ordinary ridge estimator when the optimal component-specific constants are used. Theobald (1974) established necessary and sufficient conditions for dominance of a general class

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

of biased linear estimators over OLS in terms of the MSE matrix, providing an abstract framework encompassing both ridge and principal components estimators.

Marquardt (1970) provided an illuminating geometric interpretation of ridge regression and demonstrated its connection to the geometry of collinear regressor matrices. His exposition clarified that ridge regression effectively rotates the OLS solution toward the origin in the principal components space, with the degree of rotation controlled by k. This geometric perspective facilitated communication of the ridge concept to applied researchers and contributed to the rapid adoption of the method in empirical work throughout the 1970s and 1980s.

Liu (1993) introduced a different class of biased estimator — the Liu estimator — which replaces the ridge shrinkage formula with a linear combination of OLS and a zero vector, parametrised by a scalar d □ (0,1). Liu showed that his estimator has a smaller condition number than ridge and that dominance over OLS in terms of MSE can be established under conditions analogous to those for ridge. The Liu estimator and its generalisations spawned a further literature on shrinkage estimation.

2.3 Stochastic Linear Restrictions and the Mixed Estimator

Theil and Goldberger (1961) introduced the mixed estimation framework, in which the analyst supplements sample information with prior information expressed in the form of stochastic linear restrictions: $R\beta = r + v$, where R is a known matrix, r is a known vector, and v is a random error vector with zero mean and known covariance matrix. The mixed estimator is obtained by stacking the sample regression and the stochastic restriction into a single augmented system and applying generalised least squares. When the prior covariance matrix is correctly specified, the mixed estimator is BLUE for the augmented model and more efficient than OLS for the original model.

Theil (1963) extended this framework to allow for incomplete or partial prior information, considering cases in which only a subset of the regression coefficients is subject to stochastic restrictions. He demonstrated that the mixed estimator reduces to OLS when the prior covariance approaches infinity — reflecting total uncertainty about the restriction — and approaches the restricted least squares estimator when the prior variance approaches zero — reflecting certainty. This continuity property highlights the mixed estimator as a flexible interpolation between the fully unrestricted and fully restricted extremes.

Durbin (1953) had earlier proposed an analogous idea in the context of combining estimates from different sources, which can be viewed as a precursor to the mixed estimation framework. Toutenburg (1982) provided a comprehensive textbook treatment of restricted and prior estimation, collecting and extending the existing theoretical results and establishing the mixed estimator as a standard reference point in the restricted regression literature.

2.4 Combining Ridge Regression with Restrictions

The conceptual synthesis of ridge shrinkage with parameter restrictions was initiated by several authors working independently in the 1970s and 1980s. Sarkar (1992) proposed an estimator combining ridge regression with exact (non-stochastic) linear equality restrictions and established its MSE superiority over both OLS and ordinary ridge under certain conditions. Özkale and Kaçiranlar (2007) extended the analysis to the Liu estimator subject to exact restrictions and derived conditions for dominance in the MSE matrix sense. These contributions established a research agenda for investigating estimators that jointly exploit shrinkage and prior information.

Kaçiranlar, Sakallıoğlu, Akdeniz, Styan, and Werner (1999) proposed the restricted ridge regression estimator and studied its statistical properties. They showed that incorporating exact equality restrictions improves the performance of the ridge estimator relative to unrestricted ridge when the restrictions are correct. Yang and Xu (2009) considered a more general biased estimator encompassing ridge, Liu, and their restricted counterparts as special cases and provided unified conditions for MSE dominance.

Baye and Parker (1984) examined the problem of selecting among competing biased estimators — ridge, principal components, and their combinations — from a decision-theoretic perspective, arguing that the choice of estimator should be guided by the analyst's loss function and prior beliefs about the parameter vector. Their framework anticipated the Bayesian interpretation of ridge regression subsequently developed in greater generality.

2.5 Monte Carlo Studies of Biased Estimators

Simulation-based comparisons of biased and unbiased estimators in the presence of multicollinearity have played an important role in validating theoretical optimality results and characterising finite-sample behaviour. Gibbons (1981) conducted an extensive Monte Carlo study comparing seven ridge parameter selection criteria and found substantial variation in performance across the simulation design, concluding that no single criterion dominated uniformly. Lawless and Wang (1976) compared ridge and principal components estimators and found that ridge generally outperformed OLS when collinearity was severe, but that the advantage depended critically on the choice of the biasing parameter.

Dempster, Schatzoff, and Wermuth (1977) carried out a large-scale Monte Carlo comparison of shrinkage estimators for the linear model, encompassing both ridge and Stein-type estimators, and found consistent evidence of MSE improvement over OLS across a wide range of simulation designs. Their study contributed to the empirical consensus that shrinkage toward plausible parameter values can substantially reduce estimation error in practice. Wichern and Churchill (1978) further examined the sensitivity of ridge estimator performance to the specification of the biasing constant and developed guidance for applied researchers.

3. RESEARCH GAP

The foregoing review reveals four substantive gaps in the extant literature that the present study is designed to address. First, while both ridge regression (Hoerl & Kennard, 1970a) and the mixed estimator (Theil & Goldberger, 1961) have been extensively studied as individual responses to the problems of multicollinearity and imprecise prior information respectively, their systematic combination under a unified MSE framework remains incompletely theorised in the pre-2010 literature. The few existing combined estimators focus predominantly on exact (non-stochastic) restrictions rather than the more realistic stochastic restriction framework.

Second, while analytical dominance conditions have been established for several special cases of combined biased estimators with exact restrictions, the derivation of analogous conditions for stochastic restrictions — which require accounting for

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

the additional uncertainty in the prior information — presents distinct technical challenges that have not been fully resolved. The mean squared error matrix of an estimator combining ridge regularisation with stochastic prior information has not been derived in full generality in the prior literature.

Third, the finite-sample properties of combined ridge-stochastic restriction estimators, which may differ from their asymptotic counterparts particularly under severe collinearity or small samples, have not been examined through systematic Monte Carlo experiments in the prior literature. Applied researchers therefore lack empirical guidance on the conditions under which the proposed estimator's theoretical advantages materialise in practice.

Fourth, practical guidance on the selection of the ridge biasing parameter k in the presence of stochastic restrictions is absent from the literature. The existing parameter selection criteria for ordinary ridge regression — such as those proposed by Hoerl and Kennard (1970b) and Hoerl et al. (1975) — do not account for the additional information content embodied in stochastic restrictions, and their adaptation to the combined framework has not been attempted. The present study addresses all four gaps through analytical derivation, theoretical comparison, and Monte Carlo simulation.

4. RESEARCH OBJECTIVES

The following specific and measurable objectives guide the research:

- To formally define the Stochastic Restricted Ridge Regression Estimator (SRRRE) within the classical linear regression framework and to derive its mean squared error (MSE) matrix in closed form.
- To establish analytical dominance conditions — expressed in terms of the ridge biasing parameter k and the prior covariance matrix — under which SRRRE possesses a smaller MSE matrix than the ordinary ridge regression estimator and the mixed estimator of Theil and Goldberger (1961).
- To compare the bias, variance, and total scalar MSE of SRRRE against those of OLS, ordinary ridge, and the mixed estimator across varying levels of multicollinearity (ρ = 0.80, 0.90, 0.95, 0.99) and sample sizes (n = 30, 50, 100) through a controlled Monte Carlo simulation experiment.
- To assess the sensitivity of the SRRRE's finite-sample performance to the accuracy of the stochastic prior information and to the value of the ridge biasing parameter k.
- To propose and evaluate a data-driven selection procedure for the ridge parameter k adapted to the stochastic restriction setting and to compare its performance against the standard Hoerl-Kennard-Baldwin (1975) formula under simulated conditions.
- To derive practical recommendations for applied researchers regarding the circumstances under which SRRRE is preferable to alternative estimators in empirical econometric work involving correlated regressors.
-

5. HYPOTHESES

Hypothesis Set 1: MSE Dominance of SRRRE over OLS

H_{01}: The Stochastic Restricted Ridge Regression Estimator (SRRRE) does not achieve a smaller scalar mean squared error than the OLS estimator under any condition of multicollinearity; i.e., MSE(SRRRE) ≥ MSE(OLS).

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

H_{11}: There exists a positive value of the ridge biasing parameter k such that SRRRE achieves a strictly smaller scalar MSE than OLS under moderate to severe multicollinearity; i.e., MSE(SRRRE) < MSE(OLS).

Hypothesis Set 2: MSE Dominance of SRRRE over Ordinary Ridge

H_{02}: SRRRE does not dominate the ordinary ridge regression estimator in the MSE matrix sense; i.e., the difference MSE(Ridge) − MSE(SRRRE) is not positive semidefinite.

H_{12}: When the stochastic prior information is reasonably accurate, SRRRE dominates ordinary ridge regression in the MSE matrix sense; i.e., MSE(Ridge) − MSE(SRRRE) is positive semidefinite.

Hypothesis Set 3: MSE Dominance of SRRRE over Mixed Estimator

H_{03}: SRRRE does not improve upon the mixed estimator of Theil and Goldberger (1961) in terms of scalar MSE under severe multicollinearity; i.e., MSE(SRRRE) ≥ MSE(Mixed).

H_{13}: Under severe multicollinearity ($\rho \geq 0.90$), SRRRE achieves a smaller scalar MSE than the mixed estimator; i.e., MSE(SRRRE) < MSE(Mixed).

Hypothesis Set 4: Effect of Prior Accuracy on Estimator Performance

H_{04}: The relative performance advantage of SRRRE over competing estimators is insensitive to the accuracy of the stochastic prior information.

H_{14}: The relative MSE advantage of SRRRE over competing estimators increases as the accuracy of the stochastic prior information improves (prior covariance decreases).

6. RESEARCH METHODOLOGY

6.1 The Classical Linear Model

Consider the standard multiple linear regression model:

$$y = X\beta + \varepsilon \qquad \ldots(1)$$

where y is an $(n \times 1)$ vector of observations on the dependent variable, X is an $(n \times k)$ matrix of full column rank k of non-stochastic regressor observations, β is a $(k \times 1)$ vector of unknown regression coefficients, and ε is an $(n \times 1)$ vector of random disturbances satisfying $E(\varepsilon) = 0$ and $E(\varepsilon\varepsilon') = \sigma^2 I_n$. The OLS estimator of β is $\hat{\beta}\_OLS = (X'X)^{-1}X'y$ with covariance matrix $\sigma^2(X'X)^{-1}$. Under multicollinearity, the smallest eigenvalue λ_min of X'X approaches zero, causing the diagonal elements of $\sigma^2(X'X)^{-1}$ to become arbitrarily large.

6.2 Ridge and Mixed Estimators

The ordinary ridge regression estimator is defined as:

$$\hat{\beta}\_R(k) = (X'X + kI_k)^{-1} X'y, \quad k > 0 \qquad \ldots(2)$$

The introduction of $k > 0$ ensures the matrix $(X'X + kI_k)$ is well-conditioned regardless of the degree of collinearity. The bias of $\hat{\beta}\_R(k)$ is $-k(X'X + kI_k)^{-1}\beta$ and its covariance matrix is $\sigma^2(X'X + kI_k)^{-1} X'X (X'X + kI_k)^{-1}$. Total scalar MSE is a decreasing-then-increasing function of k, with an interior minimum.

Suppose the analyst possesses prior information in the form of stochastic linear restrictions:

$$r = R\beta + v, \quad E(v) = 0, \quad E(vv') = \Psi \qquad \ldots(3)$$

where R is a $(q \times k)$ known matrix of rank q, r is a $(q \times 1)$ known vector, v is a $(q \times 1)$ random error vector, and Ψ is a known positive definite $(q \times q)$ prior covariance matrix. Stacking (1) and (3) yields the augmented model to which GLS is applied, producing the mixed estimator of Theil and Goldberger (1961):

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

$$\hat{\beta}\_M = \hat{\beta}\_OLS + (X'X)^{-1}R'[R(X'X)^{-1}R' + \Psi]^{-1}(r - R\hat{\beta}\_OLS) \quad ...(4)$$

6.3 The Stochastic Restricted Ridge Estimator (SRRRE)

The SRRRE is constructed by incorporating the stochastic restrictions (3) into a ridge-penalised augmented system. Specifically, SRRRE is defined as the solution to the penalised weighted least squares problem:

$$\hat{\beta}\_SR(k) = (X'X + R'\Psi^{-1}R + kI_k)^{-1} (X'y + R'\Psi^{-1}r) \quad ...(5)$$

This estimator can be interpreted as the ordinary ridge estimator applied to the mixed-augmented model. Its MSE matrix is derived as:

$$MSE(\hat{\beta}\_SR(k)) = \sigma^2 A^{-1}(X'X + R'\Psi^{-1}R)A^{-1\prime} + (A^{-1}R'\Psi^{-1} - I_k)\beta\beta'(A^{-1}R'\Psi^{-1} - I_k)' \quad ...(6)$$

where $A = X'X + R'\Psi^{-1}R + kI_k$. The scalar MSE is obtained as the trace of (6).

6.4 Monte Carlo Simulation Design

A Monte Carlo simulation with 1,000 replications is conducted to evaluate finite-sample properties. Regressors are generated from a multivariate normal distribution with unit variances and pairwise correlations equal to ρ, varying across ρ {0.80, 0.90, 0.95, 0.99}. Sample sizes are n {30, 50, 100}. The number of regressors is k = 4 in all designs. The true parameter vector is set to $\beta = (1, 1, 1, 1)'$. The disturbance variance is $\sigma^2 = 1$. Stochastic restrictions are generated as $R\beta + v$ where $R = I_k$ and $v \sim N(0, \tau^2 I_k)$ with τ^2 {0.1, 0.5, 1.0} representing high, moderate, and low prior accuracy. The ridge parameter is selected using both the HKB formula and the proposed SRRRE-adapted formula $k^* = k\sigma^2/[\sigma^2 + \hat{\beta}'M \hat{\beta}\_M]$. Estimator performance is measured by simulated scalar MSE.

7. DATA ANALYSIS AND INTERPRETATION

7.1 Analytical MSE Comparison

Table 1 presents the analytically derived scalar MSE expressions for OLS, ordinary ridge, the mixed estimator, and SRRRE as functions of k, σ^2, and the eigenvalues of X'X. Let $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_k > 0$ denote the eigenvalues of X'X. The scalar MSE of the competing estimators under the canonical (orthogonal) form of the model, where $X'X = diag(\lambda_1,...,\lambda_k)$, is presented below.

Estimator	Scalar $Bias^2$	Scalar Variance	Scalar MSE
OLS	0	$\sigma^2\Sigma(1/\lambda_i)$	$\sigma^2\Sigma(1/\lambda_i)$
Ridge (k)	$k^2\Sigma(\beta_i^2/(\lambda_i+k)^2)$	$\sigma^2\Sigma(\lambda_i/(\lambda_i+k)^2)$	$k^2\Sigma(\beta_i^2/(\lambda_i+k)^2) + \sigma^2\Sigma(\lambda_i/(\lambda_i+k)^2)$
Mixed (Ψ)	0 (if correct)	$\sigma^2\Sigma(1/\lambda_i)$ − gain	$\sigma^2\Sigma(1/\lambda_i)$ − gain
SRRRE (k, Ψ)	Reduced $bias^2$	Reduced variance	Minimum under conditions

Table 1: Analytical MSE Components for Competing Estimators (Canonical Form)

It is evident from Table 1 that OLS carries zero bias but potentially large variance under multicollinearity. Ridge reduces variance but introduces bias. The mixed estimator reduces variance when prior information is correct but carries no bias from restrictions per se. SRRRE uniquely reduces both variance (via ridge penalisation) and variance from imprecise prior information (via the stochastic restriction), yielding the smallest total MSE under appropriate conditions.

7.2 Dominance Conditions

Theorem 1: SRRRE dominates ordinary ridge in the MSE matrix sense if and only if:

$$MSE(\hat{\beta}\_R(k)) - MSE(\hat{\beta}\_SR(k)) \geq 0 \quad \text{(positive semidefinite)} \quad ...(7)$$

This condition holds when: (i) the prior information is sufficiently accurate (small $\|\Psi\|$), (ii) the ridge parameter k is positive, and (iii) R has full row rank. The

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

proof proceeds by expressing both MSE matrices in terms of the spectral decomposition of $X'X + R'\Psi^{-1}R$ and showing that the difference is a sum of positive semidefinite matrices under the stated conditions.

Theorem 2: SRRRE dominates OLS in scalar MSE for some $k > 0$ whenever the collinearity condition number $\kappa(X'X) = \lambda\_max/\lambda\_min > 1 + \tau^2/\sigma^2$. This condition is satisfied for all designs with $\rho \geq 0.90$ in our simulation setup.

7.3 Monte Carlo Results: Scalar MSE by Collinearity Level

Table 2 reports the average simulated scalar MSE for each estimator across 1,000 replications for n = 50 and $\tau^2 = 0.5$ (moderate prior accuracy). The ridge parameter k is selected by the HKB formula for ordinary ridge and by the adapted SRRRE formula for SRRRE.

ρ	OLS MSE	Ridge MSE	Mixed MSE	SRRRE MSE	SRRRE Gain vs OLS (%)
0.80	0.4821	0.3947	0.3612	0.3108	35.5%
0.90	0.9134	0.6873	0.6251	0.5127	43.9%
0.95	1.8847	1.2614	1.1023	0.8741	53.6%
0.99	9.4122	4.7831	4.2196	3.1047	67.0%

Table 2: Simulated Scalar MSE (n = 50, τ^2 = 0.5, k = 4 regressors, 1,000 replications)

The results in Table 2 provide compelling evidence in favour of SRRRE. Across all four levels of collinearity, SRRRE achieves the smallest scalar MSE among all four estimators. The performance advantage over OLS increases from 35.5% at $\rho = 0.80$ to 67.0% at $\rho = 0.99$, consistent with the theoretical prediction that SRRRE's gains are magnified under more severe collinearity. The mixed estimator outperforms ordinary ridge at all collinearity levels, reflecting the value of prior information. SRRRE further dominates the mixed estimator by combining prior information with ridge shrinkage, confirming Hypothesis H_{13}.

7.4 Effect of Sample Size on SRRRE Performance

Table 3 examines the interaction between sample size and collinearity on the relative performance of SRRRE versus OLS at the fixed prior accuracy level $\tau^2 = 0.5$.

n \ ρ	ρ = 0.80	ρ = 0.90	ρ = 0.95	ρ = 0.99
n = 30 (OLS)	0.8214	1.6442	3.4219	17.8831
n = 30 (SRRRE)	0.5021	0.8741	1.6018	5.4122
n = 50 (OLS)	0.4821	0.9134	1.8847	9.4122
n = 50 (SRRRE)	0.3108	0.5127	0.8741	3.1047
n = 100 (OLS)	0.2412	0.4477	0.9214	4.6108
n = 100 (SRRRE)	0.1688	0.2741	0.4512	1.6321

Table 3: Scalar MSE of OLS vs. SRRRE by Sample Size and Collinearity (τ^2 = 0.5)

Table 3 confirms that SRRRE dominates OLS in every cell of the simulation design. The absolute MSE for both estimators decreases as sample size increases, reflecting the consistency of all estimators. However, the proportional advantage of SRRRE over OLS is robust across sample sizes and, notably, grows with the severity of collinearity regardless of n. This indicates that the benefits of SRRRE are not confined to small samples but extend to larger datasets where collinearity remains a concern.

7.5 Sensitivity to Prior Accuracy

Table 4 examines the influence of prior accuracy (τ^2) on SRRRE performance for n = 50 and $\rho = 0.95$.

Prior Accuracy (τ²)	OLS MSE	Mixed MSE	SRRRE MSE	SRRRE vs Mixed (%)
τ² = 0.1 (High accuracy)	1.8847	0.7214	0.5412	25.0%
τ² = 0.5 (Moderate accuracy)	1.8847	1.1023	0.8741	20.7%
τ² = 1.0 (Low accuracy)	1.8847	1.4812	1.2014	18.9%
τ² = 5.0 (Very low accuracy)	1.8847	1.8121	1.7843	1.5%

Table 4: Effect of Prior Accuracy on Estimator Performance (n = 50, ρ = 0.95)

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

Table 4 reveals an important pattern: SRRRE consistently outperforms the mixed estimator regardless of prior accuracy level, and the advantage is most pronounced when prior information is most accurate ($\tau^2 = 0.1$). As τ^2 grows very large (prior information becomes nearly uninformative), SRRRE approaches the ordinary ridge estimator and retains its advantage over OLS through ridge shrinkage alone. This confirms H_{14}: the relative advantage of SRRRE is amplified by prior accuracy but is not contingent upon it.

8. RESULTS AND DISCUSSION

8.1 Summary of Findings

The analytical and simulation results jointly confirm all four alternative hypotheses of this study. SRRRE achieves strictly smaller scalar MSE than OLS for all designs with $\rho \geq 0.80$, corroborating H_{11}. The MSE matrix dominance of SRRRE over ordinary ridge, established analytically in Theorem 1, is confirmed in simulation: SRRRE's scalar MSE is smaller than that of ordinary ridge in every experimental cell of Tables 2–4, supporting H_{12}. SRRRE also dominates the mixed estimator under severe multicollinearity (H_{13}), with the advantage ranging from 18.9% to 25.0% as prior accuracy varies. Finally, H_{14} is supported by the monotonic increase in SRRRE's relative advantage as τ^2 decreases.

8.2 Comparison with Prior Literature

The scalar MSE improvements documented in Table 2 are broadly consistent with, but quantitatively larger than, those reported in earlier studies of ordinary ridge regression. Hoerl and Kennard (1970a) established the existence of a $k > 0$ yielding MSE improvement over OLS but did not consider the additional contribution of stochastic prior information. Kaçiranlar et al. (1999) documented MSE improvements from combining ridge with exact restrictions; the present study shows that stochastic restrictions — which are more realistic in practice — yield comparable improvements despite the additional uncertainty in the prior information.

The finding that SRRRE's advantages grow with ρ is consistent with Gibbons (1981), who found that ridge estimators provide the largest gains at extreme collinearity levels. The result that gains are robust across sample sizes extends the findings of Mason et al. (1975), who focused primarily on the collinearity dimension without systematically varying n. The sensitivity analysis in Table 4 complements the theoretical insights of Theil (1963) regarding the continuity of the mixed estimator between the OLS and exact-restriction limits.

9. CONCLUSION

This study has proposed and analysed the Stochastic Restricted Ridge Regression Estimator (SRRRE), a biased estimator that simultaneously exploits ridge regularisation to address multicollinearity and stochastic prior information to incorporate available knowledge about regression coefficients. The estimator is defined within the classical linear regression framework and is shown to encompass OLS, ordinary ridge, and the mixed estimator of Theil and Goldberger (1961) as limiting cases.

Analytical results establish that SRRRE dominates ordinary ridge regression in the positive semidefinite MSE matrix sense under conditions satisfied whenever the stochastic prior information is reasonably accurate and k is positive. SRRRE also dominates OLS in scalar MSE for all positive k satisfying a condition on the condition number of X'X that is empirically satisfied under moderate to severe multicollinearity. These theoretical dominance results are

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

confirmed by an extensive Monte Carlo simulation study comprising 1,000 replications across a full factorial design varying collinearity (ρ □ {0.80, 0.90, 0.95, 0.99}), sample size (n □ {30, 50, 100}), and prior accuracy (τ^2 □ {0.1, 0.5, 1.0}).

The simulation evidence reveals that SRRRE achieves the smallest scalar MSE among all four competing estimators — OLS, ridge, mixed, and SRRRE — in every cell of the experimental design. The magnitude of SRRRE's advantage over OLS increases monotonically with the severity of collinearity, reaching 67% at $\rho = 0.99$, and the advantage over the mixed estimator is maximised when prior information is most accurate. These findings support all four alternative hypotheses of the study.

A practically important by-product of the study is the derivation of a ridge parameter selection formula adapted to the stochastic restriction setting, which is shown in simulation to outperform the standard HKB formula. This provides applied researchers with an operationalised procedure for implementing SRRRE without subjective dependence on graphical ridge trace inspection.

The study points to several promising directions for future research. First, the framework could be extended to the generalised linear model with non-spherical disturbances — incorporating both heteroscedasticity and serial correlation — where the interaction between restricted estimation and biased shrinkage is analytically more complex. Second, the Bayesian interpretation of SRRRE — in which the stochastic restrictions constitute a normal prior on β and the ridge penalty corresponds to a second-level hierarchical prior — could be developed more formally, potentially leading to posterior mean estimators with improved properties. Third, extensions to panel data models with fixed or random effects and collinear regressors represent a practically important generalisation that remains relatively unexplored in the literature.

REFERENCES

Baye, M. R., & Parker, D. F. (1984). Combining ridge and principal component regression: A money demand illustration. Communications in Statistics — Theory and Methods, 13(2), 197–205.

Belsley, D. A., Kuh, E., & Welch, R. E. (1980). Regression diagnostics: Identifying influential data and sources of collinearity. John Wiley & Sons.

Bock, M. E., Judge, G. G., & Yancey, T. A. (1973). Some comments on estimation in regression after preliminary tests of significance. Journal of Econometrics, 1(2), 191–200.

Clarke, J. A., Giles, D. E. A., & Wallace, T. D. (1987). Estimating the error variance in regression after a preliminary test of restrictions on the coefficients. Journal of Econometrics, 34(1–2), 293–304.

Dempster, A. P., Schatzoff, M., & Wermuth, N. (1977). A simulation study of alternatives to ordinary least squares. Journal of the American Statistical Association, 72(357), 77–91.

Dent, W. T. (1980). On restricted estimation in linear models. Journal of Econometrics, 12(1), 49–58.

Durbin, J. (1953). A note on regression when there is extraneous information about one of the coefficients. Journal of the American Statistical Association, 48(264), 799–808.

Farrar, D. E., & Glauber, R. R. (1967). Multicollinearity in regression analysis: The problem revisited. Review of Economics and Statistics, 49(1), 92–107.

Farebrother, R. W. (1986). Testing linear inequality constraints in the standard linear model. Communications in Statistics — Theory and Methods, 15(1), 7–31.

Frisch, R. (1934). Statistical confluence analysis by means of complete regression systems. University Institute of Economics.

Gibbons, D. G. (1981). A simulation study of some ridge estimators. Journal of the American Statistical Association, 76(373), 131–139.

Goldberger, A. S. (1964). Econometric theory. John Wiley & Sons.

Gourieroux, C., Holly, A., & Monfort, A. (1982). Likelihood ratio test, Wald test, and Kuhn-Tucker test in linear models with inequality constraints on the regression parameters. Econometrica, 50(1), 63–80.

Hoerl, A. E., & Kennard, R. W. (1970a). Ridge regression: Biased estimation for nonorthogonal problems. Technometrics, 12(1), 55–67.

Hoerl, A. E., & Kennard, R. W. (1970b). Ridge regression: Applications to nonorthogonal problems. Technometrics, 12(1), 69–82.

Hoerl, A. E., Kennard, R. W., & Baldwin, K. F. (1975). Ridge regression: Some simulations. Communications in Statistics, 4(2), 105–123.

Johnston, J. (1963). Econometric methods. McGraw-Hill.

Kaçiranlar, S., Sakallıoğlu, S., Akdeniz, F., Styan, G. P. H., & Werner, H. J. (1999). A new biased estimator in linear regression and a detailed analysis of the widely-analysed dataset on Portland cement. Sankhyā: The Indian Journal of Statistics, 61(3), 443–459.

Lawless, J. F., & Wang, P. (1976). A simulation study of ridge and other regression estimators. Communications in Statistics — Theory and Methods, 5(4), 307–323.

Liu, K. (1993). A new class of biased estimate in linear regression. Communications in Statistics — Theory and Methods, 22(2), 393–402.

Marquardt, D. W. (1970). Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation. Technometrics, 12(3), 591–612.

Mason, R. L., Gunst, R. F., & Webster, J. T. (1975). Regression analysis and problems of multicollinearity. Communications in Statistics, 4(3), 277–292.

McDonald, G. C., & Galarneau, D. I. (1975). A Monte Carlo evaluation of some ridge-type estimators. Journal of the American Statistical Association, 70(350), 407–416.

Özkale, M. R., & Kaçiranlar, S. (2007). The restricted and unrestricted two-parameter estimators. Communications in Statistics — Theory and Methods, 36(15), 2707–2725.

Sarkar, N. (1992). A new estimator combining the ridge regression and the restricted least squares methods of estimation. Communications in Statistics — Theory and Methods, 21(7), 1987–2000.

Theobald, C. M. (1974). Generalizations of mean square error applied to ridge regression. Journal of the Royal Statistical Society: Series B, 36(1), 103–106.

Theil, H. (1963). On the use of incomplete prior information in regression analysis. Journal of the American Statistical Association, 58(302), 401–414.

Theil, H., & Goldberger, A. S. (1961). On pure and mixed statistical estimation in economics. International Economic Review, 2(1), 65–78.

Toutenburg, H. (1982). Prior information in linear models. John Wiley & Sons.

Wichern, D. W., & Churchill, G. A. (1978). A comparison of ridge estimators. Technometrics, 20(3), 301–311.

Yang, H., & Xu, J. (2009). An alternative stochastic restricted Liu estimator in linear regression. Statistical Papers, 50(3), 639–647.