International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in

Shrinkage Estimation in Linear Regression Under Multicollinearity and Biased Prior Information: A Unified Ridge-Mixed Framework with Compatibility Testing

G. Mokesh Rayalu[1] , K. Murali[2], Bandi Ramanjineyulu[3*]

Assistant Professor, Statistics & OR Division, School of Advanced Sciences, VIT University, Vellore
Assistant Professor, Sri Padmavathi College of Computer Science and Technology, Tiruchanoor, Tirupati
Senior Process Associate ,TCS , Bangalore,India
Corresponding Author : ramanji.bandi@gmail.com

Abstract

Multicollinearity among regressors and the incorporation of potentially biased prior information are among the most pervasive challenges in applied econometric modeling, yet they have largely been treated as separate problems in the statistical literature. The ordinary least squares (OLS) estimator, while unbiased and efficient under Gauss–Markov conditions, exhibits catastrophically inflated variance when regressors are highly collinear, rendering coefficient estimates unstable and inferentially unreliable. Ridge regression addresses multicollinearity by introducing a scalar bias penalty that reduces variance at the cost of introducing estimator bias. The Mixed Regression Estimator (MRE) of Theil and Goldberger (1961) incorporates stochastic prior information to improve efficiency relative to OLS but is known to be severely degraded by bias in the prior restrictions. The present paper proposes two new estimators—the Stochastic Restriction Ridge Estimator (SRRE) and the Ridge Regression with Prior Information Estimator (RRPIE)—that simultaneously address multicollinearity and incorporate potentially biased stochastic prior information through a unified ridge-mixed framework. Theoretical MSE dominance conditions are derived analytically. An extended compatibility test statistic for the joint assessment of sample-prior information concordance in the presence of ridge shrinkage is developed and its asymptotic distribution under both null and local alternative hypotheses is characterized. Monte Carlo simulation experiments across 10,000 replications confirm that SRRE achieves the lowest total mean square error (TMSE) across a wide range of multicollinearity severity levels, prior information bias magnitudes, and sample sizes, dominating OLS, standard ridge, and the conventional MRE. An empirical application to a cross-country economic growth regression model illustrates the practical advantage of the proposed estimators, with the compatibility test confirming the adequacy of the prior information used. The findings have implications

for applied econometrics in settings characterized by limited sample information and theoretically motivated parameter restrictions.

Keywords: ridge regression, mixed regression estimator, stochastic prior information, multicollinearity, shrinkage estimation, mean square error, compatibility test, biased estimation, Theil-Goldberger estimator, econometric inference

1. Introduction

The classical linear regression model, estimated by ordinary least squares (OLS), rests on a set of regularity conditions—including non-stochastic regressors, homoscedastic spherical errors, and full rank of the design matrix—whose joint satisfaction guarantees the BLUE property of the Gauss–Markov theorem. In applied econometric practice, however, these conditions are routinely violated. Among the most consequential departures is multicollinearity—the presence of approximately linear relationships among the explanatory variables—which, while leaving the OLS estimator formally unbiased, inflates its variance to the point where coefficient estimates become statistically unreliable, sign-reversal is frequent, and inferences about individual regressor effects are rendered ambiguous (Hoerl & Kennard, 1970a; Judge et al., 1985).

Parallel to the multicollinearity problem is the challenge of incorporating non-sample information into the estimation process. Applied econometricians frequently possess prior knowledge about regression coefficients—from economic theory, previous empirical studies, institutional knowledge, or meta-analytic synthesis—that can meaningfully constrain the parameter space and improve estimation efficiency. Theil and Goldberger (1961) formalized this intuition in the Mixed Regression Estimator (MRE), which combines sample likelihood information with stochastic prior restrictions on regression coefficients. The MRE dominates OLS in mean square error when prior restrictions are correctly specified (i.e., unbiased), but Yancey, Judge, and Bock (1974) demonstrated that MRE can be dominated by OLS when prior restrictions are biased—a vulnerability that substantially limits its practical applicability in settings where perfect prior knowledge is unavailable.

Remarkably, despite the coexistence of multicollinearity and uncertain prior information in many applied econometric contexts, the two problems have been largely treated in isolation in the statistical literature. Ridge regression methods address multicollinearity without incorporating prior information; MRE methods incorporate prior information without addressing the variance inflation caused by multicollinearity. The development of a unified estimation framework that simultaneously combats multicollinearity through shrinkage and incorporates stochastic prior information while remaining robust to prior bias constitutes a significant unaddressed gap in the econometric estimation literature.

The present paper addresses this gap by proposing and analyzing two new shrinkage estimators—the Stochastic Restriction Ridge Estimator (SRRE) and the Ridge Regression with Prior Information Estimator (RRPIE)—that are derived by applying ridge-type shrinkage within the Theil–Goldberger augmented regression framework. Analytical conditions for the dominance of these estimators over OLS, ridge, and MRE in the MSE matrix sense are derived. An extended compatibility test statistic—assessing the concordance of sample and prior information in the presence of ridge shrinkage—is constructed and characterized asymptotically. Monte Carlo simulation studies confirm the analytical findings, and an empirical application to a cross-country growth regression model illustrates the practical benefits of the proposed framework.

2. Review of Literature

2.1 Multicollinearity and Ridge Regression

The problem of multicollinearity in linear regression—its consequences, detection, and remediation—has been one of the most extensively studied topics in applied econometrics since Hoerl and Kennard's seminal contributions in the late 1960s. Farrar and Glauber (1967) provided an influential diagnostic framework for detecting and measuring multicollinearity using chi-square, F, and t tests applied to various orthogonality measures of the design matrix, and documented its pervasiveness in applied econometric datasets. Their work catalyzed a sustained research program on multicollinearity remediation that spawned ridge regression, principal components regression, partial least squares, and a variety of related shrinkage methodologies.

Hoerl and Kennard (1970a) introduced ridge regression as a biased estimation alternative to OLS that trades a controlled amount of bias for a substantial reduction in variance, yielding a lower total mean square error over a range of ridge parameter values. Their foundational theoretical result demonstrated the existence of a positive ridge parameter k for which the ridge estimator dominates OLS in the scalar MSE sense for any value of the true parameter vector—a result that established the theoretical superiority of ridge regression as a routine alternative to OLS in the presence of multicollinearity. Hoerl and Kennard (1970b) developed practical methods for ridge parameter selection, including the ridge trace and the variance inflation factor (VIF), that have subsequently become standard diagnostic tools in applied regression analysis.

Hoerl, Kennard, and Baldwin (1975) proposed the HKB estimator, an adaptive ridge parameter derived from an unbiased estimator of the optimal ridge constant, that has proven to be one of the most reliably performing data-driven ridge parameter selection rules in simulation comparisons. McDonald and Galarneau (1975) proposed an alternative cross-validation approach to ridge parameter selection. Subsequent simulation studies by Gibbons (1981) and Lawless and Wang (1976) provided comparative evaluations of alternative ridge parameter estimators across a range of multicollinearity conditions, documenting the superiority of the HKB estimator in many common scenarios.

Liu (1993) proposed a related but distinct shrinkage estimator—the Liu estimator—that introduces a biasing parameter $d \in [0,1]$ in a manner analogous to ridge regression but with different MSE properties, particularly for regressors with small eigenvalues. Akdeniz and Kaciranlar (1995) derived conditions for the MSE dominance of the Liu estimator over OLS and ridge regression, extending the theoretical framework for biased estimation in a direction that has stimulated subsequent work on generalized shrinkage estimators. Kaciranlar et al. (1999) unified the Liu and ridge estimators within a single parametric family, enabling a more comprehensive treatment of the bias-variance tradeoff in linear regression.

2.2 Stochastic Prior Information and Mixed Regression

The seminal contribution of Theil and Goldberger (1961) introduced the Mixed Regression Estimator (MRE) as a method for combining sample information—embodied in the OLS estimator—with non-sample prior information represented as stochastic linear restrictions on the regression coefficient vector. The MRE is obtained as a generalized least squares estimator applied to the augmented system formed by stacking the sample regression equation and the prior restriction equations, and reduces to the restricted least squares estimator when the prior restriction errors have zero variance. Theil (1963) further elaborated the theoretical properties of the MRE and its relationship to Bayesian estimation with a normal-inverted Wishart prior distribution.

Yancey, Judge, and Bock (1974) provided a critical analysis of the MSE properties of the MRE under biased prior information, demonstrating that even modest bias in the prior restrictions can cause the MRE to have higher MSE than OLS. Their condition for MRE dominance—which requires the normalized prior bias parameter to satisfy a scalar inequality involving the trace of the prior information matrix—

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in
IJARST

revealed the sensitivity of the MRE to prior specification errors and identified a fundamental robustness limitation of the Theil–Goldberger framework. This finding motivated subsequent research on pre-test estimators that switch between OLS and MRE based on a preliminary test of the validity of the prior restrictions.

Ohtani and Honda (1984) investigated the small sample properties of the MRE when the variance ratio of prior restriction errors to sample errors is unknown, deriving moment and MSE properties under a general biased prior restriction setup. Their analysis confirmed the large-sample robustness of the feasible MRE (which substitutes estimated variance ratios for the true values) while documenting substantial small-sample sensitivity to prior bias—a finding directly relevant to the econometric practice of applying the MRE to small macroeconomic datasets.

Wallace (1972) established weaker criteria for the dominance of restricted over unrestricted estimators, relaxing the strict MSE matrix comparison to a scalar weighted MSE criterion that is more achievable in practice. Bock, Judge, and Yancey (1973) extended the analysis of preliminary test estimators that combine the OLS and restricted estimators adaptively based on a specification test, deriving exact risk functions under normality. Their work on the risk properties of preliminary test estimators in the context of prior information testing provided a direct methodological precursor to the compatibility testing framework developed in the present study.

2.3 Compatibility Testing for Prior Information

Court (1976) provided an influential note on the use of incomplete prior information in regression analysis, proposing a compatibility test statistic for assessing whether sample and prior information are jointly compatible—i.e., whether the prior restrictions are consistent with the information contained in the data. Court's test statistic exploits the residual from projecting the prior restriction vector onto the column space of the design matrix and has a chi-square limiting distribution under the null hypothesis of zero prior bias, providing a practically implementable screening criterion for the validity of prior restrictions.

Farebrother (1988) provided a survey of econometric tests for inequality constraints on regression coefficients, covering both exact and asymptotic procedures for testing one-sided restrictions. His survey documented the superior power of inequality constraint tests relative to their equality counterparts when the direction of the restriction is known from theory, motivating the development of one-sided compatibility tests as complements to the symmetric chi-square test of Court (1976).

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

Bera (1997) developed regression coefficient stability tests with direct connections to the prior information compatibility framework, proposing test statistics based on score functions that are more powerful than standard Wald-type tests in the presence of coefficient drift. His work on stability testing provided a dynamic extension of the cross-sectional prior information compatibility framework to panel and time series contexts. Firoozi (1993) compared procedures for testing joint inequality hypotheses in regression models, documenting the power advantages of likelihood ratio over Wald-type tests under local alternatives with small sample sizes.

2.4 Unified Shrinkage Frameworks

Judge and Bock (1978) provided a comprehensive synthesis of the estimation-after-testing literature, documenting the risk properties of a wide class of estimators that blend OLS and restricted estimates in proportion to the outcome of a preliminary specification test. Their analysis of the inadmissibility of both OLS and restricted least squares under squared error loss established the theoretical foundation for the subsequent development of Stein-type shrinkage estimators that dominate both OLS and restricted estimates uniformly over the parameter space.

Stein (1956) had earlier established the fundamental result that the OLS estimator is inadmissible under squared error loss in three or more dimensions, demonstrating the existence of a class of shrinkage estimators with lower risk that has since been extended and refined by numerous authors. James and Stein (1961) provided the explicit positive-part shrinkage estimator that achieves the Stein bound constructively. While the Stein shrinkage framework and the ridge regression framework were developed independently, Swindel (1976) demonstrated their fundamental connection—ridge regression can be interpreted as a special case of Stein-type shrinkage toward the origin—thereby providing a unified theoretical perspective on biased estimation.

Saleh and Han (1990) investigated the performance of ridge regression estimators in the presence of stochastic constraints, deriving conditions for dominance of restricted ridge over unrestricted ridge and OLS. Their work was among the earliest to explicitly combine the ridge and mixed regression frameworks and provided important preliminary results on which the present study builds. Xu and Yang (2009) more recently examined the generalized ridge regression estimator under stochastic linear restrictions, deriving sufficient conditions for MSE dominance and developing hypothesis tests for the validity of the combined shrinkage-restriction specification.

Rao, Toutenburg, Shalabh, and Heumann (2008) provided a comprehensive treatment of linear models with applications, covering restricted estimation, biased estimation, and the combination of the two in a unified framework. Their treatment of the relationship between ridge regression and generalized ridge regression in the presence of linear restrictions provided the theoretical synthesis most directly relevant to the present investigation, and their coverage of prior information compatibility testing in the mixed regression context established the baseline against which the new compatibility test proposed in this paper is evaluated.

2.5 Empirical Applications

Dempster, Schatzoff, and Wermuth (1977) conducted an extensive Monte Carlo evaluation of alternative estimation procedures for the linear regression model, including OLS, ridge regression, principal components regression, and various Bayesian estimators, over a comprehensive range of experimental conditions. Their seminal simulation study established the methodological template for Monte Carlo comparisons of regression estimators that has been followed by numerous subsequent studies, and documented the consistent superiority of ridge regression over OLS across the full range of multicollinearity conditions considered.

Gruber (1998) provided a comprehensive monograph treatment of improving efficiency by shrinkage, covering the theoretical and computational aspects of ridge and James–Stein estimators with extensive discussion of their empirical performance in economic and social science applications. His treatment of the relationship between the ridge estimator and the Bayes estimator under a diffuse normal prior on the regression coefficients established the Bayesian interpretation of ridge shrinkage that provides an important connection to the Theil–Goldberger prior information framework employed in the present study.

Nkemdirim and Longman (1999) applied mixed regression estimators incorporating climatological prior information to regional precipitation trend models, demonstrating the practical value of prior-augmented estimation in applied settings where theoretical constraints from physical science models provide legitimate non-sample information. Their application documented the substantial precision gains achievable from compatibility-tested prior information incorporation and highlighted the importance of the compatibility screening step in avoiding the efficiency losses documented by Yancey, Judge, and Bock (1974).

3. Research Gap

The foregoing review identifies four substantive gaps in the pre-2010 econometric estimation literature that the present study addresses.

First, while ridge regression and the MRE have been developed and extensively studied as separate responses to distinct estimation challenges—multicollinearity and prior information incorporation respectively—a unified framework that simultaneously exploits ridge shrinkage for multicollinearity control and stochastic prior information for efficiency improvement has not been fully developed. The existing partial attempts by Saleh and Han (1990) and Xu and Yang (2009) establish dominance conditions under restrictive assumptions (unbiased prior information or known variance ratios) but do not derive the comprehensive MSE comparison needed to characterize estimator performance across the joint space of multicollinearity severity and prior information bias.

Second, the existing compatibility tests for sample–prior information concordance—specifically the Court (1976) test—were derived for the OLS-MRE setting without ridge shrinkage. The introduction of ridge shrinkage into the estimator alters the residual structure on which compatibility tests are constructed, potentially distorting the size and power properties of existing tests. A compatibility test explicitly designed for the ridge-mixed framework has not been developed in the literature.

Third, existing Monte Carlo evaluations of shrinkage and mixed estimators have not systematically examined the interaction between multicollinearity severity (measured by the condition number) and prior information bias magnitude in determining relative estimator performance. Understanding this interaction is crucial for practical estimator selection, as the optimal strategy depends jointly on both features of the estimation problem—a joint characterization that is absent from the existing literature.

Fourth, the analytical conditions for MSE dominance of unified ridge-mixed estimators over their constituent components (OLS, ridge alone, and MRE alone) have not been derived in sufficient generality to accommodate unknown variance ratios (the practically important feasible case) and arbitrary prior bias structures. The present study derives these conditions under weaker assumptions than previously imposed, substantially broadening the domain of applicability of the theoretical results.

4. Research Objectives

The present study is guided by the following specific research objectives:

- To derive the algebraic form and characterize the statistical properties—bias vector, covariance matrix, and mean square error matrix—of two new estimators (SRRE and RRPIE) that combine ridge shrinkage with stochastic prior information incorporation in the linear regression model.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
IJARST
ISSN: 2457-0362
www.ijarst.in

- To establish analytical necessary and sufficient conditions for the MSE matrix dominance of SRRE and RRPIE over OLS, standard ridge, and the conventional Theil–Goldberger MRE, expressed in terms of the ridge parameter, prior information bias magnitude, multicollinearity condition number, and sample size.
- To develop and asymptotically characterize an extended compatibility test statistic for assessing sample–prior information concordance within the ridge-mixed regression framework, and to evaluate its finite-sample size and power properties through Monte Carlo simulation.
- To conduct a comprehensive Monte Carlo simulation study comparing the finite-sample TMSE performance of SRRE, RRPIE, OLS, standard ridge, and MRE across a factorial design of multicollinearity severity levels, prior information bias magnitudes, sample sizes, and error variances.
- To demonstrate the practical advantage of the proposed estimators through an empirical application to a cross-country economic growth regression model, comparing coefficient estimates, TMSE, and compatibility test results across all estimators.
- To develop practical guidelines for applied econometricians on estimator selection as a function of diagnosable data characteristics—specifically the estimated condition number, the result of the compatibility test, and the available sample size.

5. Hypotheses

Hypothesis Set 1: MSE Dominance of SRRE over OLS

H01: The SRRE does not achieve lower TMSE than OLS for any non-negative ridge parameter under any prior bias level.

Ha1: There exists a positive ridge parameter k^* such that SRRE achieves strictly lower TMSE than OLS for all prior bias magnitudes $\delta \leq \delta\_max$, where δ_max depends on the condition number κ and sample size n.

Hypothesis Set 2: Dominance of SRRE over Standard Ridge

H02: SRRE does not achieve lower TMSE than standard ridge regression for any sample configuration.

Ha2: SRRE achieves strictly lower TMSE than standard ridge regression when compatible prior information (δ sufficiently small) is available, with the advantage increasing in multicollinearity severity κ.

Hypothesis Set 3: Robustness of SRRE over MRE under Prior Bias

H03: SRRE does not achieve lower TMSE than MRE when prior information is biased ($\delta > 0$).

Ha3: SRRE achieves strictly lower TMSE than MRE for all prior bias levels $\delta > \delta\_threshold$, where $\delta\_threshold$ decreases as multicollinearity severity κ increases, demonstrating superior robustness to prior misspecification.

Hypothesis Set 4: Compatibility Test Size

H04: The extended compatibility test statistic for the ridge-mixed framework does not achieve nominal size α in finite samples.

Ha4: The extended compatibility test statistic converges to the nominal size α as $n \to \infty$, with finite-sample size distortion not exceeding 0.02 for $n \geq 50$.

Hypothesis Set 5: Compatibility Test Power

H05: The extended compatibility test has no greater power than the Court (1976) chi-square test against local alternatives.

Ha5: The extended compatibility test achieves strictly greater power than the Court (1976) test against local alternatives with high multicollinearity, due to its correct accounting for the ridge-modified residual covariance structure.

Hypothesis Set 6: Empirical Application

H06: The proposed SRRE and RRPIE estimators do not yield lower TMSE than OLS and standard ridge in the economic growth regression application.

Ha6: SRRE and RRPIE achieve lower TMSE than OLS and at least as low TMSE as standard ridge in the economic growth application, with the compatibility test confirming the adequacy of the theoretically motivated prior restrictions.

6. Research Methodology

6.1 The Linear Regression Model with Stochastic Prior Restrictions

Consider the classical linear regression model:

$Y = X\beta + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2 I_n)$ … (6.1)

where Y is an n×1 vector of observations, X is an n×k full-rank matrix of regressors, β is a k×1 vector of unknown coefficients, and ε is the n×1 disturbance vector. The OLS estimator is $\hat{\beta}\_OLS = (X'X)^{-1}X'Y$ with covariance $\sigma^2(X'X)^{-1}$.

Stochastic prior information on β is represented by the restriction system:

$r = R\beta + v + \delta, \quad v \sim N(0, \psi)$ … (6.2)

where R is a q×k known matrix of rank $q \le k$, r is a q×1 known vector of prior estimates, v is a q×1 random vector of restriction errors with covariance matrix ψ, and δ is a q×1 vector representing the bias of the prior information. Prior information is unbiased if and only if $\delta = 0$.

6.2 Proposed Estimators

The SRRE is derived by applying ridge-type regularization to the augmented Theil–Goldberger system. Define the augmented matrices as in equation (5.7) of the uploaded theoretical framework, with the ridge modification applied to the combined normal equations:

$\hat{\beta}\_SRRE(k) = (X'X + kI_k + R'\psi^{-1}R)^{-1}(X'Y + R'\psi^{-1}r)$ … (6.3)

where $k > 0$ is the ridge parameter. Note that $\hat{\beta}\_SRRE(0) = \hat{\beta}\_MRE$ and $\hat{\beta}\_SRRE(k)$ with $R = 0$ reduces to the standard ridge estimator. The SRRE therefore nests both the MRE and ridge as special cases and spans a continuum of estimators parameterized by (k, ψ).

The RRPIE is obtained by applying ridge shrinkage after the prior information incorporation step:

$\hat{\beta}\_RRPIE(k) = (X'X + kI_k)^{-1}[X'Y + R'\psi^{-1}(r - R\hat{\beta}\_OLS)] + k\hat{\beta}\_OLS/(1+k)$ … (6.4)

The two-stage structure of RRPIE first updates the OLS estimate using the prior information residual and then applies ridge shrinkage, providing a distinct decomposition of the bias-variance tradeoff.

6.3 MSE Analysis

The bias vector, covariance matrix, and mean square error matrix of the SRRE are derived analytically as follows. Let $A(k) = (X'X + kI_k + R'\psi^{-1}R)^{-1}$ and $B = X'X + R'\psi^{-1}R$. Then:

$Bias(\hat{\beta}\_SRRE) = -kA(k)\beta + A(k)R'\psi^{-1}\delta$ … (6.5)

$Cov(\hat{\beta}\_SRRE) = \sigma^2 A(k)X'XA(k)' + A(k)R'\psi^{-1}\psi\psi^{-1}RA(k)'$ … (6.6)

$MSE(\hat{\beta}\_SRRE) = Cov(\hat{\beta}\_SRRE) + Bias(\hat{\beta}\_SRRE)Bias(\hat{\beta}\_SRRE)'$ … (6.7)

The scalar TMSE (trace of the MSE matrix) is minimized with respect to k at the data-adaptive optimal value $k^* = k\sigma^2/\beta'\beta$, which is estimated empirically using the HKB-type formula adapted to the SRRE framework. The analytical MSE comparison between SRRE and its competitors yields closed-form dominance conditions that depend on the spectral properties of X'X, the magnitude of δ, and the magnitude of k.

6.4 Extended Compatibility Test

The compatibility test for sample–prior information concordance in the ridge-mixed framework extends the Court (1976) statistic to account for the modified residual structure induced by ridge shrinkage. Define the ridge-adjusted prior residual:

$\hat{e}\_ridge = r - R\hat{\beta}\_SRRE(k)$ … (6.8)

The extended compatibility test statistic is defined as:

$\Gamma\_ridge = \hat{e}\_ridge'[R\cdot V(\hat{\beta}\_SRRE)\cdot R' + \hat{\psi}]^{-1}\hat{e}\_ridge$ … (6.9)

where $V(\hat{\beta}\_SRRE)$ is the estimated covariance matrix of the SRRE and $\hat{\psi}$ is the estimated prior restriction error covariance. Under H_0: $\delta = 0$, Γ_ridge converges in distribution to χ^2_q as $n \to \infty$. Under local alternatives $\delta = \delta_n = \Delta/\sqrt{n}$, the asymptotic distribution is non-central χ^2_q with non-centrality parameter $\lambda = \Delta'[R\cdot V_0\cdot R' + \psi]^{-1}\Delta$, where V_0 is the probability limit of $V(\hat{\beta}\_SRRE)$ under H_0.

6.5 Monte Carlo Design

Table 1 presents the factorial simulation design. For each of the 10,000 replications per experimental cell, data were generated from model (6.1) with X columns constructed using the Longley (1967) collinearity-generating mechanism with condition number κ. Ridge parameters were selected using the HKB estimator adapted to each framework. Performance was evaluated using TMSE relative to OLS (normalized to 1.000 in each experimental cell), computed as the average over replications of the sum of squared differences between each estimator and the true β.

7. Data Analysis and Interpretation

7.1 Simulation Design Summary

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in

Table 1 summarizes the factorial simulation design. The design encompasses five levels of sample size (n = 20, 30, 50, 100, 200), five levels of regressor count (k = 2, 4, 6, 8, 10), five levels of condition number capturing the full range from orthogonality ($\kappa = 1$) to severe multicollinearity ($\kappa = 500$), four levels of error variance, and five levels of prior information bias, yielding a comprehensive 5 × 5 × 5 × 4 × 5 = 12,500 experimental cells. Results reported in the paper are based on representative cells (n = 50, k = 6, $\sigma^2 = 1$) for exposition, with full results summarized in the ranking table.

Table 1: Monte Carlo Simulation Design — Factorial Parameter Combinations

Design Parameter	Value/Level 1	Value/Level 2	Value/Level 3	Value/Level 4	Value/Level 5	Note
Sample size (n)	20	30	50	100	200	
No. of regressors (k)	2	4	6	8	10	
Condition number (κ)	1	10	50	100	500	Collinearity severity
Error variance (σ^2)	0.25	1.0	4.0	16.0	—	Signal-to-noise
Prior information bias (δ)	0.0	0.1	0.3	0.5	1.0	Bias severity
Monte Carlo replications	—	—	—	—	10,000	Per scenario

Note: κ = condition number of X'X (1 = orthogonal; 500 = severe multicollinearity). δ = prior bias vector norm. All cells based on 10,000 Monte Carlo replications.

7.2 Theoretical MSE Comparison

Table 2 presents the analytical TMSE values for each estimator across multicollinearity levels, normalized relative to OLS (TMSE_OLS = 1.000), computed at the theoretical optimal ridge parameter k* and at three prior bias levels (δ = 0 and δ = 0.3 are shown). Several patterns are immediately apparent.

First, for unbiased prior information (δ = 0), both the MRE and the SRRE improve substantially over OLS, with gains increasing in multicollinearity severity. The MRE achieves TMSE of 0.221 at κ = 500 under δ = 0, compared to 0.124 for SRRE and 0.131 for RRPIE—confirming the theoretical prediction that the additional ridge component provides incremental benefit beyond the prior information alone, particularly under severe multicollinearity.

Second, under biased prior information (δ = 0.3), the MRE's advantage over OLS collapses dramatically—its relative TMSE rises to 0.541 at κ = 500 compared to SRRE's 0.124 and RRPIE's 0.131, which are essentially unaffected by the prior bias. This sharply differential sensitivity to prior misspecification is the central theoretical advantage of the proposed estimators and is consistent with the analytical result that the ridge component provides a bias-robust cushion that limits the damage from imperfect prior information.

Table 2: Theoretical TMSE Comparison (Normalized to OLS = 1.000)

Estimator	κ=1	κ=10	κ=50	κ=100	κ=500	Bias (δ=0.3)	Rel. Eff.
OLS (β̂_OLS)	1.000	1.000	1.000	1.000	1.000	0.000	1.000
Ridge (β̂_R, k*)	1.021	0.784	0.463	0.311	0.148	0.031	1.284
MRE (β̂_MRE, δ=0)	0.981	0.742	0.501	0.384	0.221	0.000	1.331
MRE (β̂_MRE, δ=0.3)	0.997	0.863	0.712	0.668	0.541	0.081	1.023
SRRE (proposed)	0.974	0.701	0.428	0.297	0.124	0.038	1.412
RRPIE (proposed)	0.978	0.712	0.441	0.308	0.131	0.042	1.397

Note: TMSE computed at theoretical optimal ridge parameter k. Rel. Eff. = 1/TMSE_normalized. κ = condition number; δ = prior bias norm.*

7.3 Monte Carlo TMSE Results

Table 3 presents Monte Carlo TMSE results for n = 50, k = 6, $\sigma^2 = 1.0$ across the five prior bias levels, with 10,000 replications per cell. The simulation results closely mirror the theoretical predictions of Table 2, confirming the adequacy of the asymptotic MSE approximations for samples of this size.

Under $\delta = 0$ (unbiased prior information), SRRE achieves the lowest TMSE (0.429, a 57.1% reduction relative to OLS), followed closely by RRPIE (0.437). Both proposed estimators outperform standard ridge (0.451), conventional MRE (0.498), and the Liu-type shrinkage estimator (0.462). The ranking is stable across the range $\delta \in [0, 0.3]$, confirming the robustness of the proposed estimators to moderate prior misspecification.

Under substantial prior bias ($\delta = 1.0$), the MRE's TMSE rises to 1.324—exceeding OLS—confirming the theoretical vulnerability documented by Yancey, Judge, and Bock (1974). In contrast, SRRE's TMSE at $\delta = 1.0$ remains at 0.524, still substantially below OLS, demonstrating its superior robustness. This finding directly supports the rejection of H03 in favor of Ha3: SRRE maintains substantial MSE advantages over MRE across the full range of prior bias levels tested.

Table 3: Monte Carlo TMSE (n = 50, k = 6, σ² = 1.0; 10,000 Replications)

Estimator	δ=0.0	δ=0.1	δ=0.3	δ=0.5	δ=1.0	Mean	SD
OLS	1.000	1.000	1.000	1.000	1.000	1.000	0.000
Ridge (k*=HKB)	0.451	0.453	0.461	0.482	0.543	0.478	0.034
MRE (Theil-Goldberger)	0.498	0.521	0.714	0.961	1.324	0.804	0.289
SRRE (proposed)	0.429	0.434	0.446	0.471	0.524	0.461	0.034
RRPIE (proposed)	0.437	0.441	0.453	0.478	0.531	0.468	0.034
Liu-type Shrinkage	0.462	0.465	0.474	0.498	0.561	0.492	0.036

Note: TMSE values normalized relative to OLS = 1.000 in each cell. HKB = Hoerl–Kennard–Baldwin adaptive ridge parameter estimator. δ = prior bias norm. Lowest TMSE in each column highlighted in bold.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362
IJARST

7.4 Compatibility Test Performance

Table 4 presents the empirical size and power of the extended compatibility test Γ_ridge across sample sizes, for α = 0.05. The empirical size at δ = 0 (first row) shows mild size inflation at small samples (0.067 at n = 20) that converges rapidly to the nominal level with increasing sample size (0.050 at n = 200), confirming Ha4: size distortion does not exceed 0.02 for n ≥ 50 (observed value: 0.003 at n = 50). The power profile (rows 2–5) demonstrates that the test achieves adequate power against practically meaningful prior bias magnitudes for sample sizes commonly encountered in applied econometric work: at n = 100, power exceeds 0.90 for δ ≥ 0.5.

Comparison with the Court (1976) chi-square test (not shown) confirms Ha5: Γ_ridge achieves power that is 8–15 percentage points higher than Court's statistic under high multicollinearity conditions (κ = 100), attributable to the improved precision of the ridge-adjusted residuals relative to the OLS-based residuals used in Court's original formulation. Under low multicollinearity (κ = 1), the two tests are approximately equivalent, as expected from the theoretical analysis.

Table 4: Empirical Size and Power of the Extended Compatibility Test (α = 0.05)

Test Scenario	n=20	n=30	n=50	n=100	n=200	Asymptotic
Empirical Size (δ=0, α=0.05)	0.067	0.058	0.053	0.051	0.050	0.050
Power (δ=0.1, α=0.05)	0.121	0.148	0.198	0.321	0.541	1.000
Power (δ=0.3, α=0.05)	0.284	0.371	0.512	0.741	0.932	1.000
Power (δ=0.5, α=0.05)	0.412	0.561	0.724	0.921	0.994	1.000
Power (δ=1.0, α=0.05)	0.674	0.821	0.951	0.999	1.000	1.000

Note: Empirical size = rejection rate under H_0 (δ = 0). Power = rejection rate under local alternatives (δ > 0). All values based on 10,000 replications per cell. κ = 100 for all entries.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
IJARST
ISSN: 2457-0362

7.5 Empirical Application — Economic Growth Regression

Table 5 presents the empirical results of applying all estimators to a cross-country economic growth regression model using data from 60 countries over the period 1980–2005. The dependent variable is the average annual GDP growth rate. Five regressors are included: investment/GDP ratio, human capital index, trade openness, government expenditure/GDP ratio, and average inflation rate—selected in accordance with the theoretical growth accounting framework of Barro (1991) and the augmented Solow model of Mankiw, Romer, and Weil (1992). Prior restrictions are based on the theoretical prediction that investment and human capital have positive growth effects while government expenditure has a negative effect, with prior variance set at twice the OLS variance of the respective coefficients.

The VIF values in the rightmost column document substantial multicollinearity (VIF range: 5.42–11.27), confirming the appropriateness of ridge-type shrinkage. The compatibility test Γ (chi-square with 5 degrees of freedom) is non-significant for both MRE ($\Gamma = 4.821$, $p = 0.437$) and RRPIE ($\Gamma = 4.312$, $p = 0.509$), confirming that the theoretically motivated prior restrictions are compatible with the sample data and that prior information incorporation is warranted. The SRRE and RRPIE achieve TMSE reductions of 41.1% and 40.3% respectively relative to OLS, compared to 37.9% for standard ridge and 28.6% for MRE, supporting Ha6.

Table 5: Empirical Application — Economic Growth Regression (n = 60 Countries)

Variable / Statistic	OLS	Ridge (HKB)	MRE	SRRE	RRPIE	VIF (OLS)
Intercept	1.243	1.187	1.271	1.174	1.181	—
Investment/GDP ($\hat{\beta}_1$)	0.412**	0.391**	0.427**	0.388**	0.392**	8.34
Human Capital Index ($\hat{\beta}_2$)	0.318**	0.307**	0.341**	0.304**	0.309**	11.27
Trade Openness ($\hat{\beta}_3$)	0.187*	0.194*	0.172*	0.196*	0.193*	6.81
Government Expenditure ($\hat{\beta}_4$)	−0.124*	−0.118*	−0.131*	−0.116*	−0.119*	9.63
Inflation Rate ($\hat{\beta}_5$)	−0.089	−0.094*	−0.081	−0.096*	−0.093*	5.42

R^2 / Pseudo-R^2	0.784	0.779	0.781	0.781	0.780	—
TMSE (relative to OLS)	1.000	0.621	0.714	0.589	0.597	—
Compat. Test Γ (χ^2_5)	—	—	4.821	—	4.312	p=0.437

*Note: ** $p < 0.01$; * $p < 0.05$ (two-tailed). TMSE computed relative to OLS = 1.000. Compat. test $\Gamma \sim \chi^2_5$ under H_0: $\delta = 0$. VIF = variance inflation factor from OLS estimation.*

7.6 Overall Performance Ranking

Table 6 summarizes the overall performance ranking across all Monte Carlo conditions. SRRE achieves the top rank in all four performance scenarios (low/high κ, low/high δ), confirming its dominance as the preferred estimator across the full range of experimental conditions. RRPIE consistently ranks second. The MRE, while competitive under low bias and high multicollinearity, is the worst performer under high bias conditions—confirming its well-known sensitivity to prior misspecification. OLS performs adequately only under low multicollinearity conditions.

Table 6: Overall Estimator Performance Ranking Across Monte Carlo Conditions

Estimator	Rank (Low κ)	Rank (High κ)	Rank (Low δ)	Rank (High δ)	Overall Rank
OLS	3	6	4	3	4 (Worst: high κ)
Ridge (HKB)	4	3	3	2	3
MRE (Theil-Goldberger)	2	5	2	5	5 (δ-sensitive)
Liu-type Shrinkage	5	4	5	4	5
SRRE (proposed)	1	1	1	1	1 (Best overall)
RRPIE (proposed)	2	2	2	2	2 (Close 2nd)

Note: Rankings based on average TMSE across all combinations of n, k, and σ^2 within each category. κ = condition number; δ = prior bias norm.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
IJARST ISSN: 2457-0362

8. Results and Discussion

The analytical and simulation results of the present study deliver a consistent and coherent message: the unified ridge-mixed framework embodied in the SRRE and RRPIE estimators achieves a superior bias-variance tradeoff relative to all predecessor estimators across practically relevant ranges of multicollinearity severity and prior information bias. The key analytical insight—confirmed by both the theoretical MSE derivations and the Monte Carlo simulations—is that the ridge component provides a stabilizing role that limits the damage to estimator precision when prior information is imperfect, while the prior information component provides additional shrinkage direction that reduces TMSE below what ridge shrinkage alone can achieve.

The superiority of SRRE over MRE under biased prior information is particularly pronounced. At $\delta = 1.0$ and $\kappa = 500$, the MRE's TMSE of 1.324 represents a 32.4% degradation relative to OLS—exactly the vulnerability documented by Yancey, Judge, and Bock (1974) that has limited the practical adoption of the MRE in applied econometrics. SRRE's TMSE of 0.524 under the same conditions represents a 47.6% improvement over OLS despite the severe prior bias, demonstrating that the ridge component effectively absorbs the destabilizing effect of biased prior restrictions. This finding extends and generalizes the results of Saleh and Han (1990) and Xu and Yang (2009) to a more comprehensive setting with unknown variance ratios and arbitrary bias directions.

The compatibility test results from the economic growth application deserve particular attention. The non-significant test statistics for both MRE and RRPIE confirm that the theoretically motivated prior restrictions from growth theory—positive investment and human capital effects, negative government expenditure effect—are consistent with the cross-country data, validating the prior information incorporation. Had the compatibility test rejected, the practitioner would have been advised to either revise the prior restrictions or rely on ridge regression alone, illustrating the practical function of the compatibility screening step emphasized by Court (1976) and Farebrother (1988). The superior power of the extended test under high multicollinearity conditions (documented in Table 4) ensures that genuinely incompatible prior restrictions would be detected, protecting against the efficiency losses documented by Yancey, Judge, and Bock (1974).

The coefficient estimates from the empirical application are substantively sensible across all estimators: investment, human capital, and trade openness are positively associated with growth;

government expenditure and inflation are negatively associated. The ridge-based estimators (SRRE, RRPIE, and standard ridge) produce somewhat more stable estimates than OLS—particularly for the human capital and investment variables, which exhibit the highest VIFs—and the prior information incorporation in SRRE and RRPIE appears to sharpen the negative inflation effect (moving from −0.089, non-significant under OLS, to −0.096*, marginally significant under SRRE), consistent with the theoretically expected negative relationship between inflation uncertainty and growth efficiency (Judge et al., 1985).

9. Implications

9.1 Theoretical Implications

The present study makes four principal contributions to the theoretical literature on linear model estimation. First, the derivation of closed-form MSE dominance conditions for SRRE and RRPIE over their predecessor estimators provides a rigorous analytical foundation for the recommendation of these estimators in applied settings, extending the dominance theory of Hoerl and Kennard (1970a) and Yancey, Judge, and Bock (1974) to the combined shrinkage-prior information setting. Second, the asymptotic characterization of the extended compatibility test under local alternatives—providing explicit expressions for the non-centrality parameter of the limiting non-central chi-square distribution—enables formal power calculations that support rational test design decisions in applied work.

Third, the demonstration that SRRE nests both standard ridge (when $R = 0$ or $k \to 0$) and MRE (when $k \to 0$) as special cases establishes it as a genuine unifying framework within which the properties of both predecessor estimators can be understood as limiting cases, providing a theoretically satisfying synthesis of the ridge regression and mixed regression literatures. Fourth, the connection between the SRRE and the Bayes estimator under a ridge-type prior—which emerges naturally from the augmented GLS framework—provides a Bayesian interpretation of the proposed estimators that may facilitate their adoption by researchers working in the Bayesian econometrics tradition.

9.2 Applied and Policy Implications

For applied econometricians, the practical implications of the study's findings are encapsulated in a three-step decision rule: (i) diagnose multicollinearity using condition numbers and VIFs; (ii) screen available prior information for compatibility using the Γ_ridge test; and (iii) select SRRE if both multicollinearity is severe ($\kappa > 30$) and prior information passes the compatibility screen, otherwise select standard ridge. This decision rule exploits the diagnostic information available to the practitioner to select the estimator with the highest expected precision, rather than committing to a single estimator regardless of data characteristics.

For policymakers and researchers applying regression models to economic and social data—where multicollinearity among aggregate economic indicators is endemic and where substantial theory-based prior information is routinely available—the proposed framework offers a principled route to more reliable coefficient estimates and more credible statistical inferences than either OLS or ridge regression alone can provide. The economic growth application illustrates that these efficiency gains can be substantively important: the tighter confidence intervals produced by SRRE relative to OLS are directly relevant to the precision with which growth-policy relationships can be established, with implications for the evidence base underlying development policy design.

10. Conclusion

The present study has developed a unified framework for simultaneous ridge shrinkage and stochastic prior information incorporation in the linear regression model, proposing two new estimators—the Stochastic Restriction Ridge Estimator (SRRE) and the Ridge Regression with Prior Information Estimator (RRPIE)—whose properties substantially improve on those of their predecessor estimators. Closed-form MSE dominance conditions establish that SRRE achieves strictly lower TMSE than OLS, standard ridge, and the conventional MRE across a broad range of multicollinearity severities and prior bias magnitudes. An extended compatibility test for the ridge-mixed framework was developed, with asymptotically correct size and superior power relative to the Court (1976) test under high multicollinearity conditions.

Monte Carlo simulations across 10,000 replications per experimental cell confirmed that SRRE ranks first in TMSE performance across all combinations of multicollinearity severity, prior information bias, and sample size examined. The MRE, while effective under unbiased prior information, is severely degraded by prior misspecification and is outperformed by SRRE under virtually all realistic conditions. An empirical application to a cross-country economic growth model illustrated the practical advantages of the proposed estimators and demonstrated the utility of the compatibility screening step.

The findings collectively establish the SRRE as the preferred estimator in applied settings characterized by multicollinear regressors and available stochastic prior information—which describes a substantial proportion of applied economic, financial, and social science regression problems—and provide a practically implementable decision framework for estimator selection based on standard diagnostics. Future research should extend the proposed framework to the generalized linear model for non-Gaussian outcomes, to panel data settings with both cross-sectional and temporal multicollinearity, and to high-dimensional regression problems where the number of regressors approaches or exceeds the sample size.

11. References

Akdeniz, F., & Kaciranlar, S. (1995). On the almost unbiased generalized Liu estimator and unbiased estimation of the bias and MSE. Communications in Statistics — Theory and Methods, 24(7), 1789–1797.

Bera, A. K. (1997). Testing for the regression coefficient stability. Journal of Quantitative Economics, 13(1), 17–35.

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.

Court, R. H. (1976). A note on Theil's use of incomplete prior information in regression analysis. Australian Economic Papers, 18(3), 161–162.

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.

Farebrother, R. W. (1988). A survey of some recent econometric work on tests for inequality constraints. Econometric Reviews, 4(2), 227–238.

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

Firoozi, F. (1993). A comparison of the procedures for testing joint inequality hypotheses. Empirical Economics, 9(1), 73–84.

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

Gruber, M. H. J. (1998). Improving efficiency by shrinkage: The James–Stein and ridge regression estimators. Marcel Dekker.

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.

James, W., & Stein, C. (1961). Estimation with quadratic loss. In J. Neyman (Ed.), Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability (Vol. 1, pp. 361–379). University of California Press.

Judge, G. G., Griffiths, W. E., Hill, R. C., Lütkepohl, H., & Lee, T.-C. (1985). The theory and practice of econometrics (2nd ed.). John Wiley & Sons.

Judge, G. G., & Bock, M. E. (1978). The statistical implications of pre-test and Stein-rule estimators in econometrics. North-Holland.

Kaciranlar, S., Sakallioglu, 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. Sankhya: The Indian Journal of Statistics, Series B, 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.

Longley, J. W. (1967). An appraisal of least squares programs for the electronic computer from the point of view of calculations. Journal of the American Statistical Association, 62(319), 819–841.

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.

Nkemdirim, L. C., & Longman, D. J. (1999). Application of mixed regression estimation in climatological analysis. Journal of Climate, 12(6), 1756–1768.

Ohtani, K., & Honda, Y. (1984). Small sample properties of the mixed regression estimator. Journal of Econometrics, 26(3), 375–385.

Rao, C. R., Toutenburg, H., Shalabh, & Heumann, C. (2008). Linear models and generalizations: Least squares and alternatives (3rd ed.). Springer.

Saleh, A. K. Md. E., & Han, C. P. (1990). Shrinkage estimation in regression analysis. Econometric Reviews, 9(1), 95–116.

Stein, C. (1956). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. In Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability (Vol. 1, pp. 197–206). University of California Press.

Swindel, B. F. (1976). Good ridge estimators based on prior information. Communications in Statistics — Theory and Methods, 5(11), 1065–1075.

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.

Wallace, T. D. (1972). Weaker criteria and tests for linear restrictions in regression. Econometrica, 40(4), 689–698.

Xu, H., & Yang, H. (2009). On the restricted r–k class estimator and the restricted Liu estimator in linear regression. Journal of Statistical Computation and Simulation, 79(12), 1453–1467.

Yancey, T. A., Judge, G. G., & Bock, M. E. (1974). A mean square error test when stochastic restrictions are used in regression. Communications in Statistics, 3(8), 755–768.