PREDICTIONS IN LINEAR REGRESSION MODEL WITH MEASUREMENT ERRORS

Anil K. Srivastava and Shalabh*

I. INTRODUCTION

Prediction is an important aspect of the scientific study of any phenomenon. In the traditional framework of regression analysis, the predictions are obtained either for the average value of the study variable or for its actual value. This has limited utility in the sense that either of the two predictions are not all that enough in general. Situations are not uncommon where we need to consider both these predictions simultaneously. For example, in a consumer market the life guarantee of an item on sale is of paramount importance. The promoter of the item concerned claims the guarantee on the basis of its average life prediction while the consumer takes decision regarding its purchase on the basis of its actual life prediction. The classical theory of prediction can safeguard the interest of either the promoter or the consumer at a time but can not vouchsafe for both simultaneously. This can be achieved by defining a target function of the average and the actual value predictions of the study variable with possibly different weights assigned to them. This composite function would take care of the two predictions simultaneously and would allow us to assign desired weightage to them according to their respective importance in any given application; see, for example, Zellner (1994).

Another major limitation of the classical least squares theory of prediction is that it is based on the assumption that the observations on the variables involved in the model are free from measurement errors. In practice, however,

Key-words and Phrases: Linear Regression Model; Measurement Errors; Ultra-structural Model, Structural and Functional Forms; Target Function; Least Squares Predictor and Immaculate Predictor.

*Department of Statistics, Lucknow University, Lucknow, India.
this requirement is hardly ever met. The measurement error is quite common in practice due to many reasons. For example, in certain applications, it becomes difficult to quantify the variables in a desired manner such as market conditions. On the other hand, the use of proxy variables and imperfect measurement techniques may sometimes cause the error in observations. When the measurement errors are small enough so as to become negligible, their presence does not alter the inferences deduced from regression analysis much. But when they are not so, their presence does disturb the quality of inferences appreciably which often leads to invalid conclusions.

The least squares predictor under the assumption of observations being free from measurement errors, is well armed to provide the best linear and unbiased predictions in the classical linear regression model. But when measurement errors creep in, the least squares predictor becomes not only biased but inconsistent also. With this in mind, we have attempted to purge it for contamination arising due to measurement errors by applying corrections to the least squares estimates of the regression parameters involved in the predictor; see, for example, Fuller (1987). This provides an immaculate predictor which is found to be consistent. Study of asymptotic properties of this predictor is the subject matter of this article.

The plan of the paper is as follows. In Section 2, we describe the measurement error model and its variants, and propose a target function for the average and actual value prediction of the study variable in a composite manner which allows weightage to these predictions according to practitioner’s choice at will. We then consider attenuation of the least squares predictor for this target function to purge it of the contamination prevailing in it due to measurement errors and propose an immaculate predictor. Properties of this predictor are then analysed in Section 3 and finally some remarks are placed in Section 4.

II. THE SPECIFICATION OF FRAMEWORK AND TARGET PREDICTORS

Let us consider the linear regression model which connects the study variable \( Y_i \) and an explanatory variable \( X_i \) as follows

\[
Y_i = \alpha + \beta X_i \quad ; \quad i = 1, 2, \ldots, n, \quad (2.1)
\]

where \( \alpha \) is the intercept term and \( \beta \) is the slope parameter. The study and the explanatory variables both are assumed to be contaminated with measurement errors such that the observed values of \( Y_i \) and \( X_i \) are respectively given by

\[
Y_i = Y_i + u_i \quad (2.2)
\]
where $u_i$ and $v_i$ are respective measurement error terms associated with study variable $Y_i$ and the explanatory variable $X_i$.

It is assumed that the conventional disturbance term in the regression relationship is subsumed in the error term $u_i$ and $u_1, u_2, \ldots, u_n$ are independently and identically distributed with mean 0 and variance $\sigma_u^2$. Similarly, $v_1, v_2, \ldots, v_n$ are also assumed to be distributed independently and identically with mean 0 and variance $\sigma_v^2$. Further, we assume that $X_1, X_2, \ldots, X_n$ have possibly unequal means, say, $m_1, m_2, \ldots, m_n$ respectively with same variance so that $X_i$'s can be expressed as

$$X_i = m_i + w_i$$

(2.4)

where $w_i$ is a random variable with mean 0 and variance $\sigma_w^2$. Finally, we assume that $u_i, v_i$ and $w_i$ are stochastically independent of each other.

This completes specification of the ultra-structural model. It reduces to the functional variant of the measurement error model when $X_1, X_2, \ldots, X_n$ are fixed rather than random. Similarly, it reduces to the structural variant of the measurement error model when $X_1, X_2, \ldots, X_n$ are random with same mean and the same variance; see, for example, Dolby (1976).

Writing

$$d = \frac{\sigma_w^2}{(S_{mn} + \sigma_w^2)}; \quad S_{mn} = \frac{1}{n} \sum_{i=1}^{n} (m_i - \bar{m})^2$$

(2.5)

$$\theta = \frac{\sigma_v^2}{(S_{mn} + \sigma_w^2 + \sigma_v^2)}$$

(2.6)

where $\bar{m}$ is the mean of $m_1, m_2, \ldots, m_n$, we observe that $d = 0$ for the functional model and $d = 1$ for the structural model. Thus, $d$ provides a measure of departure of ultra-structural model from its two extremes, viz., the functional and the structural forms.

Similarly, we observe that $\theta = 0$ when there is no measurement error involved in the model, that is, when the given model is classical linear regression model. Thus, a non-zero value of $\theta$ provides a measure of departure of the ultra-structural model from the classical one.

We now consider the problem of simultaneous prediction of the average value $E(Y_k)$ and the actual value $y_k$ of the study variable and propose the following target function
with \( \lambda \) as a weighing constant lying between 0 and 1. For a specific choice of \( \lambda = 0\), this target function provides the average value prediction of the study variable. Similarly, for \( \lambda = 1\), it provides the actual value prediction. For any other value of \( \lambda \) between 0 and 1, it considers simultaneous prediction of the average and actual values both with desired weightage of our choice which may be governed by the nature of the problem and practitioner’s preference.

Assuming that \( Y_t \) and its counterpart \( X_t \) satisfy all the specifications of the measurement error model, it is natural to employ the following predictor for \( Y_t \) and its expected value \( E(Y_t) \) both

\[ p = \alpha + \beta x_k \]  

(2.8)

provided the regression parameters \( \alpha \) and \( \beta \) are known. Unfortunately, \( \alpha \) and \( \beta \) are hardly ever known in practice and that renders this predictor inoperational. A simple solution to make it feasible is to replace the unknown \( \alpha \) and \( \beta \) in \( p \) by their respective least squares estimates. Thus employing

\[ \hat{\beta}_c = \frac{s_{xy}}{s_{xx}} \]  

(2.9)

\[ \hat{\alpha}_c = \bar{y} - \hat{\beta}_c \bar{x} \]  

(2.10)

with

\[ s_{xx} = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2; \quad \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i, \]

(2.11)

\[ s_{xy} = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}); \quad \bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i \]

in (2.8), we obtain the following feasible predictor

\[ \hat{p}_c = \bar{y} + \hat{\beta}_c (x_k - \bar{x}) \]  

(2.12)

for \( Y_t \) and \( E(Y_t) \) both.

Employing this \( \hat{p}_c \) for \( Y_t \) and \( E(Y_t) \) in (2.7), we get the predictor for the target function \( P \) which incidently turns out to be the same as \( \hat{p}_c \) itself because of the convex nature of the target function. Thus, \( \hat{p}_c \) can be regarded as a feasible predictor for the target function \( P \).
A basic problem with this predictor is that it is inconsistent as the least squares estimates $\hat{\alpha}_c$ and $\hat{\beta}_c$ of the regression parameters in errors in variables setup come out to be inconsistent. If, however, some additional information is available which can be incorporated in estimation procedure, these regression parameters can be estimated consistently; see for instance, Judge et al. (1985, Chapter 13). Following this, let us suppose that $\sigma^2_c$ is known apriori. This information is often available through the past experiences of long association with similar kinds of experiments, or at times, it is suggested by some theoretical considerations also. Utilizing this information in the estimation procedure, we can purge the least squares estimates, $\hat{\alpha}_c$ and $\hat{\beta}_c$ of the involved contamination due to measurement errors in observations and obtain immaculate estimates to these regression parameters which are not inconsistent. They are evaluated as

\[
\begin{align*}
\hat{\beta}_{\text{purged}} &= b_t = \frac{S_{xy}}{(S_{xx} - \sigma^2_c)} \\
\hat{\alpha}_{\text{purged}} &= a_t = \bar{y} - b_t \bar{x}.
\end{align*}
\]

Replacing unknown $\alpha$ and $\beta$ in (2.8) by these immaculate estimates $a_t$ and $b_t$ respectively, we obtain the following feasible predictor

\[
\hat{\rho}_t = \bar{y} - b_t (x_t - \bar{x})
\]

which is consistent and which can be termed as an immaculate predictor for the target function $P$.

III. EFFICIENCY OF FEASIBLE PREDICTORS

The exact expressions for the properties of the two feasible predictors considered in Section 2 turn out to be fairly involved and as such it becomes difficult to deduce any meaningful inference from them. We, therefore, employ the large sample asymptotic theory to assess the behaviour of these feasible predictors. Assuming the sample size $n$ to be sufficiently large, we derive the asymptotic approximations for their predictive bias and predictive mean squared error upto the order $O_1$. These approximations are presented in the following theorem.

**Theorem:** The large sample asymptotic approximations for the predictive bias and the predictive mean squared error of the two feasible predictors $\hat{\rho}_c$ and $\hat{\rho}_t$, of the target function $P$, to order $O_1$ are given by
Pred. Bias \( \hat{\beta}_c \) = \( E(\hat{\beta}_c - P) \)
\[ = -\theta (m_k - \bar{m}) \beta \] (3.1)

Pred. Bias \( \hat{\beta}_t \) = \( E(\hat{\beta}_t - P) \)
\[ = 0 \] (3.2)

Pred. MSE \( \hat{\beta}_c \) = \( E(\hat{\beta}_c - P)^2 \)
\[ = [\theta^2 (m_k - \bar{m})^2 + \sigma^2_r] \beta^2 \] (3.3)

Pred. MSE \( \hat{\beta}_t \) = \( E(\hat{\beta}_t - P)^2 \)
\[ = -\sigma^2_r \beta^2 \] (3.4)

where

\[ F_8 = (1 - \delta \theta)^2 + [\lambda^2 \theta r + (1 - \delta \theta - \lambda)^2 d] \left( \frac{1 - \theta}{\theta} \right) \]

with \( \delta = 0 \) or 1 and \( r = \frac{\sigma^2_u}{(1 - \theta) \sigma^2_r \beta^2} \). (3.5)

These results are derived in Appendix.

It is interesting to observe from (3.1) and (3.2) that the immaculate predictor \( \hat{\beta}_t \) is asymptotically unbiased while the least squares predictor \( \hat{\beta}_c \) is not except for the structural variant of the ultra-structural model where \( m_k = \bar{m} \). Both these predictors are equally good as far as their asymptotic bias is concerned for the classical linear regression model for which \( \theta = 0 \).

Examining the expression for the predictive mean squared error of the least squares predictor \( \hat{\beta}_c \), we observe from (3.3) that

Pred. MSE \( \hat{\beta}_c \) = \[ \theta^2 (m_k - \bar{m})^2 + \sigma^2_r (1 - \theta) (1 - \theta - f_c) \] \[ \beta^2 \] (3.6)

where

\[ f_c = \begin{cases} \frac{d(1 - \theta)^2}{\theta} , & \text{if } \lambda = 0 \\ d\theta + r , & \text{if } \lambda = 1 \end{cases} \] (3.7)

which clearly shows that the least squares predictor \( \hat{\beta}_c \) will have smaller variability when \( \lambda = 0 \) as compared to the case when \( \lambda = 1 \), if and only if,
Thus, the least squares predictor \( \hat{\beta}_c \) will give better prediction for the average value of the study variable than that of its actual value when and only when the condition (3.8) holds true. The reverse will happen otherwise, i.e., it will provide better prediction for the actual value when

\[
r > \frac{(1-2\theta) d}{\theta}.
\]  

(3.9)

For simultaneous prediction of the average and actual values both, we observe that the least squares predictor \( \hat{\beta}_c \) will have smallest variability when scaling factor \( \lambda \) in the target function \( P \) is chosen as

\[
\lambda = \frac{(1-\theta) d}{(r\theta + d)} = \lambda_c, \text{ (say)}.
\]  

(3.10)

Thus, if we choose \( \lambda = \lambda_c \) in the target function \( P \), the least squares predictor \( \hat{\beta}_c \) will provide the best predictions for the target function \( P \) in totality.

Exploring further, we observe from (3.10) that \( \lambda_c = 0 \) for \( d = 0 \) and \( \lambda_c = 1 \) for \( \theta = 0 \). This implies that the predictor \( \hat{\beta}_c \) will provide the best predictions for the average value of the study variable when the given measurement error model is in the functional form, i.e., when \( d = 0 \). Similarly, this predictor will provide the best predictions for the actual value of the study variable when the given model is free from measurement errors, i.e., when \( \theta = 0 \).

Let us now examine the predictive mean squared error approximations for the immaculate predictor \( \hat{p}_i \) which can be written as

\[
\text{Pred. MSE (} \hat{p}_i \text{)} = \left[ \sigma^2 + \left( \frac{1-\theta}{\theta} \right) f_i \right] \beta^2.
\]  

(3.11)

where

\[
f_i = \begin{cases} 
    d & \text{if } \lambda = 0 \\
    r\theta & \text{if } \lambda = 1.
\end{cases}
\]  

(3.12)

From this we observe that the immaculate predictor \( p_i \) will provide better predictions for the average value of the study variable than those of its actual value when and only when

\[
r > \frac{d}{\theta}.
\]  

(3.13)
and the reverse happens otherwise, i.e., it provides better predictions for the actual value when

$$r < \frac{d}{\theta}. \hspace{1cm} (3.14)$$

For composite prediction of the average and actual values, we observe from (3.11) that the immaculate predictor \( \hat{\rho}_I \) will have the smallest variability when we choose \( \lambda \) in the target function \( P \) as

$$\lambda = \frac{d}{(r \theta + d)} = \lambda_I, \hspace{1cm} \text{(say)}, \hspace{1cm} (3.15)$$

implying thereby that with this value of \( \lambda = \lambda_I \) in the target function, the immaculate predictor \( \hat{\rho}_I \) will provide the best composite predictions for the average and actual values of the study variable.

It is interesting to observe from (3.10) and (3.15) that

$$\lambda_c < \lambda_I \hspace{1cm} (3.16)$$

which implies that the best predictions of the target function \( P \) with least squares predictor \( \hat{\rho}_c \) give lesser importance to the actual value of the study variable as compared to those with immaculate predictor \( \hat{\rho}_I \).

It is further observed that \( \lambda_I = 0 \) when \( d = 0 \) and \( \lambda_I = 1 \) when \( \theta = 0 \). Thus, we conclude that the immaculate predictor \( \hat{\rho}_I \) will also provide the best predictions for the average value of the study variable in the functional form of the measurement error model as happens in the case of least squares predictor. Similarly, it provides the best predictions for the actual value in the classical linear regression model.

Finally, comparing the performance of the immaculate predictor \( \hat{\rho}_I \) with that of the least squares predictor \( \hat{\rho}_c \), we observe from (3.3) and (3.4) that the immaculate predictor \( \hat{\rho}_I \) will provide better predictions for the target function \( P \) than those obtained by the least squares predictor when

$$\text{Pred. MSE } (\hat{\rho}_c) - \text{Pred. MSE } (\hat{\rho}_I) \hspace{1cm} (3.17)$$

$$= \theta^2 (m_k - m)^2 - \sigma^2 \{ \theta (2 - \theta) - (2\lambda - (2 - \theta)) (1 - \theta) d \} \beta^2$$

is positive which is not possible for any \( \lambda, d \) or \( \theta \) and therefore we conclude that the immaculate predictor will not be able to dominate the least squares predictor for any kind of predictions up to the order of our approximations.
Situation may however, change if we consider higher order approximations. This needs to be explored.

IV. SOME REMARKS

We have studied the problem of simultaneous prediction of average and actual values of the study variable in a linear regression model with measurement errors and have proposed a composite target function which allows weightage to the two kinds of predictions according to practitioner's choice. This provides a general framework from which individual predictions of average and actual values can easily be obtained. The setup of ultra-structural model considered in this paper facilitates the study of these predictions for the measurement error model in general and its two important variants, viz., the functional and the structural form, and also for the classical linear regression model which is free from measurement errors. We have considered two important predictors, viz., the least squares predictor which ignores the measurement errors in variables and the immaculate predictor which is based on corrected variables. It is observed that the least squares predictor provides the best linear and unbiased predictions for the composite target function and also for two individual prediction problems when there is no measurement error in the variables, i.e., the case of classical linear regression model. However, this optimality is lost when measurement errors enter the scene. In that case, the least squares predictor is not only biased but inconsistent also. But if we apply correction for these errors in this predictor, we have an immaculate predictor which is not only consistent but asymptotically unbiased also though it has larger predictive mean squared error than that of the least squares predictor which ignores the measurement errors in variables at least to the order of our approximations. It will be worth to explore higher order approximations for the predictive mean squared errors of the two predictors for judging their efficiency.

It is also interesting to observe that the performance properties of the two predictors remain the same at least asymptotically irrespective of the nature of study, i.e., whether we consider prediction of an outside sample value such as forecasting or a within sample value. It is also probably attributed to the order of approximations we have considered. If higher order terms than $O_p(1)$ are retained, the performance of these predictors might differ in the two different situations. It will be interesting to extend our results in this direction.

It is to be noted that in the present study we have assumed the knowledge of $\sigma^2$, the true variance of measurement errors in the explanatory variable which has been utilized to formulate the immaculate predictor. Quite often we have other kinds of apriori information also. For example, instead of $\sigma^2$, the variance
\( \sigma_u^2 \) of error term \( u \) in the model may be known apriori. Alternatively, we may have information regarding ratios of variances; see, for instance, Stanely (1988). It will be illuminating to explore which kind of apriori information is more beneficial in designing the optimal consistent predictor.

Sometimes, we have repeated observations on the study and explanatory variables both. Can these repeated observations be fruitful in improving the performance of the proposed predictors in any way? This also needs to be examined.

Another direction of further investigation is to explore the role of instrumental variables in construction of efficient predictors. Employment of instrumental variables of which grouping of observations is a special case is quite popular for estimation in measurement error model; see, for example, Judge et al. (1985, Chapter 13). This could play a role here also.

Some work on such issues is underway. Although we have confined our attention to a simple model containing just two variables, it would be interesting to conduct similar studies for other cases too like when we have more than one explanatory variable and only a few of them are subject to measurement errors. The other interesting possibility is when we have multiple study variables like in seemingly unrelated regression equations setup; see, for example, Srivastava and Giles (1987).

- ACKNOWLEDGEMENT

The financial support from the Council of Scientific and Industrial Research (CSIR) of India in the form of a Junior Research Fellowship to the second author is gratefully acknowledged.

APPENDIX

The target function \( P \) can be written as

\[
P = \lambda y_k + (1 - \lambda) E (y_k) \\
= \alpha + \beta m_k + \lambda (\beta w_k + u_k). \tag{A.1}
\]

The feasible predictor \( \hat{\beta} \) for \( P \) is given by

\[
\hat{\beta} = \hat{\alpha} + \beta x_k \tag{A.2}
\]

= \bar{y} + \beta (x_k - \bar{x})
\[\begin{align*}
\bar{y} &= \alpha + \beta m_k + \beta (w_k + v_k) + \frac{1}{\sqrt{n}} (Q_u - \beta Q_v) \\
&+ (\beta - \beta) \left[ (m_k - \bar{m}) + (w_k + v_k) - \frac{1}{\sqrt{n}} (Q_v + Q_w) \right]
\end{align*}\]

where \( \hat{\alpha} \) and \( \hat{\beta} \) are respectively the estimates of regression parameters \( \alpha \) and \( \beta \) and are given by

\[\begin{align*}
\hat{\alpha} &= \overline{y} - \beta \overline{x} \\
\hat{\beta} &= \frac{s_{xy}}{s_{xx} - q \sigma^2}
\end{align*}\]  

(A.3)

with \( q = 0 \) for the ordinary least squares estimates which are contaminated with measurement errors and \( q = 1 \) for the immaculate estimates which are purged of the contamination errors, and

\[\begin{align*}
Q_u &= \sqrt{n} \overline{u} \\
Q_v &= \sqrt{n} \overline{v} \\
Q_w &= \sqrt{n} \overline{w}
\end{align*}\]  

(A.4)

Writing \( u = \text{Col} (U_1, \ldots, u_n), \, v = \text{Col} (V_1, \ldots, v_n), \, w = \text{Col} (W_1, \ldots, w_n), \, A = \left[ I_n - \frac{1}{n} e e' \right] \) with \( e \) as a column vector with all its elements as unity, and

\[\begin{align*}
\sqrt{n} \left( \frac{u'Au}{n} - \sigma_u^2 \right) &= e_u \\
\sqrt{n} \left( \frac{v'Av}{n} - \sigma_v^2 \right) &= e_v \\
\sqrt{n} \left( \frac{w'Aw}{n} - \sigma_w^2 \right) &= e_w
\end{align*}\]  

(A.5)

\[\begin{align*}
\frac{m' A u}{\sqrt{n}} &= Z_m u, \quad \frac{m' A v}{\sqrt{n}} = Z_m v, \quad \frac{m' A w}{\sqrt{n}} = Z_m w
\end{align*}\]
we observe that $e_v$, $e_r$, $e_w$, $Z_{mu}$, $Z_{mv}$, $Z_{uw}$, $Z_{uv}$ and $Z_{vw}$ are all of order $O(1)$. Thus, we can write

$$s_{xy} = \frac{1}{n} x' Ay$$  \hfill (A.6)

$$= \beta s_{xx} - \frac{\beta}{n} \left[ \sqrt{n} (Z_{mu} + Z_{mv} + e_v) - \frac{\sigma_v^2}{\beta} (Z_{mu} + Z_{mv} + Z_{uv}) \right]$$

$$= \beta s_{xx} - \beta \sigma_v^2 \left( 1 + \frac{g}{\sqrt{n}} \right)$$

where $y = \text{Col} (y_1, \ldots, y_n)$, $x = \text{Col} (x_1, \ldots, x_n)$ and

$$g = \frac{1}{\sigma_v^2} \left[ (Z_{mv} + Z_{mv} + e_v) - \frac{1}{\beta} (Z_{mu} + Z_{mv} + Z_{uw}) \right].$$  \hfill (A.7)

Similarly,

$$s_{xx} = \frac{1}{n} x' Ax$$  \hfill (A.8)

$$= (S_{xx} + \sigma_v^2) + \frac{1}{\sqrt{n}} \left[ (e_v + e_w) + 2 (Z_{mv} + Z_{mv} + Z_{vw}) \right]$$

$$= \frac{\sigma_v^2}{\theta} \left( 1 + \frac{h}{\sqrt{n}} \right)$$

where

$$h = \frac{\theta}{\sigma_v^2} \left[ (e_v + e_w) + 2 (Z_{mv} + Z_{mv} + Z_{vw}) \right].$$  \hfill (A.9)

Utilizing these $s_{xy}$ and $s_{xx}$ in (A.3), we obtain the least squares estimate $\hat{\beta}_c$ of $\beta$ as

$$\hat{\beta}_c = \frac{s_{xy}}{s_{xx}}$$  \hfill (A.10)

$$= \frac{\beta s_{xx} - \beta \sigma_v^2 (1 + g/\sqrt{n})}{s_{xx}}$$
or

\[
(\hat{\beta}_c - \beta) = -\beta \theta \left(1 + \frac{g}{\sqrt{n}}\right) \left(1 + \frac{h}{\sqrt{n}}\right)^{-1} \\
= -\beta \theta - \frac{(g - h) \theta \beta}{\sqrt{n}} \frac{1}{n} \frac{1}{n} + \ldots
\]

and hence the estimation error of \( \hat{\beta}_c \) upto order \( 0, (1) \) is given by

\[
(\hat{\beta}_c - \beta) = -\beta \theta.
\] (A.11)

Similarly, we obtain the immaculate estimate \( \hat{\beta}_r \) of \( \beta \) as

\[
\hat{\beta}_r = \frac{s_{xy}}{s_{xx} - \sigma^2_y} = \frac{\beta s_{xx} - \beta \sigma^2_y (1 + g/\sqrt{n})}{s_{xx} - \sigma^2_y}.
\] (A.12)

so that the estimation error associated with \( \hat{\beta}_r \) is given by

\[
(\hat{\beta}_r - \beta) = -\beta \frac{g}{\sqrt{n}} \left(\frac{\theta}{1 - \theta} \right) \left(1 + \frac{h}{(1 - \theta)\sqrt{n}}\right)^{-1} \\
= 0 \quad \text{upto order } 0, (1).
\] (A.13)

Employing (A.11) in (A.2), we obtain the prediction error of the least squares predictor \( \hat{\beta}_c \), upto the order \( 0, (1) \) as

\[
(\hat{\beta}_c - P) = \beta [v_k + (1 - \lambda) w_k] - \lambda u_k - \beta \theta [(m_k - \overline{m}) + (w_k + v_k)].
\] (A.14)

Similarly, employing (A.13) in (A.2), we obtain the prediction error of the immaculate predictor \( \hat{\beta}_r \), upto the order \( 0, (1) \) as

\[
(\hat{\beta}_r - P) = \beta [v_k + (1 - \lambda) w_k] - \lambda u_k
\] (A.15)

These prediction errors are utilized to evaluate \( E(\hat{\beta} - P) \) and \( E(\hat{\beta} - P)^2 \) which provide the predictive bias and predictive mean squared error approximations for the two predictors as given in the Theorem.
REFERENCES


