Classical Measurement Error with Several Regressors*

Andrew B. Abel
Wharton School of the University of Pennsylvania
National Bureau of Economic Research

February 4, 2018

Abstract

In OLS regressions with several regressors, measurement error affects estimated coefficients through two distinct channels: a multivariate attenuation factor that applies to all coefficients and generalizes the standard attenuation factor in univariate regressions; weight shifting associated with measurement error on each mismeasured regressor that further attenuates the coefficient on that regressor and affects coefficients on other regressors. I introduce a scalar "measurement error multiplier" that indicates the contribution of measurement error to the generalized variance of the measured regressors. It multiplies the variances of the measurement errors where it appears in both the multivariate attenuation channel and the weight shifting channel.

*I thank Toni Whited for correspondence that stimulated this research and Michael Roberts for scores of helpful discussions. I also thank Frank Diebold, Joao Gomes, Richard Kihlstrom and Luke Taylor for helpful discussion and the Penn Macro Lunch for helpful comments.

"If more than one variable is measured with error, there is very little that can be said."–Greene (2003, p. 86)[1]

When the lone regressor in a univariate regression has classical measurement error, the OLS estimate of its coefficient will be biased toward zero. Although this attenuation of the OLS coefficient in a univariate regression is universally known and trivial to derive, there are no simple or widely-known generalizations to the case of OLS regressions with an arbitrary number of regressors, some or all of which suffer from classical measurement error.[2] In this paper, I provide such a generalization. In particular, I derive a simple expression for the *plims* of OLS estimates of the coefficients of an arbitrary number of regressors in the presence of classical measurement error. To make the interpretation as intuitive as I can, I introduce some new terminology. I define a *measurement error multiplier*, which is a scalar that arises in several places in the expressions for the OLS coefficients. Wherever this scalar appears, it multiplies the variance of measurement error in a particular regressor or multiplies the product of the variances of measurement errors in all the regressors. Not only is this multiplier useful in simplifying expressions, it has a natural interpretation. It is an increasing function of the share of the generalized variance of the vectors of measured regressors that is accounted for by the generalized variance of the vector of measurement errors. The larger is the share accounted for by the generalized variance of measurement errors, the larger is the measurement error multiplier.

Measurement error affects the estimated OLS coefficients through two distinct channels. One of these channels operates through what I call the *multivariate attenuation factor*. This factor is the multivariate generalization of the standard attenuation factor. It attenuates the OLS estimates on all coefficients by the same proportion. The degree of multivariate

[1] Greene includes a footnote at the end of this sentence that cites several references that provide analytic results for cases with more than one variable measured with error. I discuss some of these references in the literature review in Section 1.

[2] This assessment, which is reflected in the quotation above from Greene (2003), is shared by Wooldridge (2002), another leading econometrics textbook, which states "If measurement error is present in more than one explanatory variable, deriving the inconsistency in the OLS estimators under extensions of the CEV [classical errors-in-variables] assumptions is complicated and does not lead to very usable results." (p. 76). It also shared by Roberts and Whited (2013), which states "What if more than one variable is measured with error under the classic errors-in-variables assumption? Clearly, OLS will produce inconsistent estimates of all the parameter estimates. Unfortunately, little research on the direction and magnitude of these inconsistencies exists because biases in this case are typically unclear and complicated to derive..." (p.498)

attenuation is an increasing function of the product of the measurement error multiplier and the product of the variances of the measurement errors of all regressors. If one or more regressors are measured without error, the multivariate attenuation factor is simply equal to one–that is, there is no attenuation through this channel.

The other channel is what I call *weight shifting*. When a particular variable is measured with error, some other variable(s) may help account for the variation in the dependent variable that would have been due to the particular measured variable if it were measured without error. This weight shifting further attenuates the OLS coefficient on the particular mismeasured regressor and can affect the OLS coefficients on other variables. Of course, when account is taken of measurement error affecting all regressors, the further attenuation of the coefficient of the particular mismeasured regressor could be exacerbated or reversed by the weight shifting from other mismeasured regressors. The degree of weight shifting associated with measurement error in any particular regressor is an increasing function of the product of the measurement error multiplier and the variance of the measurement error in that regressor.

Taking account of the two channels described above, the *plim* of the vector of coefficients estimated by OLS is the product of (1) the scalar multivariate attenuation factor, which is less than or equal to one; (2) a matrix capturing weight shifting, which is itself a weighted average of the identity matrix and another matrix reflecting interactions among the measured regressors; and (3) the true value of the coefficient vector. After deriving this simple expression for the *plim* of the estimated coefficient vector, I apply this expression to cases with one or more irrelevant regressors. An irrelevant regressor is a regressor for which the true coefficient in the underlying linear model is zero. It is well known that an irrelevant variable can have a nonzero estimated coefficient if it is correlated with a relevant variable that is measured with error. In this case a nonzero coefficient will cause false rejection of the null hypothesis that the irrelevant variable has a zero coefficient. I demonstrate a more troubling finding that an irrelevant variable can have a nonzero estimated coefficient even if it is uncorrelated with any relevant variable, if the irrelevant variable is correlated with another irrelevant variable that happens to be correlated with a relevant variable measured with error.

After a brief review of the literature in Section 1, I describe the basic linear model, including regressors with classical measurement error, in Section 2. I present the simple analytics and interpretation of the impact of measurement error in Section 3, where I devote subsection 3.1 to the definition and interpretation of the measurement error multiplier, subsection 3.2 to a derivation of the bias induced by measurement error, and subsection 3.3 to a simple expression for the *plim* of the estimated coefficient vector as the product of a scalar multivariate attenuation factor, a matrix representing weight shifting, and the vector of true coefficients in the underlying linear model. In Section 4, I show that measurement error can lead to false rejection of the null hypothesis that an irrelevant variable has a zero coefficient. Section 5 concludes and an Appendix contains various proofs.

1 Literature Review

Theil (1965) provides an early analysis of the impact of classical measurement error on OLS estimates in regressions with two regressors, both of which are measured with error. He derives approximate[3] expressions for the biases in the estimated coefficients. Interpreting his findings in the nomenclature of this paper, Theil shows that the multivariate attenuation factor is an increasing function of the correlation between the explanatory variables. Theil also shows the "the influence of [one the regressors] on y is partly allocated to [the other regressor], (p. 215), which is essentially the weight shifting channel that I described in the introduction and analyze in Section 3.

Levi (1973) examines the case with an arbitrary $k \geq 2$ regressors in which one variable is observed with measurement error and the other $k-1$ regressors are observed without error. He shows that the OLS estimate of the coefficient on the variable measured with error is attenuated toward zero. As I will show, the presence of one or more variables measured without error in this framework implies that the multivariate attenuation factor equals one, that is, there is no multivariate attenuation. The attenuation noted by Levi is indeed operative, but it operates through the weight shifting channel I introduce.

Griliches (1986) examines OLS regressions with two regressors and presents expressions for the *plims* of the biases in these coefficients when one or both regressors are measured with error. If only one of the regressors is measured with error, he shows that "the bias in the coefficient of the erroneous variable is 'transmitted'" to the other coefficient (p. 1479).

[3]Thiel (1965, p. 328) points out that the expressions hold "only if the error moments are sufficiently small compared to the moments of the true explanatory variable."

As in Levi (1973), in the presence of regressors measured without error, the multivariate attenuation factor is one. The attenuation of the coefficient on the mismeasured regressor and the "transmission" of bias to the other coefficient(s) are captured by the weight shifting I analyze. Griliches also provides expressions for the plims of the OLS coefficients in regressions with two regressors, both measured with error. The biases in the two coefficients contain a common factor that is related to the multivariate attenuation factor I introduce.

Garber and Klepper (1980) is the closest antecedent to the analysis in this paper. They examine OLS coefficients in regressions with an arbitrary $k > 2$ regressors, two of which are measured with error. They describe two channels of bias in the estimated coefficients on regressors measured with error: an "own" effect that captures attenuation due to spurious variation in the regressor arising from measurment error; and a "smearing" effect in which the bias in the other mismeasured regressor contaminates the coefficient of a given mismeasured regressor. The coefficients of the regressors measured without error are also contaminated by smearing from the coefficients of the mismeasured errors, which Garber and Klepper describe as a "picking up" effect. The effects that Garber and Klepper label "own" and "smearing" have a superficial similarity to the two channels I identify as multivariate attenuation and weight shifting, but there is an important difference. As I will show, if the regression includes at least one variable measured without error, as in Garber and Klepper, then the multivariate attenuation factor equals one so there is so multivariate attenuation bias. All of the bias is captured by weight shifting. The matrix that represents weight shifting in Proposition 1 is a weighted average of own-effect attenuation arising from spurious variation due to measurement error and smearing from other variables that are measured with error. Only if all of the regressors are measured with error will the scalar multivariate attenuation factor in Proposition 1 be less than 1.

My interest in the impact of measurement error was kindled by analyzing a time-honored question in corporate finance, namely, the extent to which cash flow affects a firm's capital investment, after taking account of the firm's q ratio, which is the ratio of the market value of the firm to the replacement cost of its capital The widely used q theory of investment implies that cash flow should not have any additional impact on investment after taking account of q. Nevertheless, empirical studies repeatedly find that in OLS regressions of investment on q and cash flow, both q and cash flow have significant positive coefficients. In Abel (2017) I interpret the positive estimated coefficient on cash flow as rising from measurement error in q combined with a positive correlation between q and cash flow. I derived closed-form expressions for the regression coefficients and interpreted them in terms of what I called *bivariate attenuation bias* (bivariate, rather than multivariate, since there were $k = 2$ regressors) and weight shifting. In the current paper, I extend the analysis to an arbtrary number of regressors, and that extension is far from trivial. I explore more fully the role and interpretation of the measurement error multiplier. In addition, the weight-shifting matrix, which is the weighted average of the identity matrix and the matrix of attenuation-adjusted coefficients from auxiliary regressions, is new.

2 The Linear Model

Consider a linear model relating a dependent variable y to k explanatory variables, $x_1, ..., x_k$,

$$Y = X\beta + U, \tag{1}$$

where $Y = [y_1, ..., y_N]'$ is the $N \times 1$ vector of observations on the dependent variable, X is the $N \times k$ matrix with nth row $[x_{1n}, ..., x_{kn}]$, which is the vector of nth observations on the regressors, β is the $k \times 1$ vector $[\beta_1, ..., \beta_k]'$ of coefficients on the regressors, and U is the $N \times 1$ vector $[u_1, ..., u_N]'$ where $E\{u_i\} = 0$. I assume that the regressors $x_1, ..., x_k$ are de-meaned, $X'X$ has rank k, and X and U are independent.

The regressors are measured with classical measurement error, so the econometrician sees

$$\widetilde{x}_{in} = x_{in} + \varepsilon_{in}, \tag{2}$$

where ε_{in} is the measurement error in the nth observation of x_i, $i = 1, ..., k$. The measurement errors are classical measurement errors: they have mean zero, are independent of all u_n and all x_{jn}, and are independent of each other. The variance of the measurement error ε_{in} is constant across n, so that $Var(\varepsilon_{in}) = \sigma_i^2 \geq 0$, for $n = 1, ..., N$ and $i = 1, ..., k$. The variance-covariance matrix of the vector of measurement errors $[\varepsilon_1, ..., \varepsilon_k]$ is the $k \times k$ diagonal matrix Σ, with jth diagonal element equal to $\sigma_j^2 \geq 0$.

The econometrician's observations on the regressors are represented by the $N \times k$ matrix $\widetilde{X}$ with nth row equal to $[\widetilde{x}_{1n}, ..., \widetilde{x}_{kn}]$ and the measurement errors are represented in matrix form as $\mathcal{E}$ with nth row equal to $[\varepsilon_{1n}, ..., \varepsilon_{kn}]$, so

$$\widetilde{X} = X + \mathcal{E}. \tag{3}$$

Let $\widehat{\beta}$ be the estimate of β in an OLS regression of Y on $\widetilde{X}$, so

$$\widehat{\beta} = \left(\widetilde{X}'\widetilde{X}\right)^{-1} \widetilde{X}'Y. \tag{4}$$

Define the $k \times 1$ vector $b \equiv plim\ \widehat{\beta}$, and take *plims* of both sides of equation (4) to obtain

$$b = (V + \Sigma)^{-1} V\beta, \tag{5}$$

where $V = plim\left(\frac{1}{N}X'X\right)$ is the $k \times k$ variance-covariance matrix with (i, j) element equal to $Cov\,(x_i, x_j)$, and $plim\left(\frac{1}{N}\mathcal{E}'\mathcal{E}\right) = \Sigma$. An alternative expression for b is obtained by substituting $(V + \Sigma - \Sigma)\,\beta$ for $V\beta$ in equation (5) to obtain

$$b = \beta - (V + \Sigma)^{-1} \Sigma\beta. \tag{6}$$

Equations (5) and (6) immediately deliver well-known expressions for the attenuation factor and attenuation bias in univariate regressions with classical measurement error. If $k = 1$, then V is the scalar $var\,(x_1)$, which is the variance of the true value of the regressor, and Σ is σ_1^2, which is the variance of the measurment error in that regressor. Thus, the scalar $(V + \Sigma)^{-1} V$ in equation (5) is the well-known attenuation factor $\frac{var(x_1)}{var(x_1)+\sigma_1^2} \leq 1$ and the scalar $-(V + \Sigma)^{-1} \Sigma\beta$ in equation (6) is the well-known attenuation bias $\frac{-\sigma_1^2}{var(x_1)+\sigma_1^2}\beta_1$.

The goal of this paper is to provide an intuitive interpretation of equations (5) and (6) in situations with multiple regressors. This interpretation is facilitated by introducing the concepts of *multivariate attenuation* and *weight shifting*. In developing these concepts, a measurement error multiplier arises and I will define and interpret that multiplier.

3 The Impact of Measurement Error

Consider an OLS regression with multiple regressors, some or all of which may be observed with classical measurement error. I begin by describing notational conventions:

(1) for any $k \times k$ matrix M, the $k \times k$ diagonal matrix M_D is formed from M by setting all off-diagonal elements of M equal to zero;

(2) for any $k \times k$ matrix M, the $(k-1) \times (k-1)$ matrix M_{-j} is formed from M by eliminating its jth row and jth column;

(3) for any $k \times 1$ vector z, the $(k-1) \times 1$ vector z_{-j} is formed from z by eliminating its jth element.

Under notational convention (1), V_D is the $k \times k$ diagonal matrix with jth diagonal element equal to $var(x_j)$, which is the variance of the true value of the jth regressor. Define the $k \times k$ diagonal matrix

$$S \equiv V_D^{-1}\Sigma. \tag{7}$$

The jth diagonal element of S is $s_j^2 \equiv \frac{\sigma_j^2}{var(x_j)}$, which is the variance of the measurement error in the jth regressor relative to the variance of the true value of that regressor.

3.1 Measurement Error Multiplier

To develop an expression for the bias in the coefficient vector b, define the following *measurement error multiplier.*

Definition 1 *Define the measurement error multiplier* $\omega \equiv \left[\frac{\det(V+\Sigma)-\det\Sigma}{\det V_D}\right]^{-1}$.

The rationale for this terminology is that, as I will show later, the impact of measurement error is mediated through ωs_j^2, $j = 1, ..., k$, in the weight-shifting channel and $\omega \det S$ in the multivariate attenuation channel; in both channels, the measurement error multiplier ω multiplies the variances of measurement errors.

The determinant of a variance-covariance matrix of a vector is the generalized variance of that vector.[4] Therefore, Definition 1 implies that ω is inversely related to the amount by

[4] Wilks, Samuel S., Mathematical Statistics, New York: John Wiley and Sons, Inc., 1962, p. 547: "The determinant of the variance-covariance matrix is sometimes called the *generalized variance* of" the distribution. The generalized variance of a vector can be interpreted as the amount of independent variation in that vector. For example, the generalized variance of the random vector (z_1, z_2) is $Var(z_1)\,Var(z_2) - [Cov(z_1, z_2)]^2 = \left(1 - R_{12}^2\right) Var(z_1)\,Var(z_2)$, where R_{12}^2 is the squared correlation of z_1 and z_2. As R_{12}^2 increases for given $Var(z_1)$ and $Var(z_2)$, the linear dependence between z_1 and z_2 increases, thereby reducing the independent variation in (z_1, z_2) and reducing the generalized variance of (z_1, z_2). In general, let Ω be the (non-singular) variance-covariance matrix of $z_1, ..., z_n$. If $z_1, ..., z_n$ are mutually independent, the

which the generalized variance of the measured regressors, $\det(V+\Sigma)$, exceeds the generalized variance of the measurement errors, $\det\Sigma$, normalized by the product of the variances of the true values of the regressors, $\det V_D$. Thus, when the generalized variance of the measured regressors exceeds the generalized variance of the measurement error by only a small amount, so that measurement error contributes a large part of the generalized variance of the measured regressors, the measurement error multiplier, ω, is large.

Before presenting properties of ω for general $k \geq 1$ in Lemma 1, I will illustrate ω in the simple case in which $k=2$. With $k=2$, $V = \begin{bmatrix} var(x_1) & cov(x_1,x_2) \\ cov(x_1,x_2) & var(x_2) \end{bmatrix}$, $\Sigma = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$, $V_D = \begin{bmatrix} var(x_1) & 0 \\ 0 & var(x_2) \end{bmatrix}$, and $S = \begin{bmatrix} s_1^2 & 0 \\ 0 & s_2^2 \end{bmatrix}$, so straightforward calculation yields

$$\omega = \frac{1}{1+s_1^2+s_2^2-R_{12}^2}, \qquad \text{if } k=2, \tag{8}$$

where $R_{12}^2 \equiv \frac{[cov(x_1,x_2)]^2}{var(x_1)var(x_2)}$ is the squared correlation of x_1 and x_2. An increase in R_{12}^2, which increases the correlation between the true regressors, reduces the amount of independent variation in the true regressors relative to the independent variation in measurement error, for given $var(x_1)$ and $var(x_2)$. Thus, an increase in R_{12}^2 implies that measurement error provides more of the independent variation in the measured regressors, and this increased relative contribution of measurement error to independent variation is reflected in an increased value of the measurement error multiplier ω in equation (8).

Lemma 1 *Recall* $\omega \equiv \frac{\det V_D}{\det(V+\Sigma)-\det\Sigma}$ *and* $S \equiv V_D^{-1}\Sigma$. *Define*[5] $\omega_j \equiv \frac{\det V_{D,-j}}{\det(V_{-j}+\Sigma_{-j})-\det\Sigma_{-j}}$. *Then*

1. $\omega > 0$
2. $\omega_j > 0$, *for* $j = 1, ..., k$ *and* $k \geq 2$

generalized variance is simply $\prod_{i=1}^{n} Var(z_i)$. Since Ω is positive definite, Hadamard's inequality implies that $\det\Omega \leq \prod_{i=1}^{n} Var(z_i)$. Correlation among the random variables $z_1, ..., z_n$ reduces the independent variation in these variables as indicated by a generalized variance smaller than $\prod_{i=1}^{n} Var(z_i)$.

[5] By convention, $\omega_j = 1$ when $k = 1$.

3. if $k=1$*, then* $\omega=1$

4. if $k=2$*, then* $\omega_1=\omega_2=1$

5. $\sigma_j^2\frac{\det(V_{-j}+\Sigma_{-j})}{\det(V+\Sigma)}=\frac{\omega}{\omega_j}s_j^2\frac{1+\omega_j\det S_{-j}}{1+\omega\det S}=\frac{\frac{\omega}{\omega_j}s_j^2+\omega\det S}{1+\omega\det S}$ *for* $j=1,...,k$.

6. $\frac{\omega}{\omega_j}s_j^2<1$ *for* $j=1,...,k$.

The properties in Lemma 1 will be useful in analyzing the impact of measurement error on OLS estimates.

3.2 Bias in b

To derive an expression for $(V+\Sigma)^{-1}\Sigma\beta$, which is the (negative of the) bias in b in equation (6), first use the standard formula for the inverse of a matrix[6] to obtain

$$(V+\Sigma)^{-1}=\frac{1}{\det(V+\Sigma)}C, \tag{9}$$

where C is the $k\times k$ matrix of co-factors of $V+\Sigma$: the (i,j) element of C is the (i,j) co-factor $C_{ij}=(-1)^{i+j}M_{ij}$ where M_{ij} is the (i,j) minor of $V+\Sigma$. Therefore,

$$C_{jj}=\det(V_{-j}+\Sigma_{-j})>0, \qquad j=1,...,k. \tag{10}$$

Equation (10) implies that the diagonal matrix C_D, which has (j,j) element equal to $C_{jj}>0$, is invertible so

$$(V+\Sigma)^{-1}\Sigma\beta=(V+\Sigma)^{-1}C_D^{-1}C_D\Sigma\beta. \tag{11}$$

Substitute the right hand side of equation (9) for $(V+\Sigma)^{-1}$ on the right hand side of equation (11) to obtain

$$(V+\Sigma)^{-1}\Sigma\beta=\frac{1}{\det(V+\Sigma)}CC_D^{-1}C_D\Sigma\beta. \tag{12}$$

The expression on the right hand side of equation (12) explicitly contains the co-factors of $V+\Sigma$. To obtain an expression that does not explicitly contain co-factors, I calculate the

[6] In general, the inverse of a matrix is the product of the reciprocal of the determinant of the matrix and the transpose of the matrix of co-factors. For a variance-covariance matrix, which is symmetric, the matrix of co-factors is symmetric, so I have dispensed with transposing it in equation (9).

matrix products CC_D^{-1} and $C_D\Sigma$ that appear in equation (12). To calculate CC_D^{-1}, recall that $\widetilde{x}_{-j}$ is the $(k-1)\times 1$ vector consisting of the measured regressors $\widetilde{x}_1, ..., \widetilde{x}_k$, excluding $\widetilde{x}_j$, and use the following definition.

Definition 2 *Define Γ as the $k \times k$ matrix with diagonal elements equal to zero and off-diagonal elements γ_{ij}, $i \neq j$, equal to the coefficient on $\widetilde{x}_i$ in an auxiliary OLS regression of $\widetilde{x}_j$ on $\widetilde{x}_{-j}$.* [7]

Thus, the jth column of Γ contains the coefficients of the auxiliary regression of $\widetilde{x}_j$ on the other $k-1$ regressors. Lemma 2 uses the definition of Γ to present a simple expression for CC_D^{-1}.

Lemma 2 $CC_D^{-1} = I_k - \Gamma$.

Now consider the product $\frac{1}{\det(V+\Sigma)} C_D\Sigma$, which appears in equation (12). In this product, the scalar $\frac{1}{\det(V+\Sigma)}$ is positive but can be greater than, less than, or equal to one, and the diagonal elements of the diagonal matrix C_D are co-factors of $V+\Sigma$. To obtain an alternative expression in which (1) the positive scalar is less than or equal to one, and (2) co-factors do not explicitly appear, I define the diagonal matrices W and A for $k \geq 2$.[8]

Definition 3 *For $k \geq 2$, define the $k \times k$ diagonal matrix W as diag* $\left(\frac{\omega}{\omega_1} s_1^2, ..., \frac{\omega}{\omega_k} s_k^2\right)$.

I will refer to W as a *weighting matrix* and will explain the rationale for this terminology in the discussion following Proposition 1. Note that the measurement error multiplier multiplies the variances of the measurement errors, s_j^2, in individual regressors $\widetilde{x}_j$, $j = 1, ..., k$.

Definition 4 *For $k \geq 2$, define the $k\times k$ diagonal matrix A as diag* $(1+\omega_1 \det S_{-1}, ..., 1+\omega_k \det S_{-k})$.

As I will explain in the discussion following Proposition 1, the diagonal elements of A are the reciprocals of the multivariate attenuation factors in each of the k auxiliary regressions of $\widetilde{x}_j$ on $\widetilde{x}_{-j}$, $j = 1, ..., k$. The measurement error multiplier, ω_j, multiplies $\det S_{-j}$, which

[7] Note that if $k = 1$, then Γ is a scalar and equals 0.

[8] For $k = 1$, ω_1 and $\det S_{-1}$ are undefined, so the definitions of W and A in Definitions 3 and 4, respectively, do not apply. With $k = 1$, W and A are scalars, and I adopt the conventions $W = 0$ and $A = 1$, when $k = 1$.

is the product of the variances of the measurement errors, s_i^2, $i \neq j$, in each of the $k-1$ regressors in these regressions.

The following lemma presents some properties of the diagonal matrices A and W and provides a simple expression for $\frac{1}{\det(V+\Sigma)} C_D \Sigma$ that does not directly involve co-factors.

Lemma 3 *For matrices M_1 and M_2, use the notation $M_1 < M_2$ to denote that $M_2 - M_1$ is positive definite and $M_1 \leq M_2$ to denote that $M_2 - M_1$ is positive semi-definite. Then*

1. $0 \leq W < I \leq A$.

2. $(A - I) W = (\omega \det S) I$.

3. $\frac{1}{\det(V+\Sigma)} C_D \Sigma = \frac{1}{1+\omega \det S} AW < I$.

Statement 1 of Lemma 3 implies that the diagonal elements of the weighting matrix W are non-negative and strictly less than 1. It also implies that all of the diagonal elements of A, which are the reciprocals of the multivariate attenuation factors in the k auxiliary regressions, are greater than or equal to 1. Therefore, all of the multivariate attenuation factors in the auxiliary regressions are less than or equal to one. Statement 2 of Lemma 3 is a relationship between A and W that is helpful in proving Proposition 1 below. Statement 3 of Lemma 3 provides an expression for the diagonal matrix $\frac{1}{\det(V+\Sigma)} C_D \Sigma$ that does not explicitly include co-factors, and all of the elements of this diagonal matrix are strictly less than 1.

Use Lemmas 2 and 3 to rewrite the expression for the (negative of the) bias in b in equation (12) as

$$(V + \Sigma)^{-1} \Sigma \beta = \frac{1}{1 + \omega \det S} (I_k - \Gamma) AW\beta. \tag{13}$$

3.3 Estimated Coefficients

To obtain an expression for the *plim* of OLS coefficients, substitute the expression for the (negative of the) bias from equation (13) into equation (6) to obtain

Proposition 1 *The plim of the OLS estimate of β is*

$$b = \frac{1}{1 + \omega \det S} [(I_k - W) + \Gamma AW] \beta.$$

To apply Proposition 1 to a univariate regression, use the conventions in footnote 8 that $W = 0$ and $A = 1$ when $k = 1$, along with Statement 3 of Lemma 1 that $\omega = 1$ when $k = 1$ and $\det S = s_1^2$ when $k = 1$ to obtain $b = \frac{1}{1+s_1^2}\beta_1$, which illustrates the standard attenuation in univariate regressions with classical measurement error. For $k \geq 2$, Proposition 1 shows that the coefficient vector b can be expressed as the product of a positive scalar less than or equal to one, a matrix, and the vector of true coefficients, β. I will call the scalar, which is $\frac{1}{1+\omega \det S}$, the *multivariate attenuation factor* because it multiplies all of the estimated coefficients by a factor that is less than one, provided that $\det S > 0$, that is, provided that all of the regressors are measured with error. If one or more of the regressors is measured without error, then $\det S = 0$, and the multivariate attenuation factor equals 1.

The $k \times k$ matrix $(I_k - W) + \Gamma AW$, which appears in Proposition 1, is a weighted average, with weights $I_k - W$ and W, respectively, of the identity matrix I_k (which is suppressed in this notation) and the matrix product ΓA. The weights $I_k - W$ and W sum to the identity matrix; Statement 1 of Lemma 3 implies that the weight W is positive semi-definite and the weight $I_k - W$ is positive definite. In the matrix product ΓA, which is weighted by W, the matrix Γ is the matrix of regression coefficients on $\widetilde{x}_i$, $i \neq j$, in an auxiliary regression of $\widetilde{x}_j$ on $\widetilde{x}_{-j}$, $j = 1, ..., k$, and the diagonal matrix A, adjusts these coefficients for multivariate attenuation in each of these k auxiliary regressions. If all regressors are measured without error, then $W = 0$, so that $(I_k - W) + \Gamma AW = I_k$; therefore, since the multivariate attenuation factor equals 1 in this case, $b = \beta$.

The following corollary is simply an alternative statement of Proposition 1.

Corollary 1

$$\begin{bmatrix} b_1 \\ ... \\ b_k \end{bmatrix} = \frac{1}{1+\omega \det S}\left(\begin{bmatrix} \left(1 - \frac{\omega}{\omega_1}s_1^2\right)\beta_1 \\ ... \\ \left(1 - \frac{\omega}{\omega_k}s_k^2\right)\beta_k \end{bmatrix} + \Gamma \begin{bmatrix} (1+\omega_1 \det S_{-1})\frac{\omega}{\omega_1}s_1^2\beta_1 \\ ... \\ (1+\omega_k \det S_{-k})\frac{\omega}{\omega_k}s_k^2\beta_k \end{bmatrix}\right)$$

Corollary 1 directly implies the following expression for ith element of b in terms of scalars rather than matrices,

$$b_i = \frac{1}{1+\omega \det S}\left[\left(1 - \frac{\omega}{\omega_i}s_i^2\right)\beta_i + \sum_{j \neq i}\gamma_{ij}\left(1+\omega_j \det S_{-j}\right)\frac{\omega}{\omega_j}s_j^2\beta_j\right]. \qquad (14)$$

4 Irrelevant Regressors

Define a regressor x_i as *relevant* if its true coefficient, β_i, is non-zero, and define a regressor x_j as *irrelevant* if its true coefficient, β_j, equals zero. In this section, I show that measurement error in one or more relevant variables can lead to a non-zero estimated coefficient on an irrelevant variable, thereby leading to false rejection of the null hypothesis that the coefficient on the irrelevant variable is zero. To simplify the exposition, I will focus on the case with only one relevant variable. After examining the coefficient b for an arbitrary number of irrelevant regressors, I examine the case with one relevant regressor and one irrelevant regressor. I show that if the relevant regressor is measured with error, then the estimated coefficient on the irrelevant regressor will be nonzero if the two regressors have nonzero correlation with each other. Then, I extend the analysis to $k = 3$ regressors and show that an irrelevant regressor can have a nonzero estimated coefficient, *even if it is uncorrelated with any relevant regressors*, if it is correlated with an irrelevant regressor that is correlated with a relevant regressor measured with error.

Consider the case with $k \geq 2$ regressors and assume that only one regressor is relevant. Let x_1 be the lone relevant regressor, so $\beta_1 \neq 0$ and $\beta_j = 0$, $j = 2, ..., k$. Corollary 1 implies that if $\beta_j = 0$, for $j = 2, ..., k$, then

$$b = \frac{1}{1 + \omega \det S} \begin{bmatrix} 1 - s_1^2 \frac{\omega}{\omega_1} \\ s_1^2 \frac{\omega}{\omega_1} \left(1 + \omega_1 \det S_{-1}\right) \gamma_{21} \\ \ldots \\ s_1^2 \frac{\omega}{\omega_1} \left(1 + \omega_1 \det S_{-1}\right) \gamma_{k1} \end{bmatrix} \beta_1, \quad \text{if } \beta_2 = \ldots = \beta_k = 0. \tag{15}$$

If $\widetilde{x}_1$ is measured without error, so that $s_1^2 = 0$, then $\det S = 0$ and equation (15) implies that $b' = [\beta_1, 0, ..., 0]'$ so that the OLS estimate of the vector β is consistent.

If $\widetilde{x}_1$ is measured with error, so that $s_1^2 > 0$, then the first element in the vector on the right hand side of equation (15) is positive and less than one, which attenuates the estimated coefficient $\widehat{\beta}_1$. If all of the regressors are measured with error, so that $s_j^2 > 0$, $j = 1, ..., k$, then $\det S > 0$ and the multivariate attenuation factor $\frac{1}{1+\omega \det S}$ is less than one, which further attenuates $\widehat{\beta}_1$. As for the coefficients on the other regressors, say $\widetilde{x}_j$, $j \in \{2, ..., k\}$, the estimated coefficient $\widehat{\beta}_j$ will be nonzero if and only if $\gamma_{j1} \neq 0$. The nonzero values of $\widehat{\beta}_j$, $j = 2, ..., k$, and the attenuation of $\widehat{\beta}_1$ by the factor $1 - s_1^2 \frac{\omega}{\omega_1}$ reflect weight shifting. Using the terminology of Garber and Klepper (1980), the attenuation of $\widehat{\beta}_1$ by the factor $1 - s_1^2 \frac{\omega}{\omega_1}$ reflects the "own" effect of spurious variation due to measurement error and the nonzero values of $\widehat{\beta}_j$, $j = 2, ..., k$, reflect "smearing" for the estimated coefficients of mismeasured regressors and reflect "picking up" for the estimated coefficients of correctly-measured regressors. If all of the regressors are measured with error, then $\det S > 0$ and all of the estimated coefficients are further attenuated by the multivariate attenuation factor $\frac{1}{1+\omega \det S} < 1$. Since the true value of β_j, $j = 2, ..., k$, is zero, the estimated coefficient $\widehat{\beta}_j$ will be inconsistent, and the null hypothesis $H_0 : \beta_j = 0$ will be falsely rejected, if $\gamma_{j1} \neq 0$, provided that $s_1^2 > 0$, regardless of whether any or all of the other s_i^2 are zero.

4.1 One Relevant Regressor and One Irrelevant Regressor

In this subsection, I examine the simple case with one relevant regressor and one irrelevant regressor to illustrate how weight shifting leads to a nonzero estimated coefficient on an irrelevant variable.[9] Specifically, consider the case with $k = 2$, $\beta_1 \neq 0$, and $\beta_2 = 0$. With $k = 2$, $\det S = s_1^2 s_2^2$, $\det S_{-1} = s_2^2$, and, from Statement 4 of Lemma 1, $\omega_1 = 1$. Therefore, equation (15) implies

$$b = \frac{1}{1 + \omega s_1^2 s_2^2} \begin{bmatrix} \left(1 - \omega s_1^2\right) \beta_1 \\ \omega s_1^2 \left(1 + s_2^2\right) \gamma_{21} \beta_1 \end{bmatrix}, \quad \text{if } \beta_2 = 0. \tag{16}$$

The regression coefficient γ_{21} in the second element of the vector on the right hand side of equation (16) is the coefficient on $\widetilde{x}_2$ in an auxiliary (univariate) regression of $\widetilde{x}_1$ on $\widetilde{x}_2$, that is, $\gamma_{21} = \frac{cov(\widetilde{x}_1, \widetilde{x}_2)}{var(\widetilde{x}_2)}$. Since $cov\left(\widetilde{x}_1, \widetilde{x}_2\right) = cov\left(x_1, x_2\right)$ and $var\left(\widetilde{x}_2\right) = \left(1 + s_2^2\right) var\left(x_2\right)$, it follows that $\left(1 + s_2^2\right) \gamma_{21} = \frac{cov(x_1, x_2)}{var(x_2)}$, which is the coefficient on true x_2 in a univariate OLS regression of true x_1 on true x_2. Thus, $\left(1 + s_2^2\right) \gamma_{21}$ can be viewed as the regression coefficient, *corrected for attenuation*, on $\widetilde{x}_2$ in a univariate OLS regression of $\widetilde{x}_1$ on $\widetilde{x}_2$. With that interpretion, the two elements in the vector on the right hand side of equation (16) can be viewed as weighted regression coefficients that show the impact of each of the two variables on the dependent variable, y. The weights are $1 - \omega s_1^2$ and ωs_1^2, respectively. The first element captures the direct effect of x_1 on y, through the coefficient β_1. The second

[9] As discussed in Section 1, this case is examined in Abel (2017) in the context of capital investment and its dependence on q and cash flow.

element captures the indirect effect of x_2 on y; specifically, an increase in x_2 is associated with an increase in x_1 of $(1+s_2^2)\,\gamma_{21}$, which then affects y by $(1+s_2^2)\,\gamma_{21}\beta_1$. The larger is the measurement error in $\widetilde{x}_1$, that is, the larger is s_1^2, the smaller is the weight $1-\omega s_1^2$ on β_1 in the expression for b_1 and the larger is the weight ωs_1^2 on $(1+s_2^2)\,\gamma_{21}\beta_1$ in the expression for b_2.[10] Even though x_2 is irrelevant, that is, has no direct effect on y in the underlying linear model, it will have a nonzero coefficient provided that $\widetilde{x}_1$ is measured with error and x_1 and x_2 have nonzero correlation. In effect, $\widetilde{x}_2$ provides information about the true value of the mismeasured relevant variable x_1.

4.2 An Irrelevant Regressor That Is Uncorrelated with Relevant Variables

In subsection 4.1, I showed that an irrelevant variable, $\widetilde{x}_2$, will have a nonzero OLS coefficient if it is correlated with a relevant variable, $\widetilde{x}_1$, that is measured with error. However, when $k>2$, it is possible for an irrelevant variable, $\widetilde{x}_3$, to have a nonzero OLS coefficient even if it is not correlated with any relevant variable. I illustrate this possibility in a case with $k=3$, $\beta_1\neq 0$, $\beta_2=\beta_3=0$, and $cov\,(x_1,x_3)=0$. With $k=3$ and $\beta_2=\beta_3=0$, equation (15) implies that

$$b=\frac{1}{1+\omega s_1^2 s_2^2 s_3^2}\begin{bmatrix}\left(1-s_1^2\frac{\omega}{\omega_1}\right)\beta_1\\ s_1^2\frac{\omega}{\omega_1}\left(1+\omega_1 s_2^2 s_3^2\right)\gamma_{21}\beta_1\\ s_1^2\frac{\omega}{\omega_1}\left(1+\omega_1 s_2^2 s_3^2\right)\gamma_{31}\beta_1\end{bmatrix},\quad \text{if } \beta_2=\beta_3=0. \tag{17}$$

The regression coefficients γ_{21} and γ_{31} in the vector on the right hand side of equation (17) are the OLS coefficients in an auxiliary regression of $\widetilde{x}_1$ on $\widetilde{x}_2$ and $\widetilde{x}_3$. These coefficients are presented in the following lemma.

Lemma 4 *Let γ_{21} and γ_{31} be the OLS coefficients on $\widetilde{x}_2$ and $\widetilde{x}_3$, respectively, in an OLS*

[10] An increase in s_1^2 will decrease the measurement error multiplier ω for given s_2^2 and R_{12}^2, but the reduction in ω is outweighed by the increase in s_1^2. To verify that an increase in s_1^2 increases ωs_1^2 and decreases $1-\omega s_1^2$, use equation (8) obtain $\omega s_1^2=\frac{s_1^2}{1+s_1^2+s_2^2-R_{12}^2}$ and $1-\omega s_1^2=\frac{1+s_2^2-R_{12}^2}{1+s_1^2+s_2^2-R_{12}^2}$. For given s_2^2 and R_{12}^2, an increase in s_1^2 reduces $1-\omega s_1^2$ and hence increases ωs_1^2, as asserted in the text.

regression of $\widetilde{x}_1$ *on* $\widetilde{x}_2$ *and* $\widetilde{x}_3$*. Then*

$$\begin{bmatrix} \gamma_{21} \\ \gamma_{31} \end{bmatrix} = \frac{1}{(1+s_2^2)(1+s_3^2) - R_{2,3}^2} \begin{bmatrix} (1+s_3^2)\,\theta_{2,1} - \theta_{2,3}\theta_{3,1} \\ (1+s_2^2)\,\theta_{3,1} - \theta_{3,2}\theta_{2,1} \end{bmatrix},$$

where $\theta_{i,j} \equiv \frac{cov(x_i,x_j)}{var(x_i)}$ *is the coefficient on* x_i *in a univariate OLS regression of* x_j *on* x_i*, and* $R_{2,3}^2 \equiv \frac{[cov(x_2,x_3)]^2}{var(x_2)var(x_3)}$ *is the squared correlation of* x_2 *and* x_3*.*

Proposition 2 *Suppose that (1)* $\beta_1 \neq 0$ *and* $\beta_2 = \beta_3 = 0$*; (2)* $cov\,(\widetilde{x}_1, \widetilde{x}_3) = 0$*,* $cov\,(\widetilde{x}_2, \widetilde{x}_3) \neq 0$*, and* $cov\,(\widetilde{x}_1, \widetilde{x}_2) \neq 0$*; and (3)* $s_1^2 > 0$*. Then*

$$\begin{bmatrix} \gamma_{21} \\ \gamma_{31} \end{bmatrix} = \frac{1}{(1+s_2^2)(1+s_3^2) - R_{2,3}^2} \begin{bmatrix} (1+s_3^2)\,\theta_{2,1} \\ -\theta_{3,2}\theta_{2,1} \end{bmatrix},$$

so γ_{21} *and* γ_{31} *are both nonzero and the plims of the OLS coefficients,* b_1*,* b_2*, and* b_3*, are all nonzero.*

Proposition 2 is a cautionary note about including irrelevant regressors in a regression. Even if an irrelevant regressor is uncorrelated with all relevant regressors, the irrelevant regressor can have a nonzero coefficient if it is correlated with another irrelevant regressor that is correlated with a relevant regressor measured with error. Therefore, the null hypothesis $H_0 : \beta_i = 0$ can be falsely rejected when $\widetilde{x}_i$ is an irrelevant variable (measured with or without error), even if it is uncorrelated with all relevant regressors.

5 Conclusion

I examine the impact of classical measurement error in OLS regressions with several regressors. The framework has three fundamental objects: (1) the parameter vector β, which characterizes the linear relation between the true values of the regressors and the dependent variable; (2) the variance-covariance matrix, V, of the vector of true regressors; and (3) the diagonal matrix Σ, which contains the variances of the classical measurement errors in each regressor. These measurement errors are uncorrelated with each other and with the true regressors.

I show that when there is more than one regressor, measurement error can impact OLS coefficient estimates through two distinct channels that I call multivariate attenuation and weight shifting. The strength of these channels is related to the measurement error multiplier ω, which is the reciprocal of $\frac{\det(V+\Sigma)-\det\Sigma}{\det V_D}$, where V_D is a diagonal matrix obtained from V by setting all the off-diagonal elements of V equal to zero. The measurement error multiplier ω gauges the amount of independent variation in the mismeasured regressors that is due to measurement error. It indicates the impact of measurement error on the estimated coefficients through both the multivariate attenuation channel and weight shifting channel.

The multivariate attenuation factor is a scalar $\frac{1}{1+\omega\det S}$ where $S = V_D^{-1}\Sigma$ is a diagonal matrix with elements equal to the ratio of the variance of measurement error to the variance of the true regressor. The multivariate attenuation factor, which is less than or equal to one, applies to all estimated coefficients. If any regressor is measured without error, then the multivariate attenuation factor equals one, so there is no multivariate attenuation; otherwise, the multivariate attenuation factor is less than one.

In addition to multivariate attenuation, measurement error in a regressor further attenuates the estimated coefficient on that variable and shifts weight to other regressors. If a particular variable is measured with error, it leaves room for other regressors that are correlated with that variable (or correlated with other regressors that are correlated with that variable) to pick up some of the impact of the particular variable. The measurement error multiplier ω affects the degree of this shifting. Weight shifting can lead to false rejections of the null hypothesis that an irrelevant variable has a zero coefficient, even when that irrelevant variable is uncorrelated with all of the relevant variables.

References

Abel, Andrew B., "The Analytics of Investment, q, and Cash Flow in the Presence of Measurement Error," January 2017, Journal of Financial Economics, forthcoming.

Garber, Steven and Steven Klepper, "Extending the Classical Errors-in-Variables Model," Econometrica, 48, 6 (September 1980), pp. 1541-1546.

Greene, William H., Econometric Analysis, Upper Saddle River, NJ: Pearson Education, Inc., Fifth Edition, 2003.

Griliches, Zvi, "Economic Data Issues," Chapter 25 in Handbook of Econometrics, Volume III, edited by Z. Griliches and M.D. Intriligator, Elsevier Science Publishers BV, 1986.

Levi, Maurice D., "Errors in the Variables Bias in the Presence of Correctly Measured Variables," Econometrica, 41, 5 (September 1973), pp. 985-986.

Roberts, Michael R. and Whited, Toni M., "Endogeneity in Empirical Corporate Finance," in George Constantinides, Milt Harris, and Rene Stulz, Handbook of the Economics of Finance, Elsevier, 2013, pp. 487-566.

Suvrit (https://mathoverflow.net/users/8430/suvrit), Determinant of sum of positive definite matrices, URL (version: 2011-05-20): https://mathoverflow.net/q/65501

Theil, H., Economic Forecasts and Policy, (second edition) Amsterdam: North Holland Publishing Company, 1965.

Wilks, Samuel S., Mathematical Statistics, New York: John Wiley and Sons, Inc., 1962.

Wooldridge, Jeffrey M., Econometric Analysis of Cross Section and Panel Data, Cambridge, MA: MIT Press, 2002.

Proofs

Proof. of Lemma 1. Statement 1: Since V_D is a diagonal matrix with all diagonal elements strictly positive, $\det V_D > 0$. Therefore, it suffices to prove that $\det(V+\Sigma) > \det\Sigma$. The following proof of this statement is taken from Suvrit (2011) with some minor adaptation. Since V is positive definite, $z'(V+\Sigma)z = z'Vz + z'\Sigma z > z'\Sigma z$ for every non-zero $k \times 1$ vector of real numbers. Therefore, $-z'(V+\Sigma)z < -z'\Sigma z$ for $z \neq 0$, and hence $(2\pi)^{-k/2}\int_{-\infty}^{\infty}\exp\left(-\frac{1}{2}z'(V+\Sigma)z\right)dz < (2\pi)^{-k/2}\int_{-\infty}^{\infty}\exp\left(-\frac{1}{2}z'\Sigma z\right)dz$. Now use the Gaussian integral, $(2\pi)^{-k/2}\int_{-\infty}^{\infty}\exp\left(-\frac{1}{2}z'Az\right)dz = \det A^{-1/2}$ for any positive definite $k \times k$ matrix A, to rewrite this inequality as $\det(V+\Sigma)^{-1/2} < \det\Sigma^{-1/2}$, which implies $\det(V+\Sigma) > \det\Sigma$.

Statement 2: For $k \geq 2$, Statement 2 follows immediately from Statement 1 by considering a case with $k-1$ regressors instead of k regressors.

Statement 3: If $k=1$, then $V = V_D = var(x_1)$ and $\Sigma = \sigma_1^2$, so $\omega = \frac{var(x_1)}{\left(var(x_1)+\sigma_1^2\right)-\sigma_1^2} = 1$.

Statement 4: If $k=2$, then Statement 4 follows immediately from Statement 3 by considering the case with $k-1=1$ regressor.

Statement 5: Since $\det(V+\Sigma) = \omega^{-1}\det V_D + \det\Sigma$ and $\det(V_{-j}+\Sigma_{-j}) = \omega_j^{-1}\det V_{D,-j} + \det\Sigma_{-j}$, it follows that $\sigma_j^2\frac{\det(V_{-j}+\Sigma_{-j})}{\det(V+\Sigma)} = \sigma_j^2\frac{\omega_j^{-1}\det V_{D,-j}+\det\Sigma_{-j}}{\omega^{-1}\det V_D+\det\Sigma}$. Use the facts that $(\det V_D)^{-1}\det\Sigma = \det\left(V_D^{-1}\Sigma\right) = \det(S)$ and $(\det V_{D,-j})^{-1}\det\Sigma_{-j} = \det(S_{-j})$ to obtain $\sigma_j^2\frac{\det(V_{-j}+\Sigma_{-j})}{\det(V+\Sigma)} = \sigma_j^2\frac{\det V_{D,-j}}{\det V_D}\frac{\omega_j^{-1}+\det S_{-j}}{\omega^{-1}+\det S} = s_j^2\frac{\omega_j^{-1}+\det S_{-j}}{\omega^{-1}+\det S}$, where the second equality follows from the definition $s_j^2 \equiv \frac{\sigma_j^2}{var(x_j)}$ and the fact that $\frac{\det V_{D,-j}}{\det V_D} = \frac{1}{var(x_j)}$. Finally, multiply the numerator and denominator by ω and by ω_{-j} and use the fact that $s_j^2\det S_{-j} = \det S$ to obtain $\sigma_j^2\frac{\det(V_{-j}+\Sigma_{-j})}{\det(V+\Sigma)} = \frac{\omega}{\omega_j}s_j^2\frac{1+\omega_j\det S_{-j}}{1+\omega\det S} = \frac{\frac{\omega}{\omega_j}s_j^2+\omega\det S}{1+\omega\det S}$.

Statement 6: $s_j^2\frac{\omega}{\omega_j} = \frac{\sigma_j^2}{var(x_j)}\frac{\det V_D}{\det V_{D,-j}}\frac{\det(V_{-j}+\Sigma_{-j})-\det\Sigma_{-j}}{\det(V+\Sigma)-\det\Sigma} = \frac{\sigma_j^2\det(V_{-j}+\Sigma_{-j})-\sigma_j^2\det\Sigma_{-j}}{\det(V+\Sigma)-\det\Sigma}$ $= \frac{\sigma_j^2\det(V_{-j}+\Sigma_{-j})-\det\Sigma}{\det(V+\Sigma)-\det\Sigma}$ where the first equality follows from the definitions of s_j^2, ω, and ω_j; the second equality uses $var(x_j)\det V_{D,-j} = \det V_D$; and the third inequality uses $\sigma_j^2\det\Sigma_{-j} = \det\Sigma$. It suffices to prove that $\sigma_j^2\det(V_{-j}+\Sigma_{-j}) < \det(V+\Sigma)$. To do so, define the $k \times k$ matrix A_1 to be the matrix $V+\Sigma$, except that the jth diagonal element, $var(x_j)+\sigma_j^2$, is replaced by $var(x_j)$, and define the $k \times k$ matrix A_2 to be the matrix $V+\Sigma$, with jth row replaced by zeros everywhere except for the element in the jth column, which is replaced by σ_j^2. The matrices A_1 and A_2 are constructed so that $\det(V+\Sigma) = \det A_1 + \det A_2$

and $\det A_2 = \sigma_j^2 \det(V_{-j} + \Sigma_{-j})$. Therefore, $\det(V+\Sigma) - \sigma_j^2 \det(V_{-j} + \Sigma_{-j}) = \det A_1$. Observe that A_1 defined above is the variance-covariance matrix of the measured variables $\widetilde{x}_i$, $i = 1, .., k$, except that $\widetilde{x}_j$ is replaced by the true variable x_j. Hence, $\det A_1 > 0$, so $\det(V+\Sigma) - \sigma_j^2 \det(V_{-j} + \Sigma_{-j}) = \det A_1 > 0$. ■

Proof. of Lemma 2: Let $\Omega \equiv V + \Sigma$ be the $k \times k$ variance-covariance matrix of $[\widetilde{x}_1, ..., \widetilde{x}_k]'$. The (i,j) element of Ω is $\Omega_{ij} \equiv Cov(\widetilde{x}_i, \widetilde{x}_j)$. Let $\Omega(i;j)$ be the $(k-1) \times (k-1)$ matrix obtained by deleting the ith row and jth column of Ω, and let $\Omega(i,m;j,l)$ be the $(k-2) \times (k-2)$ matrix obtained by deleting the ith row, mth row, jth column and lth column of Ω. Let $\widetilde{x}_{-j}$ be the $(k-1) \times 1$ vector obtaining by deleting the jth element of $[\widetilde{x}_1, ..., \widetilde{x}_k]'$.

Define γ_k as the $(k-1) \times 1$ vector of coefficients on $\widetilde{x}_1, ..., \widetilde{x}_{k-1}$ in a regression of $\widetilde{x}_k$ on $\widetilde{x}_1, ..., \widetilde{x}_{k-1}$, so $\gamma_k = [\Omega(k,k)]^{-1} \lambda$, where $\lambda \equiv [Cov(\widetilde{x}_1, \widetilde{x}_k), ..., Cov(\widetilde{x}_{k-1}, \widetilde{x}_k)]'$ is a $(k-1) \times 1$ vector. Since $(-1)^{i+j} \det \Omega(i,k;j,k)$ is the (i,j) co-factor of the $(k-1) \times (k-1)$ matrix $\Omega(k,k)$, the (i,j) element of $[\Omega(k,k)]^{-1}$ is $\frac{1}{\det \Omega(k,k)} (-1)^{i+j} \det \Omega(i,k;j,k)$. Therefore, γ_{ik}, the ith element of $\gamma_k = [\Omega(k,k)]^{-1} \lambda$, is $\gamma_{ik} = \frac{1}{\det \Omega(k,k)} \sum_{j=1}^{k-1} (-1)^{i+j} \det \Omega(i,k;j,k) \lambda_j$, where $\lambda_j \equiv Cov(\widetilde{x}_j, \widetilde{x}_k)$ is the jth element of λ. Since $\det \Omega(k,k) = C_{kk}$, where C_{kk} is the (k,k) co-factor of Ω, the regression coefficient γ_{ik} is

$$\gamma_{ik} = \frac{1}{C_{kk}} \sum_{j=1}^{k-1} (-1)^{i+j} \det \Omega(i,k;j,k) \lambda_j. \tag{1}$$

To calculate the (i,k) co-factor of Ω, for $i < k$, which is $C_{ik} = (-1)^{i+k} \det \Omega(i,k)$, expand $\Omega(i,k)$ along its bottom row, $[Cov(\widetilde{x}_k, \widetilde{x}_1), ..., Cov(\widetilde{x}_k, \widetilde{x}_{k-1})] = \lambda'$, to obtain $C_{ik} = (-1)^{i+k} \sum_{j=1}^{k-1} (-1)^{k-1+j} \det \Omega(i,k;j,k) \lambda_j$. Therefore,

$$C_{ik} = -\sum_{j=1}^{k-1} (-1)^{i+j} \det \Omega(i,k;j,k) \lambda_j. \tag{2}$$

Substitute equation (2) into equation (1) to obtain

$$\frac{C_{ik}}{C_{kk}} = \{ \begin{array}{ll} -\gamma_{ik}, & \text{if } i \neq k \\ 1, & \text{if } i = k \end{array} . \tag{3}$$

Since the ordering of the regressors is arbitary, equation (3) implies that

$$\frac{C_{ij}}{C_{jj}} = \{ \begin{array}{ll} -\gamma_{ij}, & \text{if } i \neq j \\ 1, & \text{if } i = j \end{array} \tag{4}$$

and hence

$$CC_D^{-1} = I_k - \Gamma. \tag{5}$$

■

Proof. of Lemma 3. Statement 1: W is a diagonal matrix with jth diagonal element $0 \leq \frac{\omega}{\omega_j} s_j^2 < 1$, where the first inequality follows from Statements 1 and 2 of Lemma 1 along with $s_j^2 \geq 0$, and the second inequality is Statement 6 in Lemma 1. Therefore, $W < I$. A is a diagonal matrix with jth diagonal element $1 + \omega_j \det S_{-j} \geq 1$, where the inequality follows from $\omega_j > 0$ (Statement 2 in Lemma 1) and $\det S_{-j} \geq 0$. Therefore $I \leq A$.

Statement 2: Definition 4 of A implies that $A - I = diag\,(\omega_1 \det S_{-1}, ..., \omega_k \det S_{-k})$, which, along with Definition 3 of W, implies $(A - I)\,W = diag\,(\omega s_1^2 \det S_{-1}, ..., \omega s_k^2 \det S_{-k})$. Use the fact that $s_j^2 \det S_{-j} = \det S$ to obtain $(A - I)\,W = diag\,(\omega \det S, ..., \omega \det S) = (\omega \det S)\,I$.

Statement 3: Since $C_D = diag\,(\det(V_{-1} + \Sigma_{-1}), ..., \det(V_{-k} + \Sigma_{-k}))$ and $\Sigma = diag\,(\sigma_1^2, ..., \sigma_k^2)$, it follows that $\frac{1}{\det(V+\Sigma)} C_D \Sigma = diag\left(\frac{\sigma_1^2 \det(V_{-1}+\Sigma_{-1})}{\det(V+\Sigma)}, ..., \frac{\sigma_k^2 \det(V_{-k}+\Sigma_{-k})}{\det(V+\Sigma)}\right)$. Now use $\sigma_j^2 \frac{\det(V_{-j}+\Sigma_{-j})}{\det(V+\Sigma)} = \frac{\omega}{\omega_j} s_j^2 \frac{1+\omega_j \det S_{-j}}{1+\omega \det S}$ from Statement 5 in Lemma 1 to obtain $\frac{1}{\det(V+\Sigma)} C_D \Sigma = diag\left(\frac{\omega}{\omega_1} s_1^2 \frac{1+\omega_1 \det S_{-1}}{1+\omega \det S}, ..., \frac{\omega}{\omega_k} s_k^2 \frac{1+\omega_k \det S_{-k}}{1+\omega \det S}\right)$. Use the definitions of W and A to obtain $\frac{1}{\det(V+\Sigma)} C_D \Sigma = \frac{1}{1+\omega \det S} AW$, which is the equality to be proved. To prove the inequality, observe that $\frac{1}{\det(V+\Sigma)} C_D \Sigma$ is a diagonal matrix with jth diagonal element $\frac{\omega}{\omega_j} s_j^2 \frac{1+\omega_j \det S_{-j}}{1+\omega \det S} = \frac{\frac{\omega}{\omega_j} s_j^2 + \omega \det S}{1+\omega \det S} < 1$, where the equality follows from $s_j^2 \det S_{-j} = \det S$, and the inequality follows from Statement 6 in Lemma 1. Therefore, $\frac{1}{\det(V+\Sigma)} C_D \Sigma = \frac{1}{1+\omega \det S} AW < I$. ■

Proof. of Proposition 1: Substitute the expression for $(V + \Sigma)^{-1} \Sigma\beta$ on the right hand side of equation (13) into equation (6) to obtain $b = \beta - \frac{1}{1+\omega \det S} (I_k - \Gamma)\, AW\beta$, which can be rewritten as $b = \frac{1}{1+\omega \det S} \left[(1 + \omega \det S)\, I_k - (I_k - \Gamma)\, AW\right] \beta$. Use Statement 2 of Lemma 3 to substitute $(A - I_k)\,W$ for $(\omega \det S)\, I_k$ to obtain $b = \frac{1}{1+\omega \det S} \left[I_k + (A - I_k)\,W - (I_k - \Gamma)\, AW\right] \beta$, which can be simplified to $b = \frac{1}{1+\omega \det S} \left[I_k - W + \Gamma AW\right] \beta$. ■

Proof. of Lemma 4: The OLS coefficients on $\widetilde{x}_2$ and $\widetilde{x}_3$, respectively, in a regression of

$\widetilde{x}_1$ on $\widetilde{x}_2$ and $\widetilde{x}_3$ are

$$\begin{bmatrix} \gamma_{21} \\ \gamma_{31} \end{bmatrix} = \begin{bmatrix} var\left(x_2\right) + \sigma_2^2 & cov\left(x_2, x_3\right) \\ cov\left(x_2, x_3\right) & var\left(x_3\right) + \sigma_3^2 \end{bmatrix}^{-1} \begin{bmatrix} cov\left(x_2, x_1\right) \\ cov\left(x_3, x_1\right) \end{bmatrix}$$

since $cov\left(\widetilde{x}_i, \widetilde{x}_j\right) = cov\left(x_i, x_j\right)$ for $i \neq j$. Use the definition $s_i^2 \equiv \frac{\sigma_i^2}{var(x_i)}$ and rewrite the equation for $\begin{bmatrix} \gamma_{21} \\ \gamma_{31} \end{bmatrix}$ to obtain

$$\begin{bmatrix} \gamma_{21} \\ \gamma_{31} \end{bmatrix} = \begin{bmatrix} \left(1 + s_2^2\right) var\left(x_2\right) & cov\left(x_2, x_3\right) \\ cov\left(x_2, x_3\right) & \left(1 + s_3^2\right) var\left(x_3\right) \end{bmatrix}^{-1} \begin{bmatrix} cov\left(x_2, x_1\right) \\ cov\left(x_3, x_1\right) \end{bmatrix}.$$

Use the standard formula for the inverse of a matrix to obtain

$$\begin{aligned} \begin{bmatrix} \gamma_{21} \\ \gamma_{31} \end{bmatrix} &= \frac{1}{\left(1 + s_2^2\right) var\left(x_2\right)\left(1 + s_3^2\right) var\left(x_3\right) - \left[cov\left(x_2, x_3\right)\right]^2} \\ &\quad \times \begin{bmatrix} \left(1 + s_3^2\right) var\left(x_3\right) & -cov\left(x_2, x_3\right) \\ -cov\left(x_2, x_3\right) & \left(1 + s_2^2\right) var\left(x_2\right) \end{bmatrix} \begin{bmatrix} cov\left(x_2, x_1\right) \\ cov\left(x_3, x_1\right) \end{bmatrix}. \end{aligned}$$

Divide the numerator and denominator by $var\left(x_2\right) var\left(x_3\right)$ to obtain

$$\begin{aligned} \begin{bmatrix} \gamma_{21} \\ \gamma_{31} \end{bmatrix} &= \frac{1}{\left(1 + s_2^2\right)\left(1 + s_3^2\right) - R_{2,3}^2} \\ &\quad \times \begin{bmatrix} \left(1 + s_3^2\right) \frac{1}{var(x_2)} & -cov\left(x_2, x_3\right) \frac{1}{var(x_2)var(x_3)} \\ -cov\left(x_2, x_3\right) \frac{1}{var(x_2)var(x_3)} & \left(1 + s_2^2\right) \frac{1}{var(x_3)} \end{bmatrix} \begin{bmatrix} cov\left(x_2, x_1\right) \\ cov\left(x_3, x_1\right) \end{bmatrix}. \end{aligned}$$

Perform the indicated multiplication to obtain

$$\begin{aligned} \begin{bmatrix} \gamma_{21} \\ \gamma_{31} \end{bmatrix} &= \frac{1}{\left(1 + s_2^2\right)\left(1 + s_3^2\right) - R_{2,3}^2} \\ &\quad \times \begin{bmatrix} \left(1 + s_3^2\right) \frac{1}{var(x_2)} cov\left(x_2, x_1\right) - cov\left(x_2, x_3\right) \frac{1}{var(x_2)var(x_3)} cov\left(x_3, x_1\right) \\ \left(1 + s_2^2\right) \frac{1}{var(x_3)} cov\left(x_3, x_1\right) - cov\left(x_2, x_3\right) \frac{1}{var(x_2)var(x_3)} cov\left(x_2, x_1\right) \end{bmatrix}. \end{aligned}$$

Finally, define the coefficient on x_i in a univariate OLS regression of x_j on x_i as $\theta_{i,j} \equiv \frac{cov(x_i, x_j)}{var(x_i)}$

and rewrite the expression for $\begin{bmatrix} \gamma_{21} \\ \gamma_{31} \end{bmatrix}$ as

$$\begin{bmatrix} \gamma_{21} \\ \gamma_{31} \end{bmatrix} = \frac{1}{\left(1+s_2^2\right)\left(1+s_3^2\right) - R_{2,3}^2} \begin{bmatrix} \left(1+s_3^2\right)\theta_{2,1} - \theta_{2,3}\theta_{3,1} \\ \left(1+s_2^2\right)\theta_{3,1} - \theta_{3,2}\theta_{2,1} \end{bmatrix}.$$

■

Proof. of Proposition 2: The assumption that $cov\left(\widetilde{x}_1, \widetilde{x}_3\right) = 0$ implies that $\theta_{3,1} = 0$. Substituting $\theta_{3,1} = 0$ into Lemma 4 yields

$$\begin{bmatrix} \gamma_{21} \\ \gamma_{31} \end{bmatrix} = \frac{1}{\left(1+s_2^2\right)\left(1+s_3^2\right) - R_{2,3}^2} \begin{bmatrix} \left(1+s_3^2\right)\theta_{2,1} \\ -\theta_{3,2}\theta_{2,1} \end{bmatrix},$$

which implies that γ_{21} and γ_{31} are both nonzero since $cov\left(x_1, x_2\right) = cov\left(\widetilde{x}_1, \widetilde{x}_2\right) \neq 0$ and $cov\left(x_2, x_3\right) = cov\left(\widetilde{x}_2, \widetilde{x}_3\right) \neq 0$ and hence $\theta_{2,1} \neq 0$ and $\theta_{3,2} \neq 0$. Substituting $\beta_1 \neq 0$, $\gamma_{21} \neq 0$, $\gamma_{31} \neq 0$ and $s_1^2 > 0$ into equation (17) and using $0 < \frac{\omega}{\omega_j} s_j^2 < 1$ from Statements 1, 2, and 6 of Lemma 1 implies that b_1, b_2, and b_3 are all nonzero.