How measurement error affects inference in linear regression

Erik Meijer[1] · Edward Oczkowski[2] · Tom Wansbeek[3]

Abstract

Measurement error biases OLS results. When the measurement error variance in absolute or relative (reliability) form is known, adjustment is simple. We link the (known) estimators for these cases to GMM theory and provide simple derivations of their standard errors. Our focus is on the test statistics. We show monotonic relations between the t-statistics and R^2s of the (infeasible) estimator if there was no measurement error, the inconsistent OLS estimator, and the consistent estimator that corrects for measurement error and show the relation between the t-value and the magnitude of the assumed measurement error variance or reliability. We also discuss how standard errors can be computed when the measurement error variance or reliability is estimated, rather than known, and we indicate how the estimators generalize to the panel data context, where we have to deal with dependency among observations. By way of illustration, we estimate a hedonic wine price function for different values of the reliability of the proxy used for the wine quality variable.

✉ Tom Wansbeek
   t.j.wansbeek@rug.nl

1   University of Southern California, Los Angeles, USA

2   Charles Sturt University, Wagga Wagga, Australia

3   Faculty of Economics and Business, University of Groningen, Nettelbosje 2, 9747 AE Groningen, The Netherlands

1 Introduction

As is well known from econometric textbooks (e.g., Baltagi 2011, sec. 5.3), measurement error in one or more regressors makes OLS estimators of linear regression models inconsistent. Often, the inconsistency will cause a bias toward zero, although this does not need not be the case and the bias can be away from zero (Wansbeek and Meijer 2000, sec. 2.3). But whatever the direction of the bias, the desire "to do something" about it has spawned a huge literature since the 1930s.

One strand in this literature is to limit the problem by deriving (asymptotic) bounds on the estimators, thus limiting the extent of the problem. In case the measurement error is confined to a single regressor, OLS is biased toward zero while reverse regression is biased away from zero, thus offering estimated bounds on the coefficient. This classical result (Frisch 1934) does not extend to the case of multiple mismeasured regressors, and then, outside information in the form of a bound on the measurement error covariance matrix is required to obtain estimators of bounds on the coefficients (Wansbeek and Meijer 2000, secs. 3.4 and 3.5).

But, not surprisingly, the focus in the literature is on coming up with a consistent estimator. One way to achieve this is through an instrumental variable. It may come from outside the model, but can also be found within the model as long as it is identified. This requires nonnormality of the regressors. Then, higher moments of the variables can be used as instruments (Geary 1942; Erickson and Whited 2002).

Another road to consistency lies open when the measurement error variance is known. Then, a consistent estimator is readily obtained by subtracting the measurement error covariance matrix from the covariance matrix of the observed regressors. Unlike in fields like physics and the medical sciences, in economics the measurement error variance is seldom known. Yet, researchers may have an idea about it or just want to understand how their results vary with its magnitude. In practice, it will often not so much be about the absolute magnitude but the magnitude relative to the observed variance, the *reliability*. For example, published psychological tests are routinely accompanied by a statement of its reliability (Fuller 1987, p. 5) and in their overview of measurement error in economics; Bound et al. (2001) present many of the results as reliabilities or correlation between the observed value and the true value (the square roots of the reliabilities), although they present results on measurement error variances as well. As this illustrates, ideas about reliability may involve fixed numbers, but in practice it will often be about numbers imported from previous research and are hence subject to sampling error, which depending on the relative sample sizes of the prior studies and current study may or may not be negligible. Buonaccorsi (2010, pp. 168–169) provides a critical assessment of the usefulness of externally estimated reliabilities.

In this paper, we consider inference for the linear regression model with measurement error in the context of these three increasingly realistic kinds of prior knowledge: known absolute variances, known reliabilities, and estimated reliabilities (or estimated measurement error variances). We will consider these three consecutively. For each case, we derive a consistent estimator of the regression coefficient and its asymptotic variance, both without and with assuming normality of the measurement error variance.

An interesting issue concerns t-values. We can distinguish three: the t-value when there were no measurement error; the t-value when there is measurement error but it is neglected; and the t-value when the measurement error is accounted for. For the first two cases, known absolute variances and known reliabilities, we show that the t-values decrease while we move along this list. (The generality of the third case, estimated reliability or measurement error variance, defies analysis.) This greatly expands the findings in Meijer and Wansbeek (2000). The issue is relevant for applied researchers for two reasons. First, a regression coefficient can become insignificant due to measurement error and, second, correcting for the measurement error will not make an insignificant coefficient significant.

The paper is organized as follows. In Sect. 2, we consider the case of known measurement error variance. We describe the model, present the adapted estimator when the measurement error variance is known and show that it is a method-of-moments (MM) estimator. We discuss estimating its variance in general and elaborate this under independence and normality. Section 3 discusses the F and t tests and shows that there is an ordering from high to low between the cases mentioned above (no measurement error, neglected measurement error, measurement error accounted for).

In Sect. 4, we turn to the case where the reliability rather than the measurement error variance is known. We derive the asymptotic variance, without and with assuming normality of the measurement errors. Analogous to Sect. 3, the issue of the ordering of the corresponding test statistics is addressed in Sect. 5.

Section 6 discusses how to handle the situation where the measurement error variance or reliability is not known but is consistently estimable, with a consistently estimated asymptotic variance.

In previous papers (Meijer et al. 2015, 2017), we have shown that panel data offer many additional possibilities for identification and estimation of measurement error models, compared to (independent) cross-sectional data. Therefore, in Sect. 7, we investigate whether the analysis up until then can be extended from a cross-sectional to a panel data context, and whether for the case of known or estimated measurement error variances or reliabilities, this makes identification and estimation easier or more difficult.

We then turn, in Sect. 8, to an empirical example. It concerns the Australian wine market. The price of wine is regressed on a number of variables including quality. Results are shown for different values of the reliability of the proxy variable that is used to quantify quality. Some concluding remarks are made in Sect. 9.

Throughout, we consider the linear regression model, which has also received most attention in the measurement error literature. A recent overview of measurement error focusing on nonlinear models is provided by Schennach (2016). Extending our results to nonlinear models is left for future research. It turns out that our estimators of the coefficients and our expressions for their asymptotic variances and the estimators of them are essentially the same as the ones presented in Fuller (1987, sec. 3.1.1), although this relation is far from obvious and our presentation is simpler and more in line with the economic literature. Some of our special cases and extensions are new, and in particular, our main contribution to the literature is given by the results comparing the magnitudes of test statistics.

2 Measurement error variance known

In this section, we introduce the model that we will study throughout and consider the case where the measurement error variance is known. Most of the results here have been described before in the literature (e.g., Fuller 1987; Wansbeek and Meijer 2000), but we present them here concisely as a reference for the rest of the paper, and we put this in a (Generalized) Method of Moments framework, which simplifies the theoretical analyses.

Consider the following linear regression model with k regressors ξ_i, measured with error v_i:

$$y_i = \xi_i' \beta + \varepsilon_i$$
$$x_i = \xi_i + v_i,$$

for $i = 1, \ldots, n$. The y_i and x_i are observed. The reduced form, after eliminating the unobserved ξ_i, is

$$y_i = x_i' \beta + u_i$$
$$u_i = \varepsilon_i - v_i' \beta.$$

Initially, assume that the observations are i.i.d. and that ε_i, ξ_i, and v_i are mutually independent with means 0, μ, and 0, respectively. We let

$$\sigma^2 = \mathbb{E}(\varepsilon_i^2) \tag{1}$$

and $\Omega = \mathbb{E}(v_i v_i')$. So

$$\sigma_u^2 = \mathbb{E}(u_i^2) = \sigma^2 + \beta' \Omega \beta. \tag{2}$$

We collect the y_i in the n-vector y and the x_i in the $n \times k$-matrix X. Let $\hat{A} = X'X/n$ and $A = \plim_{n \to \infty} \hat{A} = \mathbb{E}(\hat{A}) = \mathbb{E}(x_i x_i')$.

As is well known, major implication of this model is that the OLS estimator $\hat{\beta}_0 = (X'X)^{-1}X'y$ of β converges to $\beta_0 = A^{-1}(A - \Omega)\beta$ and is hence inconsistent, except for the trivial case that $\Omega\beta = 0$, or, equivalently, $\Omega\beta_0 = 0$. In the following, we assume that $\Omega\beta \neq 0$, that is, the model includes at least one mismeasured variable. If Ω is known, the inconsistency is easily removed by using the adapted OLS estimator

$$\hat{\beta} = (X'X - n\Omega)^{-1}X'y. \tag{3}$$

Let

$$h_i = x_i y_i - (x_i x_i' - \Omega)\beta. \tag{4}$$

Then the model assumptions imply $\mathbb{E}(h_i) = 0$, so (4) is a set of k valid moment equations. Solving $\bar{h} = 0$, with

$$\bar{h} = \frac{1}{n}\sum_i h_i = \frac{1}{n}[X'y - (X'X - n\Omega)\beta],$$

shows that the estimator $\hat{\beta}$ in (3) is a method of moments (MM) estimator.

2.1 Residual variance

The OLS-based estimator of the residual variance σ^2,

$$\hat{\sigma}_0^2 = (y - X\hat{\beta}_0)'(y - X\hat{\beta}_0)/n = y'y/n - \hat{\beta}_0'\hat{A}\hat{\beta}_0, \tag{5}$$

is also inconsistent when there is measurement error. Since $\mathbb{E}(y_i^2) = \sigma^2 + \beta'(A-\Omega)\beta$, we have

$$\plim_{n\to\infty} y'y/n = \sigma^2 + \beta'(A - \Omega)\beta, \tag{6}$$

so

$$\sigma_0^2 = \plim_{n\to\infty} \hat{\sigma}_0^2 = \sigma^2 + \beta'(A - \Omega)\beta - \beta'(A - \Omega)A^{-1}(A - \Omega)\beta > \sigma^2.$$

The strictness of the inequality is an implication of the assumption $\Omega\beta \neq 0$. Through (6) we obtain

$$\hat{\sigma}^2 = y'y/n - \hat{\beta}'(\hat{A} - \Omega)\hat{\beta}. \tag{7}$$

as a consistent estimator of σ^2.

2.2 Explained variation

We now consider the effect on R^2 and the way to correct it. Let σ_y^2 be the population variance of the y_i. This is consistently estimated by the sample variance s_y^2. Furthermore, let R_0^2 be the R^2 from the OLS regression and let $\rho_*^2 = (\sigma_y^2 - \sigma^2)/\sigma_y^2 = [\beta'(A - \Omega)\beta - \mu_y^2]/\sigma_y^2$ be the population R^2 of the regression of y_i on ξ_i, where $\mu_y = \mathbb{E}(y_i)$. Then

$$R_0^2 = \frac{s_y^2 - \hat{\sigma}_0^2}{s_y^2} = \frac{\hat{\beta}_0'\hat{A}\hat{\beta}_0 - \bar{y}^2}{s_y^2} \xrightarrow{p} \frac{\beta'(A - \Omega)A^{-1}(A - \Omega)\beta - \mu_y^2}{\sigma_y^2}$$

$$< \frac{\beta'(A - \Omega)\beta - \mu_y^2}{\sigma_y^2} = \rho_*^2.$$

So R^2 is underestimated when there is measurement error, but

$$\hat{\rho}_*^2 = \frac{\hat{\beta}'(\hat{A} - \Omega)\hat{\beta} - \bar{y}^2}{s_y^2} \tag{8}$$

is a consistent estimator of ρ_*^2.

2.3 Generalization

The assumptions we have stated above can be weakened without losing consistency of the estimator. Under weak regularity conditions, the MM estimator is consistent if $\mathbb{E}(h_i) = 0$ (or, even weaker, $\underset{n \to \infty}{\text{plim}}\, \bar{h} = 0$). A set of sufficient conditions for this is (a) $\mathbb{E}(\xi_i \varepsilon_i) = 0$, (b) $\mathbb{E}(\xi_i v_i') = 0$, (c) $\mathbb{E}(v_i \varepsilon_i) = 0$, and (d) $\mathbb{E}(v_i v_i') = \Omega$. This weaker set allows for dependence across observations (time series, panel data, clustered data) and heteroskedasticity in ε_i. It also allows for heteroskedasticity in v_i but the assumption that $\mathbb{E}(v_i v_i') = \mathbb{E}_\xi \big[\mathbb{E}_{v|\xi}(v_i v_i' \mid \xi_i)\big]$ is known but $\mathbb{E}_{v|\xi}(v_i v_i' \mid \xi_i)$ varies with ξ_i does not seem to offer much additional practical value. However, we will discuss extensions to the case where Ω is consistently estimated later, and in that situation, robustness to heteroskedasticity in v_i may be a desirable property.

2.4 The asymptotic variance

Since $\underset{n \to \infty}{\text{plim}}\, \partial \bar{h}/\partial \beta' = -(A - \Omega)$, MM theory implies that the asymptotic variance of $\hat{\beta}$ is

$$\text{avar}(\hat{\beta}) = (A - \Omega)^{-1}\big[\mathbb{E}(h_i h_i')\big](A - \Omega)^{-1}.$$

A consistent estimator of this is

$$\widehat{\text{avar}}(\hat{\beta}) = (\hat{A} - \Omega)^{-1}\big[\hat{\mathbb{E}}(h_i h_i')\big](\hat{A} - \Omega)^{-1}, \tag{9}$$

with

$$\hat{\mathbb{E}}(h_i h_i') = \frac{1}{n}\sum_i \hat{h}_i \hat{h}_i', \tag{10}$$

with $\hat{h}_i = x_i y_i - (x_i x_i' - \Omega)\hat{\beta}$. This expression was previously given in section 5.4.2 of Buonaccorsi (2010). Note that (9) is valid under heteroskedasticity of ε_i and v_i. With clustered data or other types of dependent data, the appropriate clustered or heteroskedasticity and autocorrelation consistent (HAC) covariance matrix replaces the covariance matrix (10).

We can elaborate (9) when the measurement errors are normally distributed. Then

$$\Psi = \mathbb{E}[(v_i'\beta)^2 v_i v_i'] = (\beta'\Omega\beta)\Omega + 2\Omega\beta\beta'\Omega.$$

With

$$h_i = x_i(y_i - x_i'\beta) + \Omega\beta = (\xi_i + v_i)(\varepsilon_i - v_i'\beta) + \Omega\beta = \xi_i\varepsilon_i - \xi_i v_i'\beta + v_i\varepsilon_i$$
$$- v_i v_i'\beta + \Omega\beta,$$

we obtain, using (2),

$$\begin{aligned}\mathbb{E}(h_i h_i') &= \sigma^2(A - \Omega) + \beta'\Omega\beta(A - \Omega) + \sigma^2\Omega + \Psi - \Omega\beta\beta'\Omega \\ &= (\sigma^2 + \beta'\Omega\beta)A + \Omega\beta\beta'\Omega + [\Psi - (\beta'\Omega\beta)\Omega - 2\Omega\beta\beta'\Omega] \\ &= \sigma_u^2 A + \Omega\beta\beta'\Omega,\end{aligned} \tag{11}$$

leading to

$$\mathrm{avar}(\hat{\beta}) = (A - \Omega)^{-1}(\sigma_u^2 A + \Omega\beta\beta'\Omega)(A - \Omega)^{-1}. \tag{12}$$

To make this operational, we need to replace parameters by consistent estimators. In particular, a consistent estimator of σ_u^2 is

$$\hat{\sigma}_u^2 = \hat{\sigma}^2 + \hat{\beta}'\Omega\hat{\beta}; \tag{13}$$

it can be straightforwardly verified that this is equal to $\sum_i \hat{u}_i^2/n$. So

$$\widehat{\mathrm{avar}}(\hat{\beta}) = (\hat{A} - \Omega)^{-1}(\hat{\sigma}_u^2 \hat{A} + \Omega\hat{\beta}\hat{\beta}'\Omega)(\hat{A} - \Omega)^{-1} \tag{14}$$

is a simple-structured consistent estimator when the measurement errors are normal and ξ_i, ε_i, and v_i are mutually independent.

3 Ordering of test statistics

We now turn to hypothesis testing. To obtain tractable results, we maintain the hypothesis that the measurement errors are normally distributed. Let U be a $k \times p$ matrix of full column rank, with $p < k$. Let $\tilde{\beta}$ be an estimator of β and $\tilde{V}$ be an estimator of its asymptotic variance matrix. Then a Wald test statistic for $H_0 : U'\beta = 0$ is

$$\tilde{T} = n\tilde{\beta}'U(U'\tilde{V}U)^{-1}U'\tilde{\beta}.$$

This is compared to a Chi-square distribution with p degrees of freedom. For comparing the test statistics based on different estimators, we compare their probability limits (scaled by n), $\tilde{\tau} = \plim_{n \to \infty} n^{-1}\tilde{T}$. In this comparison, we include the infeasible OLS estimator based on observing the ξ_i. For this infeasible estimator, the inconsistent OLS estimator, and the consistent MM estimator, we obtain in this order

$$\tau_\dagger = \frac{\beta' U \left[U'(A - \Omega)^{-1} U \right]^{-1} U' \beta}{\sigma^2} = \frac{\beta' Q_\dagger \beta}{\sigma^2}, \tag{15}$$

$$\tau_0 = \beta_0' U \left[U'(\sigma_0^2 A^{-1}) U \right]^{-1} U' \beta_0$$
$$= \frac{\beta'(A - \Omega) A^{-1} U (U' A^{-1} U)^{-1} U' A^{-1} (A - \Omega) \beta}{\sigma_0^2} = \frac{\beta' Q_0 \beta}{\sigma_0^2}, \tag{16}$$

$$\tau_* = \beta' U \left[U' \left\{ (A - \Omega)^{-1} (\sigma_u^2 A + \Omega \beta \beta' \Omega)(A - \Omega)^{-1} \right\} U \right]^{-1} U' \beta$$
$$= \frac{\beta' U \left[U' \left\{ (A - \Omega)^{-1} A (A - \Omega)^{-1} \right\} U \right]^{-1} U' \beta}{\sigma_u^2} - c = \frac{\beta' Q_* \beta}{\sigma_u^2} - c, \tag{17}$$

with σ^2, σ_0^2, and σ_u^2 defined in (1), (7), and (2), respectively, and $Q_\dagger$, Q_0, Q_*, and $c > 0$ implicitly defined; the latter reflects the matrix $\Omega \beta \beta' \Omega$ in the expression of τ_* and its precise form is immaterial.

3.1 Relation between the test statistics

To handle the Qs, we use the result for matrices F and H such that (F, H) is nonsingular and $F'H = 0$, and nonsingular S,

$$F(F' S^{-1} F)^{-1} F' = S - SH(H' SH)^{-1} H' S. \tag{18}$$

To prove the result, notice that both sides equal $(F, 0)$ after postmultiplication by the nonsingular matrix $(S^{-1} F, H)$. Let G be an orthogonal complement of U and consider the case where G is such that $\Omega G = 0$ or, equivalently, $WG = AG$, where $W = A - \Omega$; we will meet two instances of this below. Now,

$$Q_\dagger = U(U' W^{-1} U)^{-1} U'$$
$$= W - WG(G' WG)^{-1} G' W$$
$$= W - AG(G' AG)^{-1} G' A$$
$$Q_0 = W A^{-1} U (U' A^{-1} U)^{-1} U' A^{-1} W$$
$$= W A^{-1} [A - AG(G' AG)^{-1} G' A] A^{-1} W$$
$$= W A^{-1} W - AG(G' AG)^{-1} G' A$$
$$Q_* = U(U' W^{-1} A W^{-1} U)^{-1} U'$$
$$= W A^{-1} W - W A^{-1} WG(G' W A^{-1} WG)^{-1} G' W A^{-1} W$$
$$= W A^{-1} W - AG(G' AG)^{-1} G' A.$$

So when $\Omega G = 0$, $Q_\dagger \geq Q_0$ since $W \geq W A^{-1} W$.[1] Also, $Q_0 = Q_*$. Since $\sigma^2 < \sigma_0^2$, cf. (7), and $\sigma_0^2 < \sigma_u^2$, cf. (7) and (2), we conclude $\tau_\dagger > \tau_0 > \tau_*$.

---

[1] The inequality sign means that the difference is positive semidefinite.

3.2 *F* and *t* test

The first instance of $\Omega G = 0$ occurs when testing the null hypothesis that all coefficients except the intercept are zero. The Wald test is then the asymptotic version of the standard F test. Let the kth element of x_i be 1 and let e_k be the k-th unit vector (the k-th column of I_k). The relevant statistic is obtained by letting U be I_k without its last column, with orthocomplement $G = e_k$; clearly, $\Omega G = 0$. Thus, the null hypothesis is rejected less often when using the OLS estimator based on the observed x_i than when using (if we could) the OLS estimator based on the true ξ_i. More interestingly and somewhat paradoxically (because the estimated coefficients are typically larger and the estimated residual variance is smaller), the null hypothesis is rejected less often when using the consistent MM estimator than when using the inconsistent OLS estimator based on the x_i. Hence, the finding of a significant relation may not survive when measurement error is accounted for.

   Another interesting aspect of the ordering of the statistics is that it clearly distinguishes between the case where there is no measurement error and the case where there is measurement error but its variance is known. From a first-order perspective, there is no difference as β can be (simply) estimated consistently in both cases, but in the latter case it is harder to detect a significant relationship between the variables.

   The other instance of $\Omega G = 0$ arises when there is measurement error in a single regressor only, the first one, say, and the null hypothesis is $\beta_1 = 0$. Then Ω is proportional to $e_1 e_1'$ and $U = e_1$ so G is I_k without its first column. The Wald test statistic is then the square of the t test statistic. The same ordering as above applies, with the same comments. This generalizes a result from Meijer and Wansbeek (2000). For the case of regression with a single regressor, the result $\tau_{\dagger} > \tau_0$ was already given by Bloch (1978).[2]

4 Known reliability

Information about measurement error variances, if available, is more likely to be of the relative than the absolute form. For example, Fuller (1987, Table 1.1.1), lists the reliability of a number of socioeconomic variables, as computed from repeated measurements by the U.S. Census Bureau. Income, for instance, has a reliability of 85%. Bound et al. (2001, sec. 6) list a large amount of empirical evidence about measurement error in surveys, and most (though not all) of this is presented in terms of correlations or variance ratios, which directly translate into reliabilities. By way of another example, after performing a factor analysis of the independence of central banks, De Haan et al. (2003) produced an indicator of the latent variable "central bank independence" and listed its (estimated) reliability.

---

[2] With a minor error; in our notation, the denominator in (16) in Bloch (1978) is not σ_0^2 but σ_u^2. This does not affect the inequality.

In this case, it is natural to assume that the measurement errors of the different variables are independent. So Ω is now a diagonal matrix, and we know

$$\rho_j = \frac{\text{var}(\xi_{ij})}{\text{var}(x_{ij})} = 1 - \frac{\text{var}(v_{ij})}{\text{var}(x_{ij})} = 1 - \frac{\Omega_{jj}}{A_{jj} - \mu_j^2},$$

for $j = 1, \ldots, k$. So now Ω depends on the (unknown) diagonal elements of A, as

$$\Omega_{jj} = (1 - \rho_j)(A_{jj} - \mu_j^2).$$

The means μ_j now enter the picture as unknown parameters, requiring their own moment conditions. Let, for $i = 1, \ldots, n$,

$$W_i = \text{diag}[(1 - \rho_j)(x_{ij} - \mu_j)^2], \tag{19}$$

so $\mathbb{E}(W_i) = \Omega$. Consider the moment conditions $\mathbb{E}(h_i) = 0$, with

$$h_i = \begin{pmatrix} h_{1i} \\ h_{2i} \end{pmatrix} = \begin{pmatrix} x_i y_i - (x_i x_i' - W_i)\beta \\ x_i - \mu \end{pmatrix} \quad \text{so} \quad \bar{h} = \frac{1}{n}\begin{pmatrix} X'y - (X'X - n\bar{W})\beta \\ n(\bar{x} - \mu) \end{pmatrix}, \tag{20}$$

with $\bar{x}$, $\bar{W}$, and $\bar{h}$ the sample averages. Setting $\bar{h} = 0$ and solving for β and μ readily gives $\hat{\mu} = \bar{x}$ and

$$\hat{\beta} = (X'X - n\hat{\Omega})^{-1}X'y, \tag{21}$$

with

$$\hat{\Omega} = \text{diag}\left[(1 - \rho_j)\sum_i(x_{ij} - \bar{x}_j)^2/n\right]. \tag{22}$$

Analogously, the consistent estimator of the error variance σ^2 is now

$$\hat{\sigma}^2 = y'y/n - \hat{\beta}'(\hat{A} - \hat{\Omega})\hat{\beta}. \tag{23}$$

instead of (7). Since

$$\plim_{n\to\infty} \frac{\partial\bar{h}}{\partial(\beta', \mu')} = \mathbb{E}\left[\frac{\partial h_i}{\partial(\beta', \mu')}\right] = -\begin{pmatrix} A - \Omega & 0 \\ 0 & I_k \end{pmatrix},$$

we have instead of (9)

$$\widehat{\text{avar}}(\hat{\beta}) = (\hat{A} - \hat{\Omega})^{-1}\left(\frac{1}{n}\sum_i \hat{h}_{1i}\hat{h}_{1i}'\right)(\hat{A} - \hat{\Omega})^{-1}, \tag{24}$$

with

$$\hat{h}_{1i} = x_i y_i - (x_i x_i' - \hat{W}_i)\hat{\beta}$$
$$\hat{W}_i = \text{diag}[(1 - \rho_j)(x_{ij} - \bar{x}_j)^2].$$

Expressions (21), (23), and (24) can be found in the Stata manual's description of its **eivreg** command as of version 16 (StataCorp 2019a).[3]

4.1 Estimation in a structural equation modeling program

The linear regression model with measurement error is a special case of the general class of *structural equation models* (SEMs); see, e.g., Wansbeek and Meijer (2000, ch. 8). Most general-purpose statistical software packages have a SEM module, and there are also standalone programs for estimating them. They generally allow simple restrictions on the parameters, so estimating the model with known measurement error variance in such a program is straightforward. Estimating the model with known reliability is slightly less straightforward, however. For example, the **sem** command in Stata allows for specifying the known reliability, but it then computes (in our notation) $\hat{\Omega}$ and treats this as known, instead of the reliability itself (StataCorp 2019b, p. 577), and Lockwood and McCaffrey (2020) report that this leads to noticeably biased standard errors and propose using the bootstrap or using the theory of M-estimation to obtain correct standard errors for this procedure. However, the proper way to specify known reliability in a SEM is to impose a linear relation between the variance of ξ and the variance of the relevant element(s) of v: $\text{var}(v_{ij}) = [(1 - \rho_j)/\rho_j]\,\text{var}(\xi_{ij})$, which in Stata's **sem** procedure can be done through a specification like

```
variance(xi1@c1 e.x1@(0.2*c1/0.8))
```

where `xi1` is ξ, `e.x1` is the measurement error of the error-ridden variable `x1` (i.e., v_1), `0.8` is ρ_j (and $0.2 = 1 - \rho_j$), and `c1` indicates a free parameter. In many other structural equation programs, such linear constraints can be imposed analogously.

4.2 Asymptotic variance

Analogous to what we did in Sect. 2.4 for the case of known Ω, we derive an explicit expression for the asymptotic variance of $\hat{\beta}$ for the case of known reliabilities. We assume homoskedasticity and normality of the v_i as we can obtain a manageable expression only then. Let $G = \text{diag}[(1 - \rho_j)\beta_j]$ and let

$$\dot{x}_i = x_i - \mu$$
$$\dot{A} = \mathbb{E}(\dot{x}_i \dot{x}_i') = A - \mu\mu'$$

---

[3] In earlier versions of Stata, the expression for the asymptotic variance of $\hat{\beta}$ was incorrect. This had an important qualitative effect on the result as the t statistic increased when correcting the estimator of β with Ω while it should decrease. See Lockwood and McCaffrey (2020) for a discussion of the problem with the earlier version of **eivreg**.

$$\dot{a} = \text{vec}(\dot{A}) = \mathbb{E}(\dot{x}_i \otimes \dot{x}_i)$$
$$H_k = \sum_j e_j \otimes e_j e_j',$$

with e_j the jth unit vector of dimension k. We can now write h_{1i} as

$$h_{1i} = x_i u_i + G H_k'(\dot{x}_i \otimes \dot{x}_i)$$

and want to find an expression for $\mathbb{E}(h_{1i} h_{1i}')$.

To do so, let $P_{k,k}$ be the symmetric commutation matrix[4] of order $k^2 \times k^2$. Thus, $P_{k,k} H_k = H_k$, $\mathbb{E}(x_i u_i) = \mathbb{E}(\dot{x}_i u_i) = -\Omega\beta$, $\mathbb{E}(x_i \dot{x}_i') = \dot{A}$, and

$$\Delta = (I_k \otimes \beta'\Omega) H_k G = \text{diag}(\Omega\beta) G = G \, \text{diag}(\Omega\beta).$$

Using the method of repeated conditioning (Merckens and Wansbeek 1989; Wansbeek and Meijer 2000, p. 366) we readily obtain

$$\mathbb{E}(x_i u_i u_i x_i') = \sigma_u^2 A + 2\Omega\beta\beta'\Omega$$
$$\mathbb{E}[G H_k'(\dot{x}_i \otimes \dot{x}_i)(\dot{x}_i \otimes \dot{x}_i)' H_k G] = G H_k'[(\dot{A} \otimes \dot{A})(I_{k^2} + P_{k,k}) + \dot{a}\dot{a}'] H_k G$$
$$= 2G(\dot{A} * \dot{A})G + \Omega\beta\beta'\Omega$$
$$\mathbb{E}[\dot{x}_i u_i (\dot{x}_i \otimes \dot{x}_i)' H_k G] = -\Omega\beta\dot{a}' - (\dot{A} \otimes \beta'\Omega)(I_{k^2} + P_{k,k}) H_k G$$
$$= -\Omega\beta\dot{a}' - 2\dot{A}(I_k \otimes \beta'\Omega) H_k G$$
$$= -\Omega\beta\beta'\Omega - 2\dot{A}\Delta,$$

where "$*$" denotes the Hadamard (element-wise) product of two matrices of equal dimensions. Collecting terms we obtain

$$\mathbb{E}(h_{1i} h_{1i}') = \sigma_u^2 A + \Omega\beta\beta'\Omega + 2[G(\dot{A} * \dot{A})G - \dot{A}\Delta - \Delta\dot{A}].$$

So, with hats as usual indicating the substitution of consistent estimators, we get

$$\widehat{\text{avar}}(\hat{\beta}) = (\hat{A} - \hat{\Omega})^{-1} \left(\hat{\sigma}_u^2 \hat{A} + \hat{\Omega}\hat{\beta}\hat{\beta}'\hat{\Omega} + 2[\hat{G}(\hat{\dot{A}} * \hat{\dot{A}})\hat{G} - \hat{\dot{A}}\hat{\Delta} - \hat{\Delta}\hat{\dot{A}}] \right) (\hat{A} - \hat{\Omega})^{-1}, \tag{25}$$

with now, slightly adapting from (2), $\hat{\sigma}_u^2 = \hat{\sigma}^2 + \hat{\beta}'\hat{\Omega}\hat{\beta} = \sum_i \hat{u}_i^2 / n$, with $\hat{\Omega}$ as given in (22). So the asymptotic variance for the case of known reliabilities is different from the one for the case of known Ω, cf. (14), and quite a bit more complex.

---

[4] Its defining property is $P_{k,k}(a \otimes b) = (b \otimes a)$, where a and b are arbitrary k-vectors; see, e.g., Wansbeek and Meijer (2000, p. 361), for some of its properties.

5 Test statistics in the case of known reliability

The results for the R^2 from the case with known Ω immediately carry over to the case with known reliability, except that in the computation of $\hat{\rho}_*^2$, $\hat{\Omega}$ is used instead of Ω. The results for the Wald test also carry over, but less trivially so.

For comparing Wald tests, τ_0 and $\tau_\dagger$ are the same as before, because they do not use any information about measurement error. However, the expression for τ_* is different now. Consider first the case of the joint test of whether all coefficients except the constant are zero, that is, the Wald version of the standard F test. As discussed above, this corresponds to U being the first $k-1$ columns of I_k and its complement being e_k. Define $\Sigma = U'(A - \mu\mu')U$ and $\Omega_1 = U'\Omega U$. That is, these are the variance matrices of x and v, respectively, with their last element (corresponding to the constant) omitted. Then

$$
\begin{aligned}
\tau_* &= \beta'U[U'(A - \Omega)^{-1}U]^{-1}\Gamma^{-1}[U'(A - \Omega)^{-1}U]^{-1}U'\beta \\
&= \beta'U(\Sigma - \Omega_1)\Gamma^{-1}(\Sigma - \Omega_1)U'\beta,
\end{aligned}
$$

where the last equality follows from Lemma 1 in "Appendix A",

$$
\Gamma = \sigma_u^2\Sigma + \Omega_1 U'\beta\beta'U\Omega_1 + 2[G_1(\Sigma * \Sigma)G_1 - \Sigma\Delta_1 - \Delta_1\Sigma],
$$

and G_1 and Δ_1 are the upper-left $(k-1)\times(k-1)$ submatrices of G and Δ, respectively, or, equivalently, $G_1 = U'GU$ and $\Delta_1 = U'\Delta U$. In contrast,

$$
\begin{aligned}
\tau_0 &= \frac{\beta'(A - \Omega)A^{-1}U(U'A^{-1}U)^{-1}U'A^{-1}(A - \Omega)\beta}{\sigma_0^2} \\
&= \frac{\beta'U(\Sigma - \Omega_1)\Sigma^{-1}(\Sigma - \Omega_1)U'\beta}{\sigma_0^2}.
\end{aligned}
$$

It follows that if $\Gamma \geq \sigma_0^2\Sigma$, then $\tau_0 \geq \tau_*$. Therefore, we investigate

$$
\begin{aligned}
\Theta = \Gamma - \sigma_0^2\Sigma &= (\beta'U\Omega_1\Sigma^{-1}\Omega_1U'\beta)\Sigma + \Omega_1U'\beta\beta'U\Omega_1 \\
&\quad + 2[G_1(\Sigma * \Sigma)G_1 - \Sigma\Delta_1 - \Delta_1\Sigma],
\end{aligned} \tag{26}
$$

where we have used

$$
\sigma_u^2 - \sigma_0^2 = \beta'\Omega A^{-1}\Omega\beta = \beta'U\Omega_1U'A^{-1}U\Omega_1U'\beta = \beta'U\Omega_1\Sigma^{-1}\Omega_1U'\beta,
$$

which again uses Lemma 1. After some algebra, we find that $\Theta = R'SR$, where S is a symmetric positive semidefinite matrix, which implies that Θ is a symmetric positive semidefinite matrix and therefore $\tau_\dagger > \tau_0 \geq \tau_*$. The matrices in this expression are

$$
\begin{aligned}
R &= (\Sigma^{\frac{1}{2}} \otimes \Sigma^{-\frac{1}{2}}\Omega_1U'\beta) \\
S &= Q_{k-1}M_L Q_{k-1} + (K - L)(K - L)'
\end{aligned}
$$

$$M_L = I_{(k-1)^2} - L(L'L)^{-1}L'$$
$$Q_{k-1} = \tfrac{1}{2}(I_{(k-1)^2} + P_{k-1,k-1})$$
$$K = (\Sigma^{-\frac{1}{2}} \otimes \Sigma^{\frac{1}{2}})H_{k-1}\Pi^{\frac{1}{2}}$$
$$L = (\Sigma^{\frac{1}{2}} \otimes \Sigma^{\frac{1}{2}})H_{k-1}\Lambda^{-1}\Pi^{-\frac{1}{2}}$$
$$\Lambda = \text{diag}(\Sigma)$$
$$\Pi = \Lambda^{-1}H'_{k-1}(\Sigma \otimes \Sigma)H_{k-1}\Lambda^{-1}.$$

The matrix Q_{k-1} is a symmetric idempotent matrix (e.g., Wansbeek and Meijer 2000, p. 361), as is M_L, so it follows that S is symmetric and positive semidefinite.

This result generalizes, again after some algebra, to other tests for restrictions of the form $U'\beta = 0$ that do not involve the constant and that still satisfy $\Omega G = 0$ (with G the orthogonal complement of U) as in Sect. 3. (Hence, all mismeasured regressors are included in the test.) So, by and large, the results for known measurement error variance carry over to the case of known reliability, but with some additional restrictions.

6 Estimated reliability

Often, we may not strictly "know" the reliability (or measurement error variance), but we can consistently estimate it. Using the resulting estimate as if it is the known reliability gives consistent estimators of the parameters of interest. However, treating the estimate as the true value leads to an underestimate of the standard errors of the estimators of the coefficients of interest. The estimator of interest is a two-step estimator and the default second-step standard errors do not take the stochastic uncertainty of the first-step estimators into account.

One way to correct this would be to stack the moment conditions of the estimators of the model of interest as discussed in this paper and the moment conditions of the estimator of the measurement error variance (or reliability), using similar techniques as, for example, in Meijer and Wansbeek (2007). As discussed in that paper, if the first-step estimator is overidentified, the generalized method of moments (GMM) estimator from stacking the moment conditions differs slightly from the two-step estimator. This may not be a "problem" at all, as the joint estimator is asymptotically at least as efficient, but it may be computationally or interpretationally more complicated, or less robust to misspecification. To obtain the two-step estimator, the first-step moment conditions have to be replaced by a set of asymptotically equivalent moment conditions that just-identify the estimators, leading to a two-step MM estimator.[5]

In some cases, the measurement error variance (or reliability) is estimated from a different sample. In that case, correct standard errors can be obtained by using a relatively straightforward correction to the default standard errors.[6] Specifically, let the parameters from the first step (reliabilities, measurement error variances, possibly additional auxiliary parameters) be collected in the parameter vector κ. Then typi-

---

[5] A joint estimator can also typically be obtained by specifying both submodels appropriately in a SEM program and estimating the combined model.

[6] Or by using the multiple groups facilities of most SEM programs.

cally $\sqrt{m}(\hat{\kappa} - \kappa)$ (where m is the first-step sample size) is asymptotically normally distributed with mean zero and variance matrix V_κ, say, and the first step estimation produces a consistent estimator $\hat{V}_\kappa$. The second step moment conditions are $\bar{h}(\beta; \hat{\kappa}) = 0$ and treating $\hat{\kappa}$ as if it were the known κ, we obtain the asymptotic variance matrix $\hat{V}_\beta$, say, which is of the form $\hat{G}_\beta^{-1} \hat{V}_h (\hat{G}_\beta')^{-1}$, where $\hat{G}_\beta = \partial\bar{h}/\partial\beta'$ evaluated in $(\hat{\beta}; \hat{\kappa})$, and $\hat{V}_h$ is a consistent estimator of $\mathbb{E}(h_i h_i')$. The corrected variance matrix is obtained by writing

$$0 = \sqrt{n}\,\bar{h}(\hat{\beta}; \hat{\kappa}) = \sqrt{n}\,\bar{h}(\beta; \kappa) + \hat{G}_\beta\big[\sqrt{n}(\hat{\beta} - \beta)\big] + \frac{\sqrt{n}}{\sqrt{m}}\hat{G}_\kappa\big[\sqrt{m}(\hat{\kappa} - \kappa)\big] + o_p(1),$$

with n the second-step sample size, and $\hat{G}_\kappa = \partial\bar{h}/\partial\kappa'$ evaluated in $(\hat{\beta}; \hat{\kappa})$, and using the independence of $\sqrt{n}\,\bar{h}(\beta; \kappa)$ and $\sqrt{m}(\hat{\kappa} - \kappa)$, leading to

$$\hat{V}_{\beta,\mathrm{corr}} = \hat{V}_\beta + \frac{n}{m}\hat{G}_\beta^{-1}\hat{G}_\kappa\hat{V}_\kappa\hat{G}_\kappa'(\hat{G}_\beta^{-1})'.$$

See, for example, Inoue and Solon (2010) for a similar approach in the case of two-sample instrumental variables estimators, and Wooldridge (2002, p. 356) for an analogous approach for two-step M estimators. It is also possible to arrive at this starting from the formulas in Fuller (1987, chap. 3), but this is more involved.

We apply this general theory to the specific case of a single regressor (the first one) with measurement error. First, assume that the measurement error is estimated from an independent sample of size m to be $\hat{\lambda}$, with variance $\hat{v}_\lambda$, so $\Omega = \lambda e_1 e_1'$. Since then $\partial\bar{h}/\partial\lambda = \beta_1 e_1$, the adaptation of the expression given in (9) is

$$\widehat{\mathrm{avar}}(\hat{\beta}) = (\hat{A} - \hat{\Omega})^{-1}\Big(\hat{\mathbb{E}}(h_i h_i') + \frac{n}{m}\hat{v}_\lambda\hat{\beta}_1^2 e_1 e_1'\Big)(\hat{A} - \hat{\Omega})^{-1}.$$

Second, assume that the reliability is estimated to be $\hat{\rho}_1$, with variance $\hat{v}_{\rho_1}$. Then (19) becomes

$$W_i = (1 - \rho_1)(x_{i1} - \mu_1)^2 e_1 e_1',$$

so $\partial\bar{h}_1/\partial\rho_1 = -\beta_1 n^{-1} \sum_i (x_{i1} - \mu_1)^2 e_1$ and the adaptation of (24) is

$$\widehat{\mathrm{avar}}(\hat{\beta}) = (\hat{A} - \hat{\Omega})^{-1}\Big(\hat{\mathbb{E}}(h_{1i} h_{1i}') + \frac{n}{m}\hat{v}_{\rho_1}\hat{\beta}_1^2 (s_{x_1}^2)^2 e_1 e_1'\Big)(\hat{A} - \hat{\Omega})^{-1},$$

with $s_{x_1}^2 = \sum_i (x_{i1} - \bar{x}_1)^2/n$.

7 Extension to panel data

So far we have considered the case of a single cross section. We now consider the case of a panel data model, where measurement error issues are equally relevant, see, for example, Baltagi (2005, sec. 10.1). As documented by Meijer et al. (2015, 2017),

panel data (with independent cross-sectional units) imply additional opportunities for identifying and estimating measurement error models. We now investigate to what extent the analysis for the cross-sectional case we studied so far still essentially holds in the panel data context.

The direct generalization of the cross-sectional model to the panel data case with time dimension T is the following model,

$$y_{it} = \xi'_{it}\beta + \varepsilon_{it}$$
$$x_{it} = \xi_{it} + v_{it},$$

where $t = 1, \ldots, T$ denotes the time index, and for simplicity we assume a balanced panel. We leave the covariance structure over time of ε_{it} unrestricted. Let $\Omega_t = \mathbb{E}(v_{it}v'_{it})$ and $\Omega = \sum_{t=1}^{T}\Omega_t$. Extending (4) to the panel case, let

$$h_{it} = x_{it}y_{it} - (x_{it}x'_{it} - \Omega_t)\beta \quad \text{and} \quad h_i = \sum_{t=1}^{T} h_{it} = X'_i y_i - (X'_i X_i - \Omega)\beta, \quad (27)$$

where y_i is the vector that stacks the $y_{it}, t = 1, \ldots, T$, and X_i is the $T \times k$ matrix whose tth row is x'_{it}. If ε_{it} and ξ_{it} are uncorrelated (contemporaneous exogeneity), $\mathbb{E}(h_{it}) = 0$ and thus $\mathbb{E}(h_i) = 0$, so this is a valid moment condition and, with $X = (X'_1, \ldots, X'_n)'$ and $y = (y'_1, \ldots, y'_n)'$,

$$\hat{\beta} = (X'X - n\Omega)^{-1}X'y \quad (28)$$

is the method-of-moments estimator of β from (27). It is basically the pooled OLS estimator corrected for measurement error by using Ω, supposedly known. The usual robust estimator of its variance takes care of correlation over time and hence covers the random effects case, with the random individual effects implicitly included in ε_i.

With individual fixed effects, that is, $\varepsilon_{it} = \alpha_i + r_{it}$ with α_i potentially correlated with ξ_{it}, they need to be eliminated, which is typically done by the within transformation or first differencing (e.g., Baltagi 2005, pp. 13, 136). After such a transformation, the resulting data contain combinations of measurement errors from multiple time points: $v_{it} - \sum_{s=1}^{T} v_{is}/T$ in the case of the within transformation, and $v_{it} - v_{i,t-1}$ in the case of first differencing. The variances of these terms depend on the Ω_t in more complicated ways, and if the measurement errors are serially correlated, they also depend on the covariances between the measurement errors across time. Hence, in order to correct for measurement error, information on the measurement error structure over time has to be known in addition to knowledge of Ω. The simplest (and strongest) assumption would be that $\Omega_t = \bar{\Omega}$ does not vary over time and that the measurement errors are serially uncorrelated. Then $\text{var}(v_{it} - v_{i,t-1}) = 2\bar{\Omega}$ and $\text{var}(v_{it} - \sum_{s=1}^{T} v_{is}/T) = \bar{\Omega}(T-1)/T$, which leads to straightforward adaptations of (27) for the transformed data.

In the case of knowledge of the reliability, a leading case is also when the reliability is constant over time. First, consider the case without fixed effects. Let, as in the

cross-sectional case, all Ω_t be diagonal with

$$\Omega_t = \mathrm{diag}[(1 - \rho_j)(A_{jjt} - \mu_{jt}^2)], \tag{29}$$

with A_{jjt} the jth diagonal element of $A_t = \mathbb{E}(x_{it}x_{it}')$. Furthermore, let

$$W_{it} = \mathrm{diag}[(1 - \rho_j)(x_{ijt} - \mu_{jt})^2], \tag{30}$$

where μ_{jt} is the jth element of $\mu_t = \mathbb{E}(\xi_{it})$. Consequently, $\mathbb{E}(W_{it}) = \Omega_t$. Let $W_i = \sum_t W_{it}$ and let M be the $T \times k$ matrix with tth row equal to μ_t'. The moment condition for the cross-sectional case as given in (20) generalizes to

$$h_i = \begin{pmatrix} X_i'y_i - (X_i'X_i - W_i)\beta \\ \mathrm{vec}(X_i - M) \end{pmatrix}.$$

So, also in the case of known reliability, the analysis for a single cross-section carries over to the panel data case in a straightforward way.

Now, consider the case with fixed effects and assume the measurement errors are serially uncorrelated. Let a tilde denote the within transformation. We then obtain

$$\mathrm{var}(\tilde{v}_{it}) = \mathrm{var}\left(v_{it} - \tfrac{1}{T}\sum_{s=1}^{T} v_{is}\right) = \frac{(T-1)^2}{T^2}\Omega_t + \frac{1}{T^2}\sum_{s=1,s\neq t}^{T}\Omega_s = \Omega_t^*,$$

say, with Ω_t as in (29). In this case, let

$$h_{it} = \tilde{x}_{it}\tilde{y}_{it} - (\tilde{x}_{it}\tilde{x}_{it}' - W_{it}^*)\beta$$

with

$$W_{it}^* = \frac{(T-1)^2}{T^2}W_{it} + \frac{1}{T^2}\sum_{s=1,s\neq t}^{T} W_{is}$$

and W_{it} as in (30). Then h_{it} is a valid moment for this case. An analogous expression can be obtained in the case of first differencing.

In this section, we have only scratched the surface. The presence of panel data allows for a large number of potential assumptions about how the measurement errors evolve over time and how this can be used to estimate the coefficients consistently. Moreover, we have not discussed dynamic panel data, in which the lagged dependent variable is a regressor (e.g., Baltagi 2005, chap. 8), which is associated with a host of econometric issues that we have not discussed here. However, the cases discussed here serve as illustrations of how one can derive consistent estimators for such cases.

8 Empirical example

To illustrate the above, we estimate a hedonic price function that specifies price of wine as a function of its attributes or characteristics, see Oczkowski and Doucouliagos (2015) for a review and meta-analysis. In part, the literature recognizes that wine quality influences prices, and most studies employ a subjective quality score from a wine guide as an indicator of quality. Only a few studies, however, have recognized the consequent measurement error associated with expert quality scores only reflecting some underlying notion of latent wine quality. Oczkowski (2001) employs an instrumental variable estimator using multiple expert scores to consistently estimate the relation between price and latent quality. In contrast, Lecocq and Visser (2006) do adjust their price-quality estimates for the attenuation bias associated with expert scores; however, their adjustment formula ignores the impact of other (nonquality score) regressors on the attenuation bias, and no adjustments are made for standard errors.

Our example focuses on Australian premium wines available during 2015 and an average quality score from four expert tasters, Geddes (2015), Oliver (2015), Hooke (2015), and Halliday (2015). We estimate the equation

$$\ln(\text{Price}_i) = \beta_0 + \gamma Q_i + \beta_1 \text{Vintage}_i + \beta_2' \text{Region}_i + \beta_3' \text{Variety}_i + u_i, \quad (31)$$

where Price_i is the recommended retail price in 2015 measured in Australian dollars (Halliday 2015); Q_i is an average quality score measured out of 100; Vintage_i is the year in which the grapes were harvested; Region_i is a series of dummy variables depicting the region from where the grapes were sourced; Variety_i is a series of dummies representing the variety, blend or style of wine. Descriptive summary statistics of the data are provided in Table 1.

The quality score is an average of four expert scores, where the scores are standardized using a nonparametric distribution transformation to reflect the Halliday (2015) rating, see Cardebat and Paroissien (2015). Effectively, the other three scores are transformed to have the same quantiles as Halliday (2015). The standardized scores have similar means across the average and individual scores. However, as expected, the standard deviation for the average score (1.62) is smaller than that of the individual expert scores (2.20). The estimated standardized Cronbach's alpha reliability coefficient for the four experts is $\alpha = 0.728$.[7] The quality variable captures both the preferences of consumers for higher quality wines and the increased costs of producing better quality wines.

The vintage variable captures the preferences held by some consumers for older wines and the increased costs of producing wines which are long-lived and the costs of storing wines. In the sample, approximately 90% of wines come from the 2012, 2013,

---

[7] Cronbach's (1951) coefficient alpha is a measure of internal consistency of a scale that is a simple sum (or average) of a number of items. It is very easy to compute and if the items can be viewed as repeated measures in a simple measurement error model, it estimates the reliability of the scale. For these reasons, it has been routinely reported in psychological and educational studies that utilize such scales. More generally, it underestimates reliability and better measures are available. However, in our empirical example, the difference between its value and the estimated reliability derived from a factor analysis model is negligible, so this concern is inconsequential. See Sijtsma (2009) for a modern (and critical) review of Cronbach's alpha and its alternatives.

Table 1 Descriptive statistics

Variable	Mean	SD	Min	Max
Price	49.35	43.14	14.99	350
ln(Price)	3.689	0.588	2.707	5.858
Quality score	94.41	1.635	89.5	98.0
Vintage	2012.6	1.354	2005	2014

$n = 258$ with 11 varieties and 17 regions. Varieties and regions only included if they contain at least 10 wines. One wine with an unusually high price of $785 was omitted

and 2014 vintages, but some wines extend back to 2005. The region variables capture both the preferences of consumers and the costs of producing wines in different cool and warm climate regions. The main regions in the sample are Margret River (12.0%), Clare Valley (9.7%), and McLaren Vale (9.3%). The variety variable mainly captures consumer preferences. The main varieties in the sample are Shiraz (24.8%), Riesling (13.6%), and Chardonnay (13.2%).

We estimate (31) using the estimator (21), allowing Q_i to suffer from measurement error, but the other regressors not, for a range of reliability values for Q_i from 1.0 (uncorrected least squares) and reducing by 0.1 increments, also including the estimated reliability of 0.728 for the data set. There is a lower limit for the proposed reliability, because the implied covariance matrix of $(y_i, \xi_i')'$ needs to be positive (semi)definite. Effectively, reliabilities below this limit cannot add any additional explanatory power to the model.[8] This lower limit is the R^2 from the regression of the quality score on the other regressors in (31) and ln(Price$_i$). In our case, this is $R^2 = 0.546$, and therefore, we only present estimates for reliabilities of 0.60 and higher.

The estimates of (31) for various reliabilities are reported in Table 2. The standard attenuation bias adjustment is evident with the quality score point estimate ($\hat{\gamma}$) monotonically rising from 0.211 for no correction to 0.413 for a reliability of 0.60. For the estimated alpha of 0.728, the quality score estimate is 0.316 which constitutes an *additional* 10.5% increase in prices per quality point compared to the uncorrected estimate. This is very important economically, as correcting for measurement error on average leads to an *additional* $5.18 (in $AUD) per quality score point.

For the estimated alpha, the corrected quality coefficient estimate is approximately 50% higher than the OLS counterpart. This difference is similar to Oczkowski's (2001) finding for the difference between 2SLS and OLS estimates for latent variable models of wine reputation on price, using Australian wines assessed in 1999 and 2000 ($n = 276$). Lecocq and Visser (2006) identified the difference between measurement-error corrected and uncorrected estimates of 24% for a 1992 Bordeaux ($n = 519$) sample, 85% for a 1993 Burgundy ($n = 613$) sample and 73% for a 2001 Bordeaux ($n = 255$) sample. In general, the estimates appear to differ across time and samples, but they do point to substantial differences between measurement-error corrected and uncorrected quality-price estimates for wine. The robust standard errors for $\hat{\gamma}$ based on (24) and

---

[8] For this reason, Stata's **eivreg** program refuses to estimate the model in such cases.

Table 2 Hedonic price estimates: different reliability estimates

Reliability	R^2	Coeff.	t-value	
	Robust	Normal		
1.0	0.648	0.211	13.06	14.36
0.90	0.687	0.240	13.12	14.28
0.80	0.739	0.279	13.03	13.94
0.728	0.788	0.316	12.78	13.42
0.70	0.811	0.333	12.62	13.13
0.60	0.918	0.413	11.56	11.59

$n = 258$; vintage, regions, and varieties suppressed

the standard errors based on the normal distribution (25) lead to mostly decreasing t-ratios, though not completely monotonically for the robust ones.

As a robustness check, we have investigated some alternative specifications for the model (31): (a) using vintage dummies instead of including vintage linearly; (b) dropping the region and variety dummies; (c) both (a) and (b); (d) including the quality measure as the *only* regressor in the model. The results for model (a) are very similar to the results in the table. For models (b) and (c), the coefficient estimates increase from about 0.20 to about 0.34, so they are a bit smaller than in the table. For model (d), they increase from 0.216 to 0.361. The R^2s follow expected patterns: They are slightly higher when vintage is included as a set of dummies than when vintage is linear and substantially lower when the variety and region dummies are dropped. For model (d), R^2 increases from 0.364 to 0.606. Most interesting are the results for the t-values. In models (a) and (c), they decrease monotonically with decreasing reliability, both when robust and when normality-based standard errors are used. In model (b), the t-values are almost constant, but slightly *increasing* (from 12.21 to 12.26) with robust standard errors and slightly decreasing (from 13.25 to 13.20) as usual with normality-based standard errors. For model (d), the t-values are constant (11.65 for robust and 12.15 for normal).

In Fig. 1, we return to our reference model, but consider the situation when we know either the measurement error variance (left) or the reliability (right) and illustrate graphically the relation between the assumed measurement error variance or reliability and the estimation results. The t-values shown here are based on the robust variance estimates (9) and (24). This shows again that with increasing measurement error variance and decreasing reliability, the coefficient and the R^2 increase, but the t-value decreases, although the latter not monotonically for an assumed reliability close to 1. The t-value graphs using the estimated variances (14) and (25) based on the normality assumption (not shown) are qualitatively similar, but the t-values are a bit higher—as in Table 2—and their relation with the assumed reliability is monotonic, confirming the theoretical analysis. However, regardless of the specific assumptions made, there is no question that the coefficient of the quality rating remains highly statistically significant.

Up until now, we have assumed ignorance about the reliability of Q_i as a proxy of the true quality, Q_i^*, say. However, we can say more when we are willing to assume that
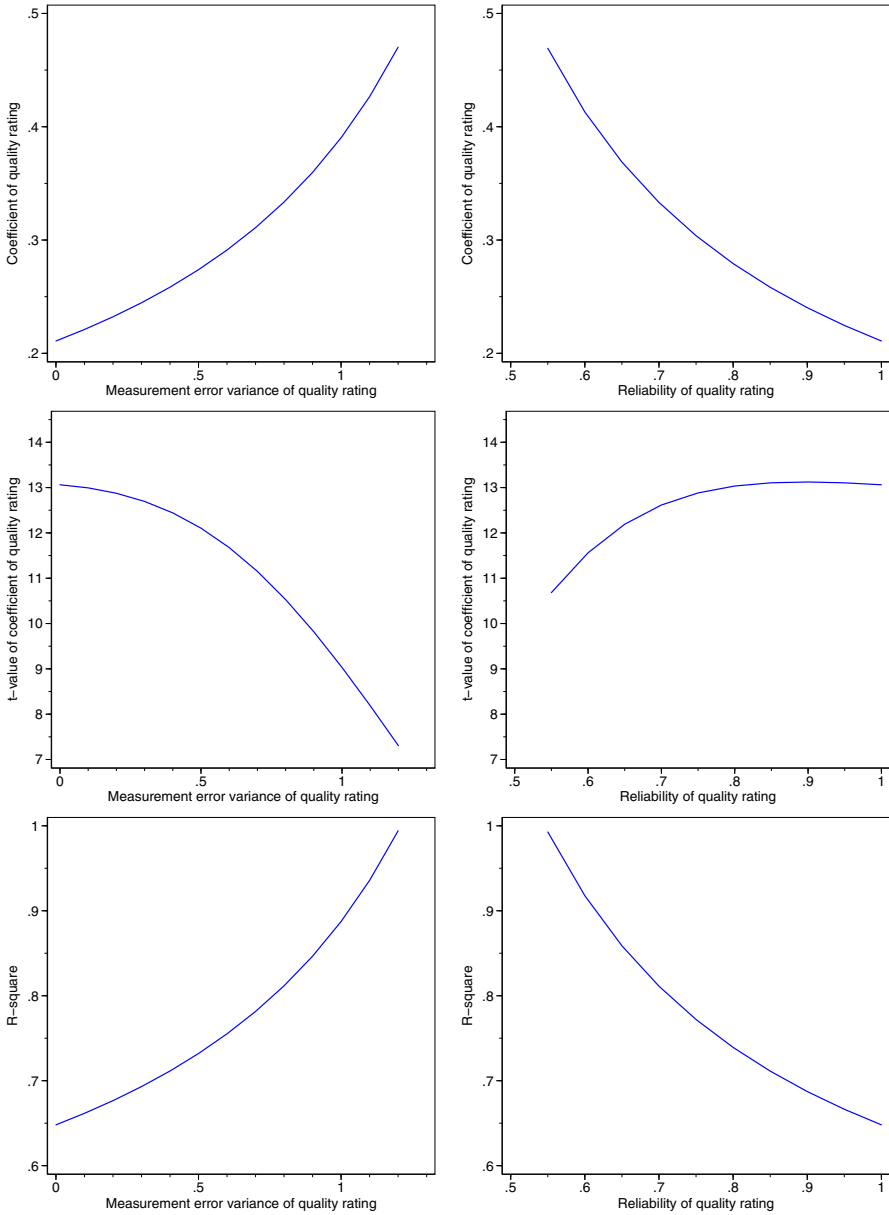
Fig. 1 Coefficient and *t*-value of the quality rating variable, and the resulting (corrected) R^2, as a function of measurement error variance (left) and reliability (right) of the quality rating, using robust standard errors

the scores Q_{i1}, Q_{i2}, Q_{i3}, and Q_{i4} given by the four expert tasters, after demeaning, satisfy a one-factor model,

$$Q_{im} = b_m Q_i^* + w_{im},$$

$m = 1, 2, 3, 4$, where the b_m are the factor loadings and the w_{im} are the error terms, with variances ω_m^2 and covariances zero. By way of normalization, we set the variance of Q_i^* equal to one. The case of no measurement error corresponds to $\omega_m^2 = 0$ for all m; the experts agree.[9] The quality variable Q_i was constructed as the average over the expert scores. So, with bars denoting the average over m,

$$Q_i = \bar{Q}_{i\cdot} = \bar{b} Q_i^* + \bar{w}_i.$$

The reliability of Q_i as a proxy for Q_i^* can now be expressed as

$$\rho = \frac{\bar{b}^2}{\bar{b}^2 + \frac{1}{4}\overline{\omega^2}}.$$

We estimated the b_m and ω_m^2 with Stata's **sem** module using the original scores, and find $\hat{\rho} = 0.7286$, which is almost identical to the Cronbach's alpha value of 0.728 mentioned earlier. With this reliability, the estimate of the quality rating coefficient is 0.316, while the implied R^2 of the regression is 0.788.

9 Discussion

It is well known that measurement error is pervasive in economic data and that it tends to bias estimators that do not correct for measurement error in the explanatory variables. We rigorously analyzed the linear regression model with measurement error, where either the variance matrix of the measurement errors is known or the reliabilities of the regressors are known. Although these cases have been discussed in the literature, we bring the results together concisely within the framework of GMM theory. We also discussed some special cases, in particular normality of the measurement errors and measurement error in only a single regressor. For these cases, the expressions simplify greatly. Furthermore, we derived expressions for the related case where measurement error variance or reliability is not known, but consistently estimated, either from the same sample or an independent sample.

Or main focus is on the effects of measurement errors on the t-statistics and hence statistical significance. We compare the t-statistic of the consistent estimator with the t-statistic of the (inconsistent) OLS estimator and the t-statistic of the (infeasible) estimator if there was no measurement error and show that they are ordered with the

---

[9] Because of the way the ratings were standardized, this would also imply that the b_m values are equal in this case. In other contexts, with a latent variable without a natural scale, we would allow the observed variables to have different scales and hence the b_m to vary, while still claiming absence of measurement error.

t-statistic of the consistent estimator being closest to zero and the t-statistic of the (infeasible) estimator being largest in absolute value. This holds for both the case with known measurement error variance and the case with known reliability. We also greatly generalized our earlier finding (Meijer and Wansbeek 2000) that the t-value decreases with the assumed measurement error variance and showed that the t-value also decreases with decreasing assumed reliability of the regressor. These results use normality of the measurement errors, as general results for robust standard errors cannot be obtained. Our empirical results suggest that the results largely carry over to robust inference, but there may be some minor departures from monotonicity.

We have also developed extensions of these estimators to panel data, which comes with a number of additional issues and opportunities. In particular, we now have to consider whether the measurement errors are serially correlated, whether they are stationary, whether there are random or fixed effects in the model of interest, and whether the model is static or dynamic. We have derived estimators for some illustrative cases in static panel data models with and without fixed effects, which also serve as guides to how one could derive estimators in a specific panel data application with more general assumptions.

We illustrated the results by estimating a hedonic regression for the price of Australian wines. We showed the sensitivity of the coefficient of the quality indicator to the assumed reliability of this indicator: This coefficient ranges from 0.2 without measurement error (reliability = 1) to 0.4 when reliability is 0.6. This also has consequences for the implied R^2 of the regression (which goes up with decreased reliability) and the t-statistic of the error-ridden regressor (which goes down with decreased reliability). However, in this particular regression, the coefficient of quality always remains statistically significant.

In the empirical study, the quality indicator was obtained as the average of four independent ratings of the quality of the same wine. By assuming a linear factor analysis model for these four ratings, we were able to estimate the reliability of the quality indicator, which is about 0.73. Taking this as the known reliability, point estimates and other statistics follow from our formulas.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflicts of interest.

Human and animal rights This article does not contain any studies with human participants or animals performed by any of the authors.

A Auxiliary lemma

Lemma 1 *Let U be the $k \times (k-1)$ matrix consisting of the first $k-1$ columns of I_k. If D is a $(k-1) \times (k-1)$ nonsingular matrix and m is a k-vector such that $m \neq UU'm$, then $X = UDU' + mm'$ is nonsingular and $U'X^{-1}U = D^{-1}$.*

Proof Define $V = (U, m)$, which is nonsingular because $m \neq UU'm$. Then $(U, e_k) = I_k = V^{-1}V = (V^{-1}U, V^{-1}m)$, where e_k is the kth column of I_k, so $V^{-1}U = U$. Defining $H = \text{diag}(D, 1)$, which is nonsingular, we have $X = VHV'$ and thus $X^{-1} = (V')^{-1}H^{-1}V^{-1}$ and $U'X^{-1}U = (V^{-1}U)'H^{-1}(V^{-1}U) = U'\text{diag}(D^{-1}, 1)U = D^{-1}$. $\qquad\square$

References

Baltagi BH (2005) Econometric analysis of panel data, 3rd edn. Wiley, New York

Baltagi BH (2011) Econometrics, 5th edn. Springer, Berlin

Bloch FE (1978) Measurement error and statistical significance of an independent variable. Am Stat 32:26–27

Bound J, Brown C, Mathiowetz N (2001) Measurement error in survey data. In: Heckman JJ, Leamer E (eds) Handbook of econometrics, vol 5. North-Holland, Amsterdam, pp 3705–3843

Buonaccorsi J (2010) Measurement error, models, methods, and applications. Chapman & Hall, Boca Raton

Cardebat J-M, Paroissien E (2015) Standardizing expert wine scores: an application for Bordeaux en primeur. J Wine Econ 10:329–348

Cronbach LJ (1951) Coefficient alpha and the internal structure of tests. Psychometrika 16:297–334

De Haan J, Leertouwer E, Meijer E, Wansbeek TJ (2003) Measuring central bank independence: a latent variables approach. Scot J Polit Econ 50:326–340

Erickson T, Whited TM (2002) Two-step GMM estimation of the errors-in-variables model using high-order moments. Econ Theory 18:776–799

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

Fuller WA (1987) Measurement error models. Wiley, New York

Geary RC (1942) Inherent relations between random variables. Proc R Ir Acad A 47:63–67

Geddes R (2015) Australian wine vintages 2016, 33rd edn. Geddes A Drink Publication, Sydney

Halliday J (2015) Australian wine companion 2016. Hardie Grant Books, Richmond

Hooke H (2015) The wine guide 2016. Bauer Media Books, Sydney. http://huonhooke.com

Inoue A, Solon G (2010) Two-sample instrumental variables estimators. Rev Econ Stat 92:557–561

Lecocq S, Visser M (2006) What determines wine prices: objective vs. sensory characteristics. J Wine Econ 1:42–56

Lockwood JR, McCaffrey DF (2020) Recommendations about estimating errors-in-variables models in stata. Stata J 20:116–130

Meijer E, Wansbeek TJ (2000) Measurement error in a single regressor. Econ Lett 69:277–284

Meijer E, Wansbeek TJ (2007) The sample selection model from a method of moments perspective. Econ Rev 26:25–51

Meijer E, Spierdijk L, Wansbeek TJ (2015) Measurement error in panel data. In: Baltagi BH (ed) The Oxford handbook of panel data. Oxford University Press, Oxford, pp 325–362

Meijer E, Spierdijk L, Wansbeek TJ (2017) Consistent estimation of linear panel data models with measurement error. J Econ 200:169–180

Merckens A, Wansbeek TJ (1989) Formula manipulation in statistics on the computer: evaluating the expectation of higher-degree functions of normally distributed matrices. Comput Stat Data Anal 8:189–200

Oczkowski E (2001) Hedonic wine price functions and measurement error. Ecoc Rec 77:374–382

Oczkowski E, Doucouliagos H (2015) Wine prices and quality ratings: a meta-regression analysis. Am J Agric Econ 91:103–121

Oliver J (2015) The Australian wine annual 2016, 19th edn. Jeremy Oliver, Canterbury

Schennach SM (2016) Recent advances in the measurement error literature. Annu Rev Econ 8:341–377

Sijtsma K (2009) On the use, the misuse, and the very limited usefulness of Cronbach's alpha. Psychometrika 74:107–120

StataCorp (2019a) Stata base reference manual: release 16. StataCorp, College Station, TX

StataCorp (2019b) Stata structural equation modeling reference manual: release 16. StataCorp, College Station, TX

Wansbeek TJ, Meijer E (2000) Measurement error and latent variables in econometrics. North-Holland, Amsterdam

Wooldridge JM (2002) Econometric analysis of cross section and panel data. MIT Press, Cambridge

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.