Cost Function Estimation with Proportional Errors in Variables

Richard E. Just and Rulon D. Pope®

Department of Agricultural and Resource Economics, University of Maryland, College Park Maryland, USA and Department of Economics, Brigham Young University, Provo, Utah, USA

ABSTRACT

A model with proportional errors in variables arising naturally in microeconomics is considered. Unlike the classical additive errors case, all OLS parameter estimates exhibit attenuation bias that does not depend on the limiting distribution of the data. The distribution of OLS estimators is developed. With no intercept, a simple correction of OLS based on mean predictions is identified that is consistent and asymptotically normal. With an intercept, a readily available additional moment based on sample data identifies the parameters. In neither case are additional restrictions or use of extra-sample data as instruments required as for common errors-in-variables methods.

Key Words: Errors in variables, Proportional errors, Estimation

JEL Classifications: C13, C20, C80

1. INTRODUCTION

It is well known that errors in variables (EIV) cause bias and inconsistency in regression estimates. Special stochastic structures of errors in variables have yielded useful insights into bias and inconsistency (e.g., Amemiya, 1990; Cragg, 1997; Garber and Klepper, 1980; Hausman, 2001; Griliches and Ringstad, 1970; Hausman, Newey, Ichimura, and Powell, 1991; Riersol, 1950, Theil, 1961). However, in the classical EIV model, $y_t = X_t^* \beta + \varepsilon_t$, $X_t = X_t^* + \delta_t$, $E(\varepsilon_t) = 0$, $t = 1,\ldots,T$, where the regressors in X_t^* are observed

® Richard E. Just, Department of Agricultural and Resource Economics, University of Maryland, College Park Maryland, USA (e-mail: rjust@arec.umd.edu)
Rulon D. Pope, Department of Economics, Brigham Young University, Provo, Utah, USA (e-mail: rulon_pope@byu.edu)

only with the errors in X_t, the direction of bias in least squares estimators has generally been ambiguous (Nelson,1995) and models are not identified without adding restrictions or extra-sample data as instruments.

This paper considers a special case where the right hand side is contaminated with a single measurement error that affects all right hand side terms or all terms except the constant term. Because of the examples that arise naturally in microeconomics, we consider proportional errors in variables (PEIV) where $y_t = \alpha + X_t^* \beta + \varepsilon_t,\ X_t = X_t^* \delta_t,\ E(\varepsilon_t) = 0,\ E(\delta_t) = 1,$ i.e., where X_t is a vector of at least two regressors each affected by the same scalar error, δ_t, multiplicatively. We suggest that such cases are likely common in microeconomic models that impose some form of homotheticity or homogeneity.

To illustrate, modern applied consumer demand theory generally flows via Roy's identity from an indirect utility function to consumer demands. The indirect utility function $V(p,m) = \overline{V}(p/m)$ where $\overline{V}$ is a function of a price vector p divided by scalar income m follows from homogeneity (Blundell, 1988). Implied consumer demands are of the form $x = f(p/m)$. Thus, errors of measurement in income cause common mutiplicative errors in all normalized prices. Because of well-known problems in measuring income, e.g., permanent versus transitory income, we suggest that errors in measurement of income are far more likely than errors in measurement of prices.

Similar problems occur in standard microeconomic models of production. A common approach in applied studies is to start with the unit cost function or some generalization of it under homotheticity (Berndt and Wood, 1975). Cost functions are thus of the form $c(p,w) = \varphi(q)h(p)$ where p now represents a vector of factor prices, q is a scalar output, and c is cost. Mismeasurement of output q thus causes common multiplicative errors of measurement in all price effects. Since factor demands are the gradient of c with respect to p via Shephard's Lemma, factor demands and the substitution effects are thus mismeasured due to the common impact of errors in q. Because many forms of production are affected by unforeseen circumstances, so that ex post production differs from ex ante planned production, we suggest that errors in measurement of the relevant planned output are far more likely than errors in measurement of prices.

Another common approach in applied production studies is the profit function approach in which the standard property of linear homogeneity in prices yields $\pi(w,p) = \overline{\pi}(w/p)$ where w is a vector of prices for inputs and p is now a scalar price of a single output. Mismeasurement of the output price thus causes common multiplicative errors in measurement of all normalized input prices. Because input prices are typically well-known at the commencement of any time-consuming production process but the resulting output price after production cannot always be well anticipated, we suggest that errors of measurement in anticipated output prices by using data on later actual output prices are far more likely than for the input prices. In each of these cases, a plausible model involves common multiplicative mismeasurement errors among regressors.

Interestingly, concern over common errors among regressors as a result of deflating all variables by a common but erroneous deflator arose relatively early in the errors-in-variables literature in economics. While these studies are salient to the type of question we consider, they generally bear little or no resemblance to modern applied microeconomic practice as discussed above. Kuh and Meyer (1955) and Briggs (1962) examined conditions under which correlations were biased and non-deflation was therefore preferred to deflation. Subsequently, Casson (1973) provided an approach for consistent estimation when the dependent variable as well as all the regressors are subject to the same proportional measurement error. In his case, multiplying the regression equation by the deflator obtains a model with measurement error only in the constant term. However, most modern microeconomic models that are popular for empirical purposes represent economic behavior where the right hand side structure does not generate deflated dependent variables.

To make clear the empirical relevance of the specific PEIV structure of the linear model in this paper in production and consumption studies, consider the widespread application of the following models.

1.1 Cost Function Estimation

Homothetic technology, stochastic production, and risk neutrality yields a cost function of the form $c = \theta(\bar{q})\varphi(r)$ where q is planned output and r is an input price vector (Shephard, 1970). Suppose $\varphi(r)$ is linear in parameters and observed output q varies about planned unobserved output $\bar{q}$ stochastically as with unanticipated weather effects on agricultural production. If production has constant returns to scale, $\theta(\bar{q}) = \bar{q}$, and $\varphi(r)$ takes on a generalized Leontief form (Diewert, 1971), then the estimated equations are typically the conditional input demands (x),

$$x_i = \bar{q}\left[b_{ii} + \sum_{j\neq i} b_{ij}(r_j/r_i)^{1/2}\right] = b_{ii}\bar{q} + \sum_{j\neq i} b_{ij}\bar{q}(r_j/r_i)^{1/2},$$

which yield an ith factor demand the form $y = X^*\beta$ where the jth column of X^* has element $\bar{q}(r_j/r_i)^{1/2}$. If planned output is mismeasured by actual output, e.g., by q where $q = \bar{q}/\delta$, $E(\delta) = 1$, then the PEIV model of this paper applies. The generalized Leontief cost function developed by Diewert and it's associated factor demands have been employed extensively in homothetic and non-homothetic forms (where the latter commonly adds an additional term involving $\bar{q}$). Homothetic examples are numerous: e.g., Izumida, Urushi, and Nakanishi (1999); Fuss (1977), Pope and Just (1996).

1.2 Consumer Demand Estimation

Any consumer demand x of the form $x = \alpha + \sum_j b_j(p_j/m)$ will also have the proportional error structure when income m is mismeasured. Some examples include Heien (1977), Burt and Brewer (1971), and Song and Hallberg (1982). Equally common in applied demand work with incomplete demand systems is to deflate all terms on the right hand side except the constant in a partial demand system or a single demand equation, which in typical linear form yields $x = \alpha + b_0(m/P^*) + \sum_j b_j(p_j/P^*)$ and is also of the form $y = X^*\beta$. Thus, all terms except the constant are proportionally mismeasured if the deflator P^* is measured with error, e.g., by P where $P = P^*/\delta$, $E(\delta) = 1$. These applications can be expanded to include particular forms of utility as in the production example above. For example, homotheticity can be expanded to include quasi-homotheticity in which case demands will be of the form $x = a(p) + b(p)m$ (Gorman 1959; Lewbel, 1987), which is required when aggregate demands are rationalized from consumer theory and aggregate income or per-capita income is used for m.

These models show that the PEIV model arises naturally in competitive demand, supply, and cost function estimation problems when key variables such as output, income, or price indexes are measured imperfectly. Thus, theoretically-grounded empirical practice calls for consistent estimation of PEIV models.

Often, these models have a separable form which, when coupled with homogeneity conditions, implies that no constant term (other than one affected by the random multiplicative factor) should be included in the regression equation. For example, any constant term in $\varphi(\cdot)$ should be multiplied by the measurement error in q in case 1 above. Yet in empirical work, the PEIV problem is generally ignored. This is particularly troublesome in applications where one knows a priori, consistent with expected profit or expected utility maximization, that the ex post realization of random output is not applicable in the cost function (e.g. in agricultural applications; see Pope and Just 1996).

The two most distinguishing features of these models as well as the PEIV model of this paper are multiplicative errors and commonality of the multiplicative error term among regressors. Consideration of multiplicative errors is, of course, not new. Typical empirical practice has often been to assume, as a matter of convenience, that errors in variables are multiplicative when data are logged, and additive when data are not logged, so that standard EIV models as they have been developed in the literature are applicable. We suggest alternatively that the choice of linear versus log linear forms is not always a matter of convenience for serious empirical work as suggested by the early work on ratio regressors cited above. Otherwise, the literature would not continue to use specific forms such as the generalized Leontief cost and expenditure functions, or the CES, Generalized McFadden, or other specific functional forms to specify estimated demands involving additive functions of the levels of variables rather than of their logs. The approach of this paper eliminates the arbitrary choice of a logged-data model for the case of proportional errors invariables because essentially the same convenience as ordinary least squares (OLS) is attained for a model additive in levels. Moreover, while some alternative specific functional forms use logged data, e.g., the Cobb-Douglas form, the approach of this paper is apparently the only one that permits consideration of the PEIV problems highlighted above in a modern second-order flexible form (forms such as the translog are not linear in logged data and thus present other non-standard problems for estimation).

In this paper, absence of a constant term is considered initially not only because this case has theoretical relevance as in case 1 above, but because the model is readily identified using standard assumptions and sample moments without extraneous data (an unusual result in the EIV literature). The magnitude and direction of the bias of least squares is unambiguously obtained. A simple correction to the least squares estimator is derived analytically to achieve consistency, and the asymptotic distribution of the corrected least squares estimator is obtained. Subsequently, the more general case with a constant is considered where an obvious instrument leads to an additional moment condition and consistent estimation. In this case, a simple adjustment to the least squares estimator is consistent although it cannot be derived in closed form. The estimator can be placed within the generalized method of moments (GMM) framework so that the asymptotic properties of the estimator are readily determined. A novel aspect of the GMM approach in this paper is that, while instruments are typically assumed nonstochastic or orthogonal to disturbances, we start with moment conditions using the instruments directly.

2. INCONSISTENCY OF OLS ESTIMATORS

Consider a standard linear regression model with no intercept and at least two regressors, $y_t = X_t^* \beta + \varepsilon_t$, $t = 1,...,T$, where regressor data on each observation contain proportional errors, $X_t = X_t^* \delta_t$, where δ_t is a scalar multiplicative error.

Assumption A. Suppose (i) standard regression assumptions apply: $E(\varepsilon_t) = 0$, $E(\varepsilon_t^2) = \sigma < \infty$, $t = 1,...,T$, $\sum_t X_{ti}^* X_{tj}^* / T \overset{a.s.}{\rightarrow} Q_{ij}$ where $Q = \{Q_{ij}\}$ is positive definite and exists, and X_t^* and ε_t are jointly independent and identically distributed (iid) random draws with bounded fourth moments;[1] and (ii) additional PEIV assumptions apply: $E(\delta_t) = 1$, $E(\delta_t^2) = \omega < \infty$, $t = 1,...,T$, and $X'X$ is invertible except on an inconsequential set of measure zero where δ_t is also jointly iid with X_t^* and ε_t with bounded fourth moments.

[1] The iid assumption can be relaxed to require only independence without much difficulty by applying Markov's version of the law of large numbers (White, 1984, p. 330).

For notational purposes let $y = (y_1,...,y_T)'$, $X^* = (X_1^{*\prime},...,X_T^{*\prime})' = \{X_{ti}^*\}$, $\varepsilon = (\varepsilon_1,\ldots, \varepsilon_T)'$, $X = (X_1',...,X_T')'$, and $\Delta = diag(\delta_1,\ldots,\delta_T)$. Then the model $y = X^*\beta + \varepsilon$ is inconsistently estimated by the OLS estimator using data $X = \Delta X^*$,

$$\hat{\beta} = (X'X)^{-1}X'y = (X'X)^{-1}X'X^*\beta + (X'X)^{-1}X'\varepsilon, \tag{1}$$

because

$$(X'X/T)^{-1} = (X^{*\prime}\Delta^2 X^*/T)^{-1} \overset{a.s.}{\rightarrow} Q^{-1}/\omega, \tag{2}$$

$$X'\varepsilon/T = X^{*\prime}\Delta\varepsilon/T \overset{a.s.}{\rightarrow} 0, \tag{3}$$

$$X'X^*/T = X^{*\prime}\Delta X^*/T \overset{a.s.}{\rightarrow} Q. \tag{4}$$

That is, the respective elements of $X^{*\prime}\Delta^2 X^*/T$, $X^{*\prime}\Delta\varepsilon/T$, and $X^{*\prime}\Delta X^*/T$ are $\sum_t X_{ti}^* X_{tj}^* \delta_t^2/T$, $\sum_t X_{ti}^*\delta_t\varepsilon_t/T$, and $\sum_t X_{ti}^* X_{tj}^*\delta_t/T$, which converge almost surely to their expectations ωQ_{ij}, 0, and Q_{ij}, respectively, by the Kolmogorov law of large numbers under Assumption A. Substituting (2)-(4) into (1) proves:

Proposition 2.1. Under Assumption A, $\hat{\beta} \overset{a.s.}{\rightarrow} \beta/\omega$.

Note that $\omega > 1$ must hold for stochastic δ_t because $\omega = E(\delta_t^2) = Var(\delta_t) + 1$. Thus, from the result in Proposition 2.1, proportional errors in variables bias regression coefficients proportionally and cause absolute underestimation of every element of β (attenuation bias). This result suggests that many elasticities for microeconomic models in the literature may be underestimated where the conditions discussed in the introduction apply.

3. INCONSISTENCY OF ESTIMATED STANDARD ERRORS

If the bias in estimates, $\hat{\beta}$, are proportional to β, then the possibility exists that tests of zero regression coefficients under OLS are appropriate. To examine this issue, consider errors in estimating the covariance matrix. Using true but unobservable data, the least squares regression estimator is $\beta^* = (X^{*\prime}X^*)^{-1}X^{*\prime}y$, which with spherical disturbances has covariance matrix $\Sigma_\beta^* = \sigma(X^{*\prime}X^*)^{-1}$ for which $T\Sigma_\beta^* \overset{a.s.}{\rightarrow} \sigma Q^{-1}$. Assuming for the moment that σ is known, if the covariance is estimated using error-laden data by $\Sigma_\beta = \sigma(X'X)^{-1}$, then (2) implies

$$T\Sigma_{\beta} \overset{a.s.}{\rightarrow} \frac{\sigma}{\omega} Q^{-1}. \tag{5}$$

Thus, the covariance matrix is also proportionally underestimated by the same factor that β is underestimated. This causes the standard errors of regression coefficients to be proportionally underestimated and the asymptotically normal test statistics for zero regression coefficients to be overestimated by a factor of $\omega^{-1/2}$. Thus, rejections of hypothesis tests of zero coefficients are overly conservative because significance is underestimated.

Next, consider the more realistic case where σ is unknown and estimated by

$$\hat{\sigma} = \hat{\varepsilon}'\hat{\varepsilon}/T,\ \hat{\varepsilon} = y - X\hat{\beta} = My,\ M = I - X(X'X)^{-1}X'.$$

Substituting $y = X^{*}\beta + \varepsilon$ and using (2)-(4) obtains

$$\hat{\sigma} \overset{a.s.}{\rightarrow} \beta'Q\beta - (\beta'Q\beta/\omega) + \sigma. \tag{6}$$

While (5) implies underestimation, the variance of the regression disturbance is overestimated by an additive term $(\omega-1)\beta'Q\beta/\omega$ that increases both in the variance of data errors and means of the regression equation. Combining results in (5) and (6) proves:

Proposition 3.1. Under Assumption A, the covariance matrix estimated in practice when σ is unknown, $\hat{\Sigma}_{\beta} = \hat{\sigma}(X'X)^{-1}$, satisfies

$$T\hat{\Sigma}_{\beta} \overset{a.s.}{\rightarrow} \frac{\sigma}{\omega} Q^{-1} + \frac{\omega-1}{\omega^{2}} \beta'Q\beta Q^{-1}. \tag{7}$$

The first right-hand term tends to underestimation while the second right-hand term tends to overestimation. Thus, the covariance matrix estimated in practice has partially offsetting asymptotic errors. The errors are exactly offsetting if $\beta'Q\beta = \sigma\omega$. The estimated covariance matrix strongly converges to a matrix that differs from the true asymptotic covariance matrix by a scalar factor ($\beta'Q\beta$ is a scalar). Hence, standard test statistics err asymptotically, but by the same scalar factor for each regression coefficient.

4. ASYMPTOTIC DISTIBUTION OF THE OLS ESTIMATOR

Having established the almost sure convergence properties of the standard OLS estimator, we next examine the asymptotic distribution of the error-laden regression estimator. Because $\hat{\beta}$ converges strongly to β/ω, we examine the asymptotic distribution of

$\sqrt{T}(\hat{\beta}-\beta/\omega)$, which yields the covariance of the standard OLS estimator. Using (1) and factoring $(X'X/T)^{-1}$ obtains

$$\sqrt{T}(\hat{\beta}-\beta/\omega)=(X'X/T)^{-1}(V/\sqrt{T}) \tag{8}$$

where

$$V\equiv X^{*\prime}\Delta X^{*}\beta+X^{*\prime}\Delta\varepsilon-(X^{*\prime}\Delta^{2}X^{*}\beta/\omega)=\sum_{t}V_{t} \tag{9}$$

and

$$V_{t}\equiv\delta_{t}X_{t}^{*\prime}X_{t}^{*}\beta+\delta_{t}\varepsilon_{t}X_{t}^{*\prime}-(\delta_{t}^{2}X_{t}^{*\prime}X_{t}^{*}\beta/\omega)=X_{t}^{*\prime}[\delta_{t}\varepsilon_{t}+(\delta_{t}-\delta_{t}^{2}/\omega)X_{t}^{*}\beta].$$

Clearly, from the latter expression, $E(V_t)=0$, V_t is serially uncorrelated, and

$$E(V_{t}V_{t}')=(\sigma\omega+\bar{y}_{t}^{2}\eta)X_{t}^{*\prime}X_{t}^{*} \tag{10}$$

where $\eta=E[(\delta_{t}-\delta_{t}^{2}/\omega)^{2}]$ and $\bar{y}_{t}=X_{t}^{*}\beta$. Thus, from (9), the central limit theorem implies asymptotic normality of $V/\sqrt{T}$ and therefore of $\sqrt{T}(\hat{\beta}-\beta/\omega)$.

Proposition 4.1. Under Assumption A,

$$\sqrt{T}(\hat{\beta}-\beta/\omega)\overset{d}{\rightarrow}N\left[0,\frac{\sigma}{\omega}Q_{-1}+\frac{\eta}{\omega^{2}}Q^{-1}RQ^{-1}\right] \tag{11}$$

where

$$\sum_{t}\bar{y}_{t}^{2}X_{t}^{*\prime}X_{t}^{*}/T\overset{a.s.}{\rightarrow}R.$$

Proof: Under Assumption A, the overall covariance matrix of V_t (considering randomness of X_t) is constant implying asymptotic normality of $V/\sqrt{T}$ by the standard central limit theorem (Loéve, 1977, p. 286A). Bounded fourth moments and conditions on ε_t and δ_t are required for existence of the covariance matrix of V_t. Using (10) in absence of serial correlation of V_t implies

$$E(VV'/T)=\sum_{t}(V_{t}V_{t}'/T)=\sum_{t}(\sigma\omega+\bar{y}_{t}^{2}\eta)X_{t}^{*\prime}X_{t}^{*}/T\overset{a.s.}{\rightarrow}\sigma\omega Q+\eta R.$$

Because $V/\sqrt{T}$ is premultiplied by $(X'X/T)^{-1}$ in (8), pre- and post-multiplication of this equation by Q^{-1}/ω completes the proof.

Comparison of the covariance matrices in (7) and (11) suggests great similarity. First, the initial term of each is identical, $\sigma Q^{-1}/\omega$. Second, $\beta'Q\beta$ in (7) simply represents an asymptotic average of $\bar{y}_t^2$ because $\sum_t y_t^2/T \overset{a.s.}{\to} \beta'Q\beta$. However, $Q^{-1}R$ is another approach to averaging $\bar{y}_t^2$. That is, $(X^{*\prime}X^*)^{-1}\sum_t \bar{y}_t^2 X_t^{*\prime}X_t^*$ is an average with weights $(X^{*\prime}X^*)^{-1}X_t^{*\prime}X_t^*$, which sum to the identity matrix, $(X^{*\prime}X^*)^{-1}\sum_t X_t^{*\prime}X_t^* = (X^{*\prime}X^*)^{-1}X^{*\prime}X^* = I$. Thus, the tests performed naively using standard OLS methods are approximations of valid tests. However, the degree of approximation is not clear.

5. A CONSISTENT ESTIMATOR BASED ON OLS

Examination of the above results suggests a simple approach for asymptotic correction of the standard regression estimator. A consistent estimator of ω facilitates the needed correction. To obtain a consistent estimator of ω, note by the Kolmogorov law of large numbers that the asymptotic mean of the dependent variable satisfies

$$\bar{y} = \sum_t y_t/T = \sum_t X_t^*\beta/T + \sum_t \varepsilon_t/T \overset{a.s.}{\to} \tilde{q}\beta \tag{12}$$

where $\sum_t X_t^*/T \equiv \bar{X}^* \overset{a.s.}{\to} \tilde{q}$. On the other hand, by Proposition 2.1,

$$\bar{X}\hat{\beta} \overset{a.s.}{\to} \tilde{q}\beta/\omega. \tag{13}$$

where $\bar{X} \equiv \sum_t X_t/T \overset{a.s.}{\to} \tilde{q}$ by the Kolmogorov law of large numbers. If $\tilde{q}\beta \neq 0$, then comparing (12) and (13) reveals that

$$\tilde{\omega} = \bar{y}/(\bar{X}\hat{\beta}) \overset{a.s.}{\to} \omega. \tag{14}$$

We regard $\tilde{q}\beta \neq 0$ as a relatively harmless assumption because $\bar{y} > 0$ is plausible, for example,

in each of the motivating examples of the introduction and would seem to be so for any similar case that motivates proportional errors. Then from Proposition 2.1,

$$\hat{\beta}^* \equiv \tilde{\omega}\hat{\beta} \overset{a.s.}{\to} \beta,$$

and from (6),

$$\tilde{\sigma} = \hat{\sigma} - \hat{\beta}'(X'X/T)\hat{\beta}(\tilde{\omega}-1) \overset{a.s.}{\to} \sigma. \tag{15}$$

because from (2) and Proposition 2.1, $\hat{\beta}'(X'X/T)\hat{\beta} \overset{a.s.}{\to} \beta'Q\beta/\omega$. This proves:

Proposition 5.1. Under Assumption A, a proportional correction of the least squares regression estimator based on the ratio of the sample mean of the dependent variable to the sample mean of regression predictions obtains a consistent estimator of the regression parameters.

6. ASYMPTOTIC INFERENCE

The correction to the least squares estimator can also be obtained as a method-of-moments estimator. Let m_i and m_{ii} be the ith direct and cross sample moments, respectively, about the origin and let μ_i be the ith direct population moment about the origin. The method-of-moments estimator is found by setting the sample moments equal to their asymptotic limits, the population moments, and solving the corresponding set of equations. These equations are

$$m_1(X) - \mu_1(X^*) = 0;\ m_1(y) - \mu_1(X^*)\beta = 0;\ m_2(X) - \mu_2(X^*)\omega = 0. \tag{16}$$

$$m_2(y) - \beta'\mu_2(X^*)\beta - \sigma = 0;\ m_{11}(X,y) - \mu_2(X^*)\beta = 0. \tag{17}$$

where, as above, $\mu_1(X^*) = \tilde{q}$ and $\mu_2(X^*) = Q$. If β has dimension K, these equations contain K, 1, K^2, 1, and K restrictions, respectively, in the $K + 1 + 1 + K + K^2 \equiv L$ parameters in $(\beta', \sigma, \omega, \tilde{q}', vec(Q)') \equiv \theta'$. This system is just identified and its solution can be shown to agree with (14)-(15). Because the system is just identified, the weighting matrix in a GMM approach is irrelevant so (14)-(15) is the method-of-moments special case of GMM estimation. The asymptotic covariance matrix of sample moments is estimated by $W = (1/T)\sum_t (g_{jt} - \overline{g}_j)(g_{it} - \overline{g}_i), i,j = 1,\ldots,L$, where g_{jt} (e.g., y_t^2) is an element of the sum in the jth sample moment and $\overline{g}_j$ is the jth sample moment, e.g., $m_2(y) = (1/T)\sum_t y_t^2$. Letting $\hat{\xi}$ represent a matrix with jth row $\partial\overline{g}_j/\partial\theta$ evaluated at $\theta = \hat{\theta}$, an estimate of the asymptotic covariance of this method-of-moments estimator is

$$\hat{\Sigma}_{\hat{\theta}} = \hat{\xi}^{-1} W \hat{\xi}^{-1}. \tag{18}$$

From standard method-of-moments results, which can be applied here due to our assumed regularity conditions (Newey and McFadden, 1994), asymptotic normality follows. Using (16)-(18) thus provides the means of inference. For example, this distribution theory facilitates a test of the hypothesis $\omega = 1$, which if rejected implies existence of significant proportional errors in the regressor variables in support of the PEIV model.

7. ADDING AN INTERCEPT

Now, suppose the model is $y_t = \alpha + X_t^* \beta + \varepsilon_t$, $t = 1,...,T$, and continue with Assumption A. For expositional purposes, it is sufficient to identify and estimate the model assuming X_t^* is a 1×2 row vector with,

$$y_t = \alpha + X_{t1}^* \beta_1 + X_{t2}^* \beta_2 + \varepsilon_t, \; t = 1,...,T. \tag{19}$$

Cases with many variables measured without error and many other variables affected by the same proportional error follow in a straight forward manner as discussed below. We first suggest the estimator with an instrumental-variable-like approach as earlier and then formalize it in a

GMM framework to draw on familiar asymptotic results.

Assumption B. Where $z_t = X_{t1}/X_{t2}$, $z = (z_1,\ldots,z_T)'$, l is a vector of ones, and $W = (l, z, X^*)$, assume $E(W_t\delta_t) = E(W_t\varepsilon_t) = 0$, $t = 1,\ldots,T$, and $(1/T)\sum_t W_t W_t' \overset{a.s.}{\to} \tilde{Q}$ where $\tilde{Q}$ is positive definite with lower right partition Q. Let all other assumptions on $X_t, X_t^*, \varepsilon_t$, and δ_t follow Assumption A.

Applying instruments $(l,X)'$ to (19) obtains for this just-identified system

$$\begin{bmatrix} l'y/T \\ X'y/T \end{bmatrix} = \begin{bmatrix} \alpha \\ \alpha \bar{X} \end{bmatrix} + \begin{bmatrix} \bar{X}^* \beta \\ (1/T) X \bar{X}^* \beta \end{bmatrix} + \begin{bmatrix} \bar{\varepsilon} \\ (1/T) X' \varepsilon \end{bmatrix}$$

where bars denote averages. These yield the familiar moment equations $m_1(y) - \alpha - m_1(X^*)\beta = 0$

and $m_2(X, y) - \alpha m_1(X)' - m_2(X, X)\beta = 0$. Because, $m_1(X) - m_1(X^*) \overset{a.s.}{\to} 0$, $m_2(X, X) \overset{a.s.}{\to} \omega Q$, $m_2(X^*, X) \overset{a.s.}{\to} Q$, and $m_2(X^*, X^*) \overset{a.s.}{\to} Q$ as noted earlier, the equivalent moment conditions in terms of observable variables are

$$m_1(y) - \alpha - m_1(X)\beta = 0, \tag{20}$$

$$m_2(X,y) - \alpha m_1(X)' - [m_2(X,X)/\omega]\beta = 0. \tag{21}$$

Solving (20) for α and substituting into (21) yields

$$\hat{\alpha} = m_1(y) - m_1(X)\hat{\beta}, \tag{22}$$

$$\hat{\beta} = \left[\frac{m_2(X,X)}{\omega} - m_1(X)'m_1(X)\right]^{-1}\left[m_2(X,y) - m_1(X)'m_1(y)\right]. \tag{23}$$

where, by Assumption A and Proposition 33 of Dhrymes (1978), the inverse in (23) exists. If ω is known, then (23) is a consistent estimator for β that requires a proportional adjustment of part of the least squares data matrix. The estimator of α in (22) is identical to the usual least squares estimator once β is estimated appropriately.

For the case where ω is unknown, the necessary additional moment equation is available using z. That is, applying instruments $(l,z,X)'$ to (19), instead of $(l,X)'$ as above, obtains an additional equation, $z'y/T = \alpha\bar{z} + (1/T)z'X^*\beta + (1/T)z'\varepsilon$, which yields the moment equation,

$$m_2(z,y) = \hat{\alpha}m_1(z) - m_2(z,X^*)\hat{\beta}. \tag{24}$$

Replacing X^* by X, noting that $m_2(z,X) - m_2(z,X^*) \xrightarrow{a.s.} 0$, and using (22) yields

$$\left[m_2(z,y) - m_1(z)m_1(y)\right] - \left[m_2(z,X) - m_1(z)m_1(X)\right]\hat{\beta} = 0,$$

which, upon substituting (23), obtains

$$\tilde{m}_2(z,y) - m_2(z,X)\left[\frac{m_2(X,X)}{\omega} - m_1(X)'m_1(X)\right]^{-1}\tilde{m}_2(X,y) = 0 \tag{25}$$

where $\tilde{m}$ represents centered moments. Although (25) does not have a closed form solution for ω, it facilitates easy numerical solution and, once solved, the unconditional estimators of β and σ can be found from (22) and (23).

To represent this approach in a common GMM setting, let[2]

[2] For the two-regressor case we abuse slightly the GMM terminology, which is generally reserved for the case with over-identifying restrictions. As noted in the conclusions for the more general case, however, such over-identifying restrictions will generally be present.

$$g(z_t,\theta)=\begin{bmatrix} y_t-\alpha-X_t'\beta \\ X_t y_t-\alpha X_t - X_t X_t'(\beta/\omega) \\ z_t y_t-\alpha z_t - z_t X'\beta_t \end{bmatrix}, \tag{26}$$

which yields $E[g(z_t,\theta)] = 0$, $t = 1,...,T$. The method of moments estimator, $\hat{\theta}$, solves

$$(1/T)\sum_t g(z_t,\hat{\theta})=0. \tag{27}$$

With two regressors, four moment equations are obtained in the four unknowns, α, β_1, β_2, and ω. To attain almost sure convergence, consistency, and asymptotic normality, we assume standard regularity conditions (see, e.g. Hansen, 1982).

Assumption C. Let θ_0 be the true parameter value in the interior of a compact parameter space, Θ. Assume g is (i) continuous and measureable for all z_t, (ii) g is first moment continuous for all $\theta\in\Theta$, (iii) $E[g(z_t,\theta)]$ exists and is finite for all $\theta\in\Theta$, and (iv) $E[g(z_t,\theta_0)] = 0$ uniquely at $\theta = \theta_0$. Further, assume the gradient of g denoted by $\partial g/\partial\theta$ satisfies properties (i)-(iii) and is full rank for all $\theta\in\Theta$.

Proposition 7.1. Under Assumptions B and C and the definition in (26), the method of moments estimator, $\hat{\theta}$, that solves (27) satisfies $\sqrt{T}(\hat{\theta}-\theta)\xrightarrow{d} N(0,G\Omega^{-1}G)$ where $G=E(\partial g/\partial\theta)$ and $\Omega=E[g(z_t,\theta)g(z_t,\theta)']$.

Proof: Conveniently, the estimator defined by (27) is just identified. Thus, given conventional regularity conditions satisfied by Assumptions B and C (e.g., Hansen, 1982), the method of moments estimator is consistent and asymptotically normal (see also Newey and McFadden, 1994).[3]

Although G is easy to compute analytically and thus estimate, this is not so for Ω because it involves higher moments. Alternatively, standard numerical GMM methods allow inference using $\hat{G}=(1/T)\sum_t \partial g/\partial\theta(z_t,\hat{\theta})$ and $\hat{\Omega}=(1/T)\sum_t[\partial g/\partial\theta(z_t,\hat{\theta})][\partial g/\partial\theta(z_t,\hat{\theta})]'$. However, these results do not facilitate an analytical comparison of OLS and GMM standard errors.

[3] While we calculate standard errors in this paper with a robust approach similar to White (1984). The covariance matrix in Proposition 7.1 can also be developed based on the specific structure of the problem. While this alternative approach may be more accurate in some settings, we opt for a more robust approach here.

8. A MONTE CARLO COMPARISON

Comparing OLS and GMM estimators, one would expect measurement errors in addition to typical sampling errors to create additional instability of parameter estimates. Although OLS estimates are inconsistent, it does not follow for small samples that OLS is dominated by a (consistent) GMM estimator. In order to shed light on these issues, we examine the simple linear model in (19) numerically.

While we have examined a number of parametric variations with similar results, we report here only a typical set of cases with parameters $\alpha = 1$, $\beta_1 = 2$, and $\beta_2 = 5$. True regressors are randomly generated from uniform distributions with respective means of 8 and 10 and ranges of 0 to 16 and 0 to 20, and are fixed throughout all replications. Regressors with proportional errors in variables are randomly generated in each iteration where the proportional error is normally distributed with a unit mean. The standard deviation of the proportional error is parametrically varied among values 0.05, 0.15, and 0.30, which corresponds, respectively, to ω = 1.0025, 1.0225, and 1.0900. The dependent variable is generated using the true regressor data, the parameters, and a random normal regression error. For all reported results, normal regression errors were generated with mean zero and standard deviation 10. The resulting mean for the dependent variable is approximately 70. The number of replications is 1,000 throughout.

Table 1 compares the OLS estimator with the GMM estimator defined by (22)-(24) or (25). All reported biases and mean squared errors are measured as proportions of the true parameters. Consider first the case with $\omega = 1.09$, which represents large but not unrealistic proportional measurement error.[4] Examining the bias column for OLS estimates, the biases are negative for all slope parameters (attenuation bias) as expected, implying with the chosen parameterization that the intercept bias is positive. All biases are large, but the intercept bias is particularly so because it depends on the attenuation results for the slope parameters and the mismeasured means of the data. Note following (22), however, that the bias of the intercept is essentially determined by the bias of the slope coefficients and that this bias becomes arbitrarily large as means of the nonconstant regressors are chosen arbitrarily far from zero. At a sample size of 40, the smallest bias is 29.4 percent.[5] With a doubling of sample size to 80, OLS biases do not change materially. This is true even for the largest reported sample size of 1,000.[6] Mean squared errors (MSE) for OLS follow much the same pattern. They diminish but not substantially with increasing sample size.

[4] See Pope and Just (1996) for a case with very large measurement errors due to the econometrician incorrectly assuming that agricultural output is known ex ante in cost function estimation.

[5] The percent biases for the three parameters follow the same qualitative order as for OLS on the true data. These are: $\alpha = 0.243$, $\beta_1 = -0.687$, $\beta_2 = -0.002$. That is, sampling error yields the largest bias for the intercept followed by the first and then the second slope parameters.

[6] In fact, the bias for α is slightly higher for 1000 observations than for 80. In the PEIV model, increasing sample size does not diminish measurement error. Thus, biases do not approach zero in sample size. We suggest that this is a plausible outcome for most problems in economics.

Turning to the GMM results in Table 1, bias and MSE fall dramatically with increasing sample size as expected. For example, a doubling of the sample size from 40 to 80 reduces bias for α from 1444.4 percent to –60.1 percent. The change in the bias of slope coefficients is similarly impressive, declining from 13-41 percent to around 1 percent. Interestingly, however, even when the sample size is quite large (1,000), the bias of slope coefficients can still be about half of one percent and the resulting constant term bias can be much larger. As for OLS, the MSE is much larger for the intercept than the slope parameters owing to the magnitude of mean regressor data. However, for GMM, the MSE falls dramatically as sample size is increased and becomes negligible at 1,000 observations with the exception of the intercept. Note that using z as an instrument permits estimation of ω with little bias based on as few as 40 observations.

Comparing estimators for the case where measurement error is lower reveals a somewhat different comparison. When measurement error is small with $\omega = 1.0025$, OLS dominates GMM with respect to both bias and MSE with 40 observations. When the sample size doubles to 80, bias is lower but MSE is larger with GMM. For large samples, bias is much lower for GMM than OLS but MSE is still larger. In both cases, bias and MSE tends to fall substantially as sample size is increased.

When measurement error has an intermediate magnitude with $\omega = 1.0225$, OLS has a much larger bias than GMM throughout the reported sample range. The bias falls dramatically with sample size for the GMM estimator but only slightly for OLS. However, MSE is far higher for GMM with 40 observations. As sample size doubles to 80, MSE for GMM is approximately 3 times larger than for OLS. For large samples, GMM dominates with respect to both bias and MSE.

In summary, this Monte Carlo experiment suggests the following qualitative conclusions (which are also consistent with unreported results):

1. For small measurement error and small sample sizes, OLS dominates with respect to both bias and MSE.
2. For intermediate levels of measurement error and intermediate sample sizes, GMM tends to yield less bias but larger MSE.
3. For large measurement error and large sample sizes, GMM tends to have both less bias and lower MSE.
4. Bias falls dramatically with sample size for GMM but does not decline appreciably with sample size for OLS.
5. In relative terms, MSE falls much faster with sample size for GMM than for OLS but is much higher for small samples than OLS.

These results suggest that PEIV estimation methods are useful for some problems in smaller sample sizes and for most problems in large samples. For example, estimation of cost functions for the case of stochastic production may typically be a case of large measurement error if unanticipated weather causes substantial production variation as in agriculture. On the other hand, demand estimation with errors in price indexes may typically have smaller relative measurement error and thus suggest use of PEIV estimation methods only with large samples.

9. AN EMPIRICAL ILLUSTRATION

In this section we apply the PEIV model to a typical problem of agricultural factor demand and technical change estimation (Binswanger 1974, Antle 1984) with the widely used aggregate agricultural data created by Ball et al. (1997). The data are annual from 1948-1994 and consist of aggregate agricultural output, q, input prices, r_i, and input quantities, x_i, for capital (i = 1), chemicals (i = 2), fuel (i = 3), feeds (i = 4), labor (i = 5), and other purchased inputs (i = 6). In addition, exogenous technical change is introduced by including the year and year squared as regressors. The equations for input demands are assumed to follow from the generalized Leontief cost function for as discussed in the introduction aside from a quadratic technical change relationship, $b_{ii} + c_i t + d_i t^2$, substituted for each b_{ii}:

$$x_{it} = b_{ii}\overline{q}_t + \sum_{j\neq i} b_{ij}\overline{q}_t (r_{jt}/r_{it})^{1/2} + c_i\overline{q}_t t + d_i\overline{q}_t t^2 + \varepsilon_{it},\ i = 1,...,6;\ t = 48,...,94.$$

As earlier, this regression can be stated in standard linear form, $y = X^*\beta + \varepsilon$, $E(\varepsilon) = 0$, where the first six columns of X^* have typical element $\{\overline{q}_t (r_{jt}/r_{it})^{1/2}\}$ with associated parameter b_{ij} and the last two columns have the year and year squared with respective coefficients c_i and d_i. We assume that actual output q_t rather than expected output $\overline{q}_t$ is used in least squares regressions where $q_t = \overline{q}_t \delta_t$, $E(\delta_t) = 1$, $t = 1,\ldots,T$, $E(\delta\delta') = \omega I_T I$, $E(\delta\varepsilon') = 0$.

To verify that necessary assumptions are plausible for this application, consider the stylized case of crop production where farmers make input decisions such as seed and fertilizer use, land preparation methods, and intensity of land use at the time of crop planting. Then weather conditions are realized during the growing season, which affect eventual actual production. Observed ex post production thus varies stochastically from unobserved planned ex ante production because input decisions are made at the time of production planning before ensuing weather is known. While exceptions to this timing can be identified, they are typically believed to be minor by comparison so that this conceptual framework represents the bulk of agricultural production.

Under expected profit maximization, standard dual cost theory applies to the optimization problem at the time of production planning, so input quantities depend on ex ante rather than ex post actual output. Thus, in terms of the notation of the model, x represents the quantity of input demanded, X represents functions of input prices times actual ex post output, and X^* represents functions of input prices times ex ante output. Because ex ante output and the associated input decisions are made before weather is realized, they are unlikely to depend on weather conditions that are not yet known, i.e., X is likely independent of δ. The errors in the input demands, on the other hand, represent errors in optimization for individual inputs. Because these decisions are made before weather is realized, they are also unlikely to depend on weather, i.e., ε is likely independent of δ. Finally, dependence of weather on farmer's earlier input decisions is implausible, i.e., δ is likely independent of X^*. For these reasons, the independence of Assumption A is highly plausible. (While we have not argued the independence of ε and X here, this is the standard assumption that has been used in conventional estimation problems of this type, and is discussed at length elsewhere.)

Table 2 presents system estimates of the model where the standard symmetry conditions ($b_{ij} = b_{ji}\ \forall\ i,j$) are imposed consistent with a common underlying generalized Leontief cost function. Because all input demand equations are affected by the same proportional errors in actual ex post output as a measurement of ex ante output, a common underlying generalized Leontief cost function also implies a common ω parameter across equations. Thus, only one estimate of the ω parameter appears in Table 2. Nevertheless, as a test of the applicable theoretical model of PEIV, we tested whether the same ω parameter applies to all equations. By estimating the model allowing distinct parameters in each equation, the estimates of the remaining parameters were almost identical (differed by no more than about 1 or 2 in the second significant digit) and the estimated parameters were almost the same, ranging from 1.00032 to 1.00092 with standard errors ranging from 0.00022 to 0.00038.[7] The chi-square statistic for testing whether all parameters are identical was 0.00000022 with 5 degrees of freedom, which overwhelmingly favors the conclusion that the same PEIV problem affects all 6 input demand equations.

As for the estimates in Table 2, a respectable 20 of 34 coefficient estimates are significant at the 5 percent level and 14 of 34 are significant at the 1 percent level. The R^2 statistics are generally high, particularly for chemicals and labor. Technical change is labor saving and capital, chemical, and other input using, which are highly plausible for U.S. agriculture over the 1948-1994 period. With the quadratic form in time, technical change switches from fuel and feed using to fuel and feed saving in the latter part of the estimation period, which also seems quite plausible. Estimated demand elasticities are appropriately negative

[7] In GMM or instrumental variables (IV) estimation, where Z is a $T \times K$ matrix of non-random instruments, the standard model has covariance $Z'\Omega Z$ associated with $Z'\varepsilon = Z'(y - X^*\beta)$ where $\Omega = E(\varepsilon\, \varepsilon') = \sigma^2 I_T$ under homoskedasticity. With PEIV, however, the corresponding covariance in terms of observables follows from $Z'(y - X\beta) \equiv Z'[(X^* - X)\beta + \varepsilon]$, which also has mean zero but generates the heteroskedastic covariance matrix $Z'[\sigma^2 I_T + (\omega - 1)D]Z$ where D is a diagonal $T \times T$ matrix with typical element $(X_t^*\beta)^2$. For the case of random instruments such as appear in X, the same principle applies but yields a more complicated error term. Because the predicted dependent variables in this application vary relatively little compared to their magnitude, and the estimated ω is much closer to 1 than the estimated σ^2 is close to zero, the errors in this application were very close to homoskedastic. For this estimation problem we have not employed the well-known heteroskedastic-autocorrelation correction (HAC) method of Newey and West (1987), which yields consistent standard errors in any case, because data are not sufficient. As the number of orthogonality conditions grows relative to sample size and number of parameters, finite sample identification may be aided but the ability to compute HAC standard errors may diminish. Due to the large number of inputs (and hence, orthogonality conditions) which exceeds the number of observations in our application, we report nonlinear three-stage least squares standard errors in Table 2.

at sample means for all inputs and at every observation for all inputs other than capital where one observation generated a very small positive elasticity of 5×10^{-5}. Estimated elasticities range over time from –0.41 to –0.94 for chemicals, –0.06 to –0.14 for fuels, –0.14 to –0.42 for feed, –0.10 to 0.48 for labor, –0.01 to –0.10 for other inputs, and range from near zero to –0.03 for capital. The vast majority of these estimates are plausible both in sign and magnitude. While other diagnostic statistics could be presented and discussed, we forego these here because, apart from considering PEIV, this is a common application with agricultural data and similar investigations have been reported elsewhere.

Turning to the parameter of central importance for this paper, the estimated common parameter ω is significantly greater than 1.0 at the 1 percent level. Thus, the results imply that PEIV is statistically significant in this application. Therefore, the results support the assertion that actual output has statistically significant errors as a representation of planned output that affect all cross price terms in generalized Leontief cost function estimation for agriculture.

10. CONCLUDING DISCUSSION

This paper explores estimation of the PEIV model where the proportional error in variables is identical across regressors other than the constant term. Such models arise naturally in many microeconomic applications. Application of least squares to such models causes attenuation bias for slope coefficients whether or not an intercept is present. Results show that standard hypothesis tests may be approximate but no simple criteria are available to determine accuracy. Alternatively, an asymptotic correction is identified as a minor modification of OLS obtaining strong consistency. The estimator converges in distribution to a multivariate normal making inference relatively convenient. While typical historical practice has been to use logged data when errors are multiplicative as a matter of convenience, many microeconomic models that generate internally consistent behavior do not admit such convenience. Further, we suggest that the methodology of this paper, particularly for cases that do not generate a constant term as in the empirical application presented here, is particularly convenient and allows intellectual honesty in the application of such structural models.

For cases with a constant term that is not subject to error, estimation can be accomplished using the ratios of observed variables subject to proportional error as instruments. Although the approach is developed here only for the case with an intercept and two mismeasured regressors, the approach can be easily expanded to the case with multiple regressors measured without error (which are treated like l above) and two or more regressors have a common proportional error. In both cases, moment equations must be specified such that the orthogonality conditions hold and identification is possible. With more than two mismeasured regressors, however, more ratios of mismeasured variables are available as instruments than necessary, which leads to overidentification. Where k variables are mismeasured, $k(k-1)$ possible ratios like z are available. For efficiency, consideration must then be given to optimal weighting matrices and instruments (Hansen, 1982).

As a final comment, equation (25) is nonlinear in ω and the mismeasured variables. We have not examined the creation of optimal moments (or instruments). Improvements are likely possible by choosing instruments based on the predicted gradient of the moment equations (see Donald and Newey 2001). This may involve polynomials in the instrument z. In this case, efficiency would dictate a true GMM use of an appropriate weighting matrix based on the covariance matrix of the moment equations. We leave these possibilities to a future paper.

REFERENCES

Amemiya, Y. (1990). Two-stage instrumental variable estimators for the nonlinear errors-in-variables model, *Journal of Econometrics*, 44, 311-332.

Antle, J.M. (1984). The structure of U.S. agricultural technology, 1910-78, *American Journal of Agricultural Economics*, 66, 414-421.

Ball, E. V., Bureau, J-C., Nehring, R. and Somwaru, A. (1997). Agricultural productivity revisited, *American Journal of Agricultural Economics*, 79, 1045-1063.

Berndt, E. R. and Wood, D.O. (1975). Technology, prices, and the derived demand for energy, *Review of Economics and Statistics*, 57, 259-268.

Binswanger, H. (1974). The Measurement of technical change biases with many factors of production, *American Economic Review*, 64, 964-976.

Blundell, R. (1988). Consumer Behavior, Theory and empirical evidence—a survey, *Economic Journal*, 98, 16-65.

Briggs, F.E.A. (1962).The influence of errors on the correlation of ratios, *Econometrica*, 30, 162-177.

Burt, O.R. and Brewer, D. (1971). Estimation of net social benefits from outdoor recreation, *Econometrica*, 39, 813-827.

Casson, M.C. (1973). Linear regression with error in the deflating variable, *Econometrica*, 41, 751-759.

Cragg, J.G. (1997). Using higher moments to estimate the simple errors-in variables model, *Rand Journal*, 28, S71-S90.

Diewert, W.E. (1971). An application of the Shephard duality theorem: a generalized Leontief, *Journal of Political Economy*, 79, 481-507.

Donald, S.G. and Newey, W.K. (2001). Choosing the number of instruments, *Econometrica*, 69, 1161-1191.

Dhrymes, P. (1978). *Mathematics for Econometrics*. Springer-Verlag: New York.

Fuss, M. (1977). The structure of technology over time: a model for testing the "putty-clay" hypothesis, *Econometrica*, 45, 1797-1821.

Garber, S. and Klepper, S. (1980). Extending the classical normal errors in variables model, *Econometrica*, 48, 1541-1546.

Gorman, W.M. (1959). Separable utility and aggregation, *Econometrica*, 27, 469-481.

Griliches, Z. and Ringstad, V. (1970). Error-in-variables bias in nonlinear contexts, *Econometrica*, 38, 368-370.

Hansen, L. P. (1982). Large sample properties of method of moments estimators, *Econometrica*, 50, 1029-1054.

Hausman, J. (2001). Mismeasured variables in econometric analysis: problems from the right and problems from the left, *Journal of Economic Perspectives*, 15, 57-67.

Hausman, J.A., Newey, W. K., Ichimura, H. and Powell, J.L. (1991). Identification and estimation of polynomial errors-in variables models, *Journal of Econometrics,* 50, 273-295.

Heien, D. (1977). The cost of the U.S. dairy price support program: 1949-1974, *Review of Economics and Statistics*, 59, 1-8.

Humphrey, D.B. (1975). Estimates of factor-intermediate substitution and separability, *Southern Economic Journal*, 41, 531-534.

Izumida, N., Urushi, H. and Nakanishi, S. (1999). An empirical study of the physican-induced demand hypothesis—the cost function approach to medical expenditure of the elderly in

Japan, *Review of Population and Social Policy*, 8, 11-25.

Kuh, E. and Meyer, J. (1955). Correlation and regression estimates when the data are ratios, *Econometrica*, 23, 400-416.

Lewbel, A. (1987). Characterizing some Gorman Engel curves, *Econometrica*, 55, 1451-1459.

Loève, M. (1977). *Probability Theory*. 4th edn. Van Nostrand: Princeton.

Nelson, D. B. (1995). Vector attenuation bias in the classical errors-in-variables model, *Economics Letters*, 49, 345-349.

Newey, W. and McFadden, D. (1994). Large sample estimation and hypothesis testing. In *Handbook of Econometrics*, Vol. 4, eds. R.F. Engle, R.F., D. McFadden, Elsevier: New York, 2111-2245.

Newey, W. (2001). Flexible simulated moment estimation of nonlinear errors-in-variables models, *Review of Economics and Statistics*, 83, 616-627.

Newey, W.K. and West, K.D. (1987). A simple positive definite heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica*, 55, 703-8.

Pope, R.D. and Just, R.E. (1996). Empirical implementation of ex ante cost functions, *Journal of Econometrics*, 72, 231-49.

Riersol, O. (1950). Identifiability of a linear relation between variables which are subject to error, *Econometrica*, 18, 375-89.

Shephard, R. (1970). *Cost and Production Functions*. Princeton University Press: Princeton, NJ.

Song, D. and Hallberg, M. (1982). Measuring producers' advantage from classified pricing of milk, *American Journal of Agricultural Economics*, 64, 1-8.

Theil, H. (1961). *Economic Forecasts and Policy*. North-Holland: Amsterdam.

White, H. (1984). *Asymptotic Theory for Econometricians*. Academic Press: New York.

Table 1. Comparisons of Ordinary Least Squares and Method of Moments Estimators

Magnitude of Errors		Ordinary Least Squares		Generalized Method of Moments	
Sample Size	Parameter Estimator	Bias[a]	Mean-Squared Error	Bias	Mean-Squared Error
		- - - - - - - - - - - - - Magnitude of error, ω = 1.0900 - - - - - - - - - - - - -			
T = 40	α	28.800	868.904	14.444	4327.61
	β_1	–0.696	0.549	–0.410	3.781
	β_2	–0.294	0.095	–0.131	0.336
	ω	–	–	0.007	0.019
T = 80	α	24.402	612.703	–0.601	1306.840
	β_1	–0.672	0.481	0.011	1.466
	β_2	–0.273	0.079	0.010	0.120
	ω	–	–	0.002	0.005
T = 1,000	α	25.097	631.164	–0.047	9.055
	β_1	–0.688	0.476	0.005	0.015
	β_2	–0.283	0.080	<–0.001	<0.001
	ω	–	–	<0.001	<0.001
		- - - - - - - - - - - - - Magnitude of error, ω = 1.0225 - - - - - - - - - - - - -			
T = 40	α	10.340	133.558	–2.908	3192.322
	β_1	–0.254	0.112	0.095	2.614
	β_2	–0.103	0.017	0.023	0.260
	ω	–	–	–0.001	0.007
T = 80	α	8.423	83.657	–1.312	249.078
	β_1	–0.243	0.082	0.040	0.248
	β_2	–0.091	0.011	0.013	0.027
	ω	–	–	–0.004	0.001
T = 1,000	α	8.772	77.890	0.017	4.969
	β_1	–0.247	0.063	0.001	0.007
	β_2	–0.097	0.010	<–0.001	<0.001
	ω	–	–	<–0.001	<0.001
		- - - - - - - - - - - - - Magnitude of error, ω = 1.0025 - - - - - - - - - - - - -			
T = 40	α	1.529	22.011	–2.072	708.953
	β_1	–0.037	0.032	0.070	0.488
	β_2	–0.015	0.004	0.015	0.069
	ω	–	–	–0.010	0.003
T = 80	α	1.006	10.639	–0.435	145.934
	β_1	–0.033	0.016	0.011	0.124
	β_2	–0.010	0.002	0.004	0.018
	ω	–	–	–0.003	<0.001
T = 1,000	α	1.099	1.931	0.029	3.784
	β_1	–0.031	0.002	<–0.001	0.004
	β_2	–0.012	<0.001	<0.001	<0.001
	ω	–	–	<0.001	<0.001

[a] Note that bias and mean squared error are measured as proportions of true parameters where $\alpha = 1$, $\beta_1 = 2$, and $\beta_2 = 5$.

Table 2. Estimates of Conditional Input Demands with PEIV for U.S. Agriculture, 1948-1994.[a]

	Demand Equation					
Coefficient	Capital	Chemicals	Fuels	Feed	Labor	Other Inputs
Capital Price	0.4684 (.4556)	[b]				
Chemical Price	–0.0339 (0.0377)	–3.7107* (0.6144)				
Fuel Price	–0.0134 (0.0408)	0.3551* (0.1213)	0.6726 (0.6434)			
Feed Price	–0.0845⁺ (0.0385)	0.3981* (0.0854)	–0.0173 (0.0709)	0.2645 (0.3730)		
Labor Price	–0.0311 (0.0700)	0.3671* (0.1006)	0.0135 (0.0951)	0.3235* (0.0914)	22.1407* (0.7689)	
Other Price	–0.0284 (0.0679)	–0.3589* (0.1296)	–0.1138 (0.1326)	0.0313 (0.1047)	0.3921⁺ (0.1846)	0.5467 (0.7053)
Year	0.0441* (.0110)	0.0947* (0.0130)	0.0265 (0.0173)	0.0246⁺ (0.0102)	–0.4895* (0.0176)	0.0333 (0.0211)
Year Squared	–0.00043* (0.00007)	–0.00060* (0.00008)	–0.00029⁺ (0.00011)	–0.00027* (0.00007)	0.00272* (0.00011)	–0.00033⁺ (0.00014)
ω	1.00046* (00013)					
R^2	0.75	0.96	0.69	0.88	0.99	0.67
Elasticities (at means)	–0.009	–0.57	–0.10	–0.25	–0.26	–0.04

[a] Standard errors appear in parentheses. Significance at the 5% level is indicated by '⁺' and at the 1% level by '*'. In the case of ω, significance corresponds to a one-sided test of H_0: $\omega = 1$ versus H_1: $\omega > 1$ while all other significance levels correspond to two-sided tests of zero coefficients.

[b] Note that blanks for estimated coefficients in the table correspond to coefficients constrained to be the same as other coefficients in the table (either by symmetry or the commonality of ω across all equations).