Department of Agricultural and Resource Economics
University of California, Davis

A Dual Least-Squares Estimator of the Errors-In-Variables Model Using Only First And Second Moments

by

Quirino Paris

Working Paper No. 14-003

June 2014

Giannini Foundation of Agricultural Economics

A Dual Least-Squares Estimator of the Errors-In-Variables Model Using Only First And Second Moments

Quirino Paris

June 2014

Abstract
The paper presents an estimator of the errors-in-variables in multiple regressions using only first and second-order moments. The consistency property of the estimator is explored by Monte Carlo experiments. Based on these results, we conjecture that the estimator is consistent. The proof of consistency, to be dealt in another paper, is based upon the assumptions of Kiefer and Wolfowitz (1956). The novel treatment of the errors-in-variables model relies crucially upon a neutral parameterization of the error terms of the dependent and the explanatory variables. The estimator does not have a closed form solution. It requires the maximization of a dual least-squares objective function that guarantees a global optimum. This estimator, therefore, includes the naïve least-squares method (when only the dependent variable is measured with error) as a special case.

Keywords: errors-in-variables, measurement errors, dual least squares, first moments, second moments, Monte Carlo

JEL: C30

Quirino Paris is a professor in the Department of Agricultural and Resource Economics, University of California, Davis.

A Dual Least-Squares Estimator of the Errors-In-Variables Model Using Only First And Second Moments

Quirino Paris

June 2014

1. Introduction

For more than a century, statisticians have attempted to solve the problem of obtaining consistent parameter estimates (intercept, slope, error variances) in a linear regression where both dependent and independent (explanatory) variables are subject to measurement errors. Traditionally, this problem has gone by the name of the errors-in-variables (EIV) model. A varying degree of success was attained during a century of efforts. One desirable objective, however, has escaped so far: the goal of obtaining an easy method for consistent parameter estimates under general and plausible assumptions using only first and second-order moments of the sample information. This goal is desirable for a number of reasons: simplicity, relation to maximum likelihood estimates under normality assumptions, global optimum, and avoidance of the difficulties associated with obtaining valid measures of the fourths and higher moments in empirical analyses.

Adcock (1878) appears to have pioneered the discussion of the EIV model by recognizing that the naïve least-squares method has "a problem." His suggestion was to compute orthogonal (to the estimated regression line) errors for which he had to assume equal variances of the errors in both dependent and independent variables. This means that he had to augment the usual sample information with additional out-of-sample knowledge about errors variances. Over time, researchers have suggested a variety of approaches that recognize the prevailing consensus about the need of additional (user supplied) information in order to obtain the desired properties of the parameter estimates. Gillard (2010) presents an overview of various approaches covering instrumental variables, maximum likelihood, the moment method and others. Several important papers by Reiersøl (1950), Neyman (1951), Wolfowitz (1954) and Kiefer and Wolfowitz (1956), however, are omitted. All these papers establish the strong consistency of maximum likelihood estimates of a rather general EIV structural regression subject only to the exclusion condition of Reiersøl (1950). His remarkable theorem states that – to achieve identification of the parameter estimates – the latent variable(s) of the EIV model cannot be distributed as normal random variables. It appears that Kiefer and Wolfowitz (1956) results have been neglected, especially in recent years. Kendall and Stuart (1979), for example, in their fourth edition of volume 2, dealing with maximum likelihood and the EIV structural model, mention neither Reiersøl (1950) nor Kiefer and Wolfowitz (1956).

Another example of omitted important literature is associated with the works of Pal (1980), Van Monfort *et al*. (1987), Cragg (1997) and Dagenais and Dagenais (1997) who discuss the method of moments in the context of the EIV model estimation but do not mention Neyman (1951), Wolfowitz (1954) and Kiefer and Wolfowitz (1956). Their idea is to use sample moments of order higher than the first and second one to obtain estimates of the slope coefficient and then derive the other parameters from relations based upon first and second moments. There are some warnings. If the latent random variable is distributed according to a symmetric distribution (normal, uniform), the third moment vanishes and it is necessary to use forth moments to estimate the slope. The variance of the estimated coefficient, then, involves eighth-order moments. The higher the moments, the more demanding the information requirement for their valid measure.

In recent years, nonparametric methods have been applied to the EIV problem (see Delaigle and Meister, 2007). But, often, the importance of knowing the structure of the regression function explicitly and the dimension of the individual parameters (elasticities) works in the direction of parametric estimation.

Also, we are not interested in sample moments of higher order. As the paper's title states, we discuss a rather simple procedure that achieves consistent parameter estimates using only first and second-order moments of the sample information. In section 2 we define the simplest EIV structural model along the lines of many predecessors and in particular of Lindley (1947), Wolfowitz (1954), Kiefer and Wolfowitz (1956) and Kendall and Stuart (1979). We consider the second-order moment relations derived by these authors who noted that for the simplest EIV structural model there are three second-order moment relations but four parameters to estimate. Hence, by considering only these second-order moments, the EIV model is not identified.

In this paper, we introduce a neutral transformation of the random errors and set up a mathematical programming model that, in principle, exhibits a global maximum and achieves consistent estimates of all parameters involved. Our estimator does not have a closed form solution. It requires a numerical optimization software like GAMS. We do not use a maximum likelihood approach. Instead, we use a dual least-squares methodology along the lines discussed by Paris (2011). In section 3, we give a brief and simple introduction to the dual of the least-squares method and note that the dual least-squares specification possesses a global maximum over the parameter space. In section 4, we extend the estimator to three latent random variables (one dependent and two "explanatory"). In section 5, we present a Monte Carlo experiment dealing with an EIV regression with two latent variables (one dependent and one "explanatory") and involving 6 parameters. In section 6, we present a Monte Carlo experiment dealing with an EIV regression with three latent variables (one dependent and two "explanatory") and involving 10 parameters. The results of both Monte Carlo experiments do not negate the conjecture that the dual LS estimator produces consistent estimates of the EIV model's parameters. In section 7, we generalize the specification of the EIV dual LS estimator to K latent "explanatory" variables in vector and matrix notation. Conclusions come in section 8.

2. The Simplest EIV Case

Following Kendall and Stuart (1979, p. 400), we state a linear relation between two latent random variables Y^* and X^* in the form of

$$Y^* = \alpha + \beta X^* \tag{1}$$

with the objective of estimating parameters α and β. As Y^* and X^* are latent variables they are not observed directly. In their place, we measure repeatedly two random variables Y and X that bear the following relations with the latent variables

$$y_i = y_i^* + u_i^* \tag{2}$$

$$x_i = x_i^* + v_i^* \tag{3}$$

$i = 1,...,N$, where u_i^* and v_i^* are $i.d.d.$ measurement errors (deviations) from the true value of the latent variables Y^* and X^*.

In this simplest case, we assume that

$$E(u_i^*) = E(v_i^*) = 0, \quad \mathrm{var}(u_i^*) = \sigma_{u^*}^2, \ \ \mathrm{var}(v_i^*) = \sigma_{v^*}^2 \quad \text{all } i,$$
$$\mathrm{cov}(u_i^*, u_j^*) = \mathrm{cov}(v_i^*, v_j^*) = 0, \qquad i \neq j, \tag{4}$$
$$\mathrm{cov}(u_i^*, v_j^*) = 0, \quad \text{all } i, j.$$

The EIV model, therefore, can be restated as

$$y_i = \alpha + \beta x_i^* + u_i^* \tag{5}$$
$$x_i = x_i^* + v_i^*. \tag{6}$$

The use of assumptions (4) allows the conclusion that

$$\begin{aligned} E(x) &= E(x^*) = \mu \\ E(y) &= \alpha + \beta\mu \end{aligned} \tag{7}$$

Lindley (1947) and Kendall and Stuart (1979) stated the second-order moment relations:

$$\mathrm{var}(y) = \beta^2 \sigma_{x^*}^2 + \sigma_{u^*}^2 \tag{8}$$
$$\mathrm{var}(x) = \sigma_{x^*}^2 + \sigma_{v^*}^2 \tag{9}$$
$$\mathrm{cov}(y, x) = \beta \sigma_{x^*}^2. \tag{10}$$

By using first- and second-order sample moments to approximate the left-hand-side population moments of relations (7)-(10), the list of EIV conditions can be restated as

$$\overline{x} = \mu \tag{11}$$
$$\overline{y} = \alpha + \beta\mu \tag{12}$$
$$m_{yy} = \beta^2 \sigma_{x^*}^2 + \sigma_{u^*}^2 \tag{13}$$
$$m_{xx} = \sigma_{x^*}^2 + \sigma_{v^*}^2 \tag{14}$$
$$m_{yx} = \beta \sigma_{x^*}^2. \tag{15}$$

Several authors, including Lindley (1947) and Kendall and Stuart (1979) acknowledge the non-identification of system (11)-(15) because it admits six parameters $(\mu, \alpha, \beta, \sigma_{u^*}^2, \sigma_{v^*}^2 \sigma_{x^*}^2)$ but only five equations. These authors emphasized the need for additional (user supplied) sample information.

We propose to tackle the estimation problem from a different angle. First, we wish to keep together the structure of the EIV relations (5) and (6) with the second-order moment relations (11)-(15). Second, we directly connect the variances $\sigma^2_{u^*}$ and $\sigma^2_{v^*}$ in (13) and (14) to relations (5) and (6) by means of a neutral but crucial parameterization of the error terms, namely

$$u_i^* = \sigma_u U_i \tag{16}$$
$$v_i^* = \sigma_v V_i \tag{17}$$

where U_i and V_i are standard normal variables. This means that $\sigma^2_{u^*} = \sigma^2_u$ and $\sigma^2_{v^*} = \sigma^2_v$ Furthermore, using (7) in a sample context,

$$\overline{y}^* = \alpha + \beta \overline{x}^* \tag{18}$$

where $\overline{y}^*$ and $\overline{x}^*$ are the sample means of the latent variables Y^* and X^* whose realizations are stated as in (2) and (3). Thus, the array of estimable relations of the simplest EIV model is assembled as follows

$$y_i = \alpha + \beta x_i^* + \sigma_u U_i \tag{19}$$
$$x_i = x_i^* + \sigma_v V_i \tag{20}$$

$$\overline{y}^* = \alpha + \beta \overline{x}^* \tag{21}$$
$$m_{yy} = \beta^2 \sigma^2_{x^*} + \sigma^2_u \tag{22}$$
$$m_{xx} = \sigma^2_{x^*} + \sigma^2_v \tag{23}$$
$$m_{yx} = \beta \sigma^2_{x^*} \tag{24}$$

$$\sum_{i=1}^N U_i / N = 0 \tag{25}$$
$$\sum_{i=1}^N U_i^2 / N = 1 \tag{26}$$
$$\sum_{i=1}^N V_i / N = 0 \tag{27}$$
$$\sum_{i=1}^N V_i^2 / N = 1 \tag{28}$$

with $\sigma_u > 0, \sigma_v > 0, \sigma_{x^*} > 0$.

This specification of the EIV model is akin to the specification of Kiefer and Wolfowitz (1956) who derived consistent ML estimates "in the presence of infinitely many incidental parameters (latent variables)."

We, however, will not maximize a likelihood function. The system of relations (19)-(28) does not have a closed form solution. In principle, it can have a solution where all the parameters have admissible values, with $\sigma_u > 0, \sigma_v > 0, \sigma_{x^*} > 0$. In other words, it is necessary to find an interior solution of the parameter space. Any boundary solution is not admissible. It is desirable, therefore, to find an interior solution that optimizes some robust statistical function. For this task, we choose the dual objective function of the least-squares (LS) method described by Paris (2011, p. 70). The reason for this choice resides in the global maximum of the dual LS specification, the numerical stability of the optimization problem and the intuitive meaning of the dual LS approach.

Using the terminology of information theory, the dual of the least-squares method corresponds to the maximization of the net value of sample information (*NVSI*). By applying this criterion to the EIV model described above, the structure of the objective function to be maximized subject to relations (19)-(28) turns out as

$$\begin{aligned} \max NVSI = {} & \sigma_u \sum_{i=1}^{N} y_i U_i / N + \sigma_v \sum_{i=1}^{N} x_i V_i / N \\ & - \sum_{i=1}^{N} [y_i - \alpha - \beta x_i^*]^2 / 2N - \sigma_v^2 \sum_{i=1}^{N} V_i^2 / 2N. \end{aligned} \tag{29}$$

It must be emphasized that the estimator developed in this section provides estimates not only of parameters α and β defining the linear regression but also of the error variances σ_u^2 and σ_v^2. And, finally, it provides estimates of the mean and variance of the latent variable X^*, μ_{x^*} and $\sigma^2_{x^*}$. Actually, with the estimate of the N values $\hat{x}_i^*$ of the latent variable it is possible to approximate rather well its entire distribution. This result is a remarkable byproduct of the original goal that was defined simply as the estimation of α and β. This general result depends crucially upon the use of all the sample information, including the sample latent variable (estimated) realizations that were overlooked in previous works of the EIV problem.

Consistency of the dual LS estimator. At present, consistency of the dual LS estimator specified by the maximization of (29) subject to relations (19)-(28) is formulated as a conjecture. Consistency of the EIV parameter estimates could be proved by using the assumptions of Kiefer and Wolfowitz (1956), but their implementation is left for another day. In this paper, we claim that if it is possible to obtain a global maximum of (29) and all the parameter estimates have admissible values (in particular, if all the variances have positive values) the estimator is consistent. In a least-squares context, if there is an interior feasible solution, that solution is a global optimal solution. For the time being, Monte Carlo experiments may either support or negate the conjecture.

3. The Dual of the Least-Squares Estimator

Given the novelty of the dual LS estimator, we give a brief outline of its structure and meaning using the familiar setup of a multiple regression model in vector and matrix notation. In this section, the mathematical symbols are completely unrelated to the EIV model discussed above. The traditional (primal) LS approach consists of minimizing the squared deviations from an average relation of, say, a linear model that consists of three parts:

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{u} \tag{30}$$

where $\mathbf{y}$ is an (*n x 1*) vector of sample observations, $\mathbf{X}$ is an (*n x k*) matrix of predetermined values, $\boldsymbol{\beta}$ is a (*k x 1*) vector of unknown parameters to be estimated by the LS method, and $\mathbf{u}$ is an (*n x 1*) vector of deviations from the quantity $\mathbf{X\beta}$. In the terminology of information theory, relation (30) may be regarded as representing the decomposition of a message into signal and noise, that is, **message = signal + noise**, with obvious correspondences with the three components of (30). Symbolically, then, the LS methodology minimizes the squared deviations (noise) subject to the model's specification

Primal
$$\min_{\mathbf{u},\boldsymbol{\beta}} LS = \mathbf{u'u}/2 \tag{31}$$

$$\text{subject to} \qquad \mathbf{y} = \mathbf{X\beta} + \mathbf{u}. \tag{32}$$

The dual of the LS method is derived using the Lagrangian function and the corresponding first order necessary conditions

$$L(\mathbf{u},\boldsymbol{\beta},\boldsymbol{\lambda}) = \mathbf{u'u}/2 + \boldsymbol{\lambda}'(\mathbf{y} - \mathbf{X\beta} - \mathbf{u}) \tag{33}$$

$$\frac{\partial L}{\partial \mathbf{u}} = \mathbf{u} - \boldsymbol{\lambda} = \mathbf{0} \tag{34}$$

$$\frac{\partial L}{\partial \boldsymbol{\beta}} = -\mathbf{X}'\boldsymbol{\lambda} = \mathbf{0}. \tag{35}$$

Using $\boldsymbol{\lambda} = \mathbf{u}$, $\mathbf{u}'\boldsymbol{\lambda} = \mathbf{u'u}$ and (35) into the Lagrangian function, the dual specification of the least-squares method results in

Dual
$$\max_{\mathbf{u}} NVSI = \mathbf{y'u} - \mathbf{u'u}/2 \tag{36}$$

$$\text{subject to} \qquad \mathbf{X'u} = \mathbf{0}. \tag{37}$$

In the dual specification, the values of the $\boldsymbol{\beta}$ parameters are obtained as Lagrange multipliers of the orthogonal constraints (37). The *NVSI* interpretation stems from the fact that in the LS environment the vector $\mathbf{u}$ acquires a double role and meaning: as a vector of deviations (noise) in the primal and as a vector of "marginal sacrifices" (that is, "prices") in the dual since $\mathbf{u} = \boldsymbol{\lambda}$. The dual LS objective function (36) has a global maximum over the parameter space $(\mathbf{u}, \boldsymbol{\beta})$. In section 2, the objective function (29) of the EIV dual LS estimator has the structure of relation (36).

4. EIV Model With Three Latent Variables

When three latent variables (one dependent and two "explanatory" latent variables) enter the EIV model, the dual LS estimator takes on the following structure:

$$\max NVSI = \sigma_u \sum_{i=1}^{N} y_i U_i / N + \sigma_{v_1} \sum_{i=1}^{N} x_{1i} V_{1i} / N + \sigma_{v_2} \sum_{i=1}^{N} x_{2i} V_{2i} / N \tag{38}$$
$$- \sum_{i=1}^{N} [y_i - \alpha - \beta_1 x_{1i}^* - \beta_1 x_{1i}^*]^2 / 2N - \sigma_{v_1}^2 \sum_{i=1}^{N} V_{1i}^2 / 2N - \sigma_{v_2}^2 \sum_{i=1}^{N} V_{2i}^2 / 2N$$

subject to

$$y_i = \alpha + \beta_1 x_{1i}^* + \beta_2 x_{2i}^* + \sigma_u U_i \tag{39}$$
$$x_{1i} = x_{1i}^* + \sigma_{v_1} V_{1i} \tag{40}$$
$$x_{2i} = x_{2i}^* + \sigma_{v_2} V_{2i} \tag{41}$$

$$\overline{y}^* = \alpha + \beta_1 \overline{x}_1^* + \beta_2 \overline{x}_2^* \tag{42}$$
$$m_{yy} = \beta_1^2 \sigma_{x_1^*}^2 + \beta_2^2 \sigma_{x_2^*}^2 + 2\beta_1\beta_2 \sigma_{x_1^* x_2^*} + \sigma_u^2 \tag{43}$$
$$m_{x_1x_1} = \sigma_{x_1^*}^2 + \sigma_{v_1}^2 \tag{44}$$
$$m_{x_2x_2} = \sigma_{x_2^*}^2 + \sigma_{v_2}^2 \tag{45}$$
$$m_{yx_1} = \beta_1 \sigma_{x_1^*}^2 + \beta_2 \sigma_{x_1^* x_2^*} \tag{46}$$
$$m_{yx_2} = \beta_1 \sigma_{x_1^* x_2^*} + \beta_2 \sigma_{x_2^*}^2 \tag{47}$$
$$m_{x_1x_2} = \sigma_{x_1^* x_2^*} \tag{48}$$

$$\sum_{i=1}^{N} U_i / N = 0 \tag{49}$$

$$\sum_{i=1}^{N} U_i^2 / N = 1 \tag{50}$$

$$\sum_{i=1}^{N} V_{1i} / N = 0 \tag{51}$$

$$\sum_{i=1}^{N} V_{1i}^2 / N = 1 \tag{52}$$

$$\sum_{i=1}^{N} V_{2i} / N = 0 \tag{53}$$

$$\sum_{i=1}^{N} V_{2i}^2 / N = 1 \tag{54}$$

where $\sigma_{x_1^* x_2^*} = \text{cov}(x_1^*, x_2^*)$. The same assumptions stated in (4) are extended to this case.

The extension to four or more latent variables is straightforward and is given in section 7.

The EIV model of this section exhibits ten parameters that must be estimated,

$(\alpha, \beta_1, \beta_2, \sigma_u^2, \sigma_{v_1}^2, \sigma_{v_2}^2, \mu_{x_1^*}, \mu_{x_2^*}, \sigma_{x_1^*}^2, \sigma_{x_2^*}^2)$.

5. Monte Carlo Experiment With Two Latent Variables

The software GAMS (1988) was used to implement the dual least-squares estimator of section 2. The following true values of the parameters and random variables were chosen for this example:

$$\begin{aligned}
&\alpha = -2.0 \\
&\beta = 0.85 \\
&\sigma_u = 1.2 \\
&\sigma_v = 1.4 \\
&U \sim Normal(0, 1.2) \\
&V \sim Normal(0, 1.4) \\
&x_i^* \sim Uniform(3, 9)
\end{aligned}$$

The "explanatory" latent variable, therefore, is distributed with mean $\mu_{x^*} = 6.0$ and variance $\sigma_{x^*}^2 = 3.0$.

The example has six parameters to be estimated. The implementation of a Monte Carlo experiment depends crucially upon the seed of the GAMS program to start the pseudo random-number generator. For this reason, we repeated 20 times the estimation with different seeds and computed the average values of the parameters at each level of the number of observations. The results are reported in Table 1.

Table 1. Monte Carlo Results of the EIV Estimator with Two Latent Variables

N	$\alpha = -2.0$	$\beta = 0.85$	$\sigma_u = 1.2$	$\sigma_v = 1.4$	$\mu_{x^*} = 6.0$	$\sigma^2_{x^*} = 3.0$
50	-1.8295	0.8242	1.1580	1.4097	5.9790	3.0967
100	-2.0384	0.8632	1.1743	1.3650	5.9730	3.0345
200	-2.0294	0.8541	1.1978	1.4021	6.0119	2.9156
500	-2.0138	0.8512	1.1839	1.3937	5.9993	3.0632
1,000	-2.0616	0.8600	1.2001	1.3960	5.9958	2.9594
2,000	-2.0368	0.8555	1.1903	1.3926	5.9911	2.9850
5,000	-1.9851	0.8463	1.1947	1.4117	6.0013	2.9887
10,000	-2.0168	0.8541	1.1998	1.4050	5.9970	2.9748

The results of Table 1 were obtained using the solver Conopt3. It is important to explore the parameter space around the possible optimal solution by repeating the computations with different initial points. It seems safe to say that the results of Table 1 do not negate the conjecture that the dual LS estimator presented here produces consistent estimates of the EIV model's parameters.

6. Monte Carlo Experiment With Three Latent Variables

The following true values of the parameters and random variables were chosen for this example:

$$\alpha = 2.5$$
$$\beta_1 = 0.85$$
$$\beta_2 = -0.5$$
$$\sigma_u = 1.2$$
$$\sigma_{v1} = 0.8$$
$$\sigma_{v2} = 1.4$$
$$U \sim Normal(0,1.2)$$
$$V_1 \sim Normal(0,0.8)$$
$$V_2 \sim Normal(0,1.4)$$
$$x_{1i}^* \sim Uniform(3,9)$$
$$x_{2i}^* \sim Uniform(5,15)$$

The "explanatory" latent variables, therefore, are distributed with mean $\mu_{x_1^*} = 6.0$, $\mu_{x_2^*} = 10.0$ and variance $\sigma^2_{x_1^*} = 3.0$, $\sigma^2_{x_2^*} = 8.3333$, respectively. The results are presented in Table 2.

Table 2. Monte Carlo results of EIV model with three latent variables

N	$\alpha =$ 2.5	$\beta_1 =$ 0.85	$\beta_2 =$ –0.5	$\sigma_u =$ 1.2	$\sigma_{v1} =$ 0.8	$\sigma_{v2} =$ 1.4	$\mu_{x_1^*} =$ 6.0	$\sigma^2_{x_1^*} =$ 3.0	$\mu_{x_2^*} =$ 10.0	$\sigma^2_{x_2^*} =$ 8.333
50	2.6104	0.8101	-0.4910	1.1210	0.7590	1.3673	6.0813	2.9495	10.1299	8.4228
100	2.3884	0.8602	-0.4975	1.1987	0.7654	1.3906	6.0493	2.8519	10.0097	8.3299
200	2.4736	0.8390	-0.4910	1.1935	0.7864	1.4167	6.0261	2.9370	9.9816	8.3605
500	2.5869	0.8446	-0.5063	1.1929	0.7856	1.3836	6.0325	3.0194	10.0663	8.1560
1,000	2.5450	0.8468	-0.5002	1.1997	0.7863	1.4092	6.0108	3.0013	10.0212	8.3982
2,000	2.5560	0.8438	-0.5026	1.1977	0.7875	1.3991	6.0192	3.0282	10.0039	8.3164
5,000	2.5937	0.8419	-0.5047	1.2061	0.7917	1.3996	6.0082	3.0099	10.0081	8.2647
10,000	2.5356	0.8447	-0.5005	1.2034	0.7899	1.4053	6.0092	3.0118	9.9987	8.3048

Also in this case, all the estimates tend toward the true values of their respective parameters rather quickly. It is interesting to note that also at the smaller sample sizes the "precision" of the estimated coefficients is rather surprising and, overall, satisfactory. A guideline has emerged to evaluate the numerical results of any empirical run: First, check whether all the estimated variances (standard deviations) are positive. If not, change the initial point. Second, record the value of the objective function. Third, rerun the program with different initial points and compare the results to the previous runs. Choose the solution (with positive values of the variances) that corresponds to the maximum value of the objective function. The GAMS software contains a solver (Baron) that computes a global optimum, if it exists, and can be used to verify whether such an optimum is associated with positive values of the variances and corresponds to the dual LS solution.

7. Generalized EIV Dual Least-Squares Estimator In Matrix Notation

There are N sample observations and K "explanatory" latent variables and all the assumptions stated in (4) apply. Define the following vectors and matrices:

$\mathbf{y}' = [y_1, y_2, ...y_N] \equiv$ vector of sample observations of dependent variable
$\mathbf{y}^{*\prime} = [y_1^*, y_2^*, ...y_N^*] \equiv$ vector of sample dependent latent variables
$\mathbf{s}' = [1,1,...,1] \equiv$ sum vector
$\boldsymbol{\beta}' = [\beta_1, \beta_2, ..., \beta_K] \equiv$ vector of slope coefficients
$\mathbf{X}^* = [\mathbf{x}_1^*, \mathbf{x}_2^*, ..., \mathbf{x}_K^*] \equiv$ matrix of latent variables
$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_K] \equiv$ matrix of sample observations of latent variables
$\mathbf{u}' = [u_1, u_2, ..u_N] \equiv$ vector of errors of dependent variable
$\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_K] \equiv$ matrix of error of latent variables

$$\ddot{\mathbf{V}} = \begin{bmatrix} \mathbf{v}_1 & & \\ & \mathbf{v}_2 & \\ & & \mathbf{v}_K \end{bmatrix} = \text{diagonal error matrix}$$

$$\ddot{\mathbf{X}}^* = \begin{bmatrix} \mathbf{x}_1^* & & \\ & \mathbf{x}_2^* & \\ & & \mathbf{x}_K^* \end{bmatrix} \equiv \text{ diagonal matrix of latent variables}$$

$$\overleftrightarrow{\mathbf{X}} = \begin{bmatrix} \mathbf{x}_1 & & \\ & \mathbf{x}_2 & \\ & & \mathbf{x}_K \end{bmatrix} = \text{diagonal matrix of observed latent variables}$$

$$\Sigma_{\mathbf{x}^*} = \begin{bmatrix} \sigma^2_{x_1^*} & \sigma^2_{x_1^* x_2^*} & \sigma^2_{x_1^* x_K^*} \\ \sigma^2_{x_2^* x_1^*} & \sigma^2_{x_2^*} & \sigma^2_{x_2^* x_K^*} \\ \sigma^2_{x_K^* x_1^*} & \sigma^2_{x_K^* x_2^*} & \sigma^2_{x_K^*} \end{bmatrix} = \text{variance/covariance matrix of latent variables}$$

$$\mathbf{M}_{\mathbf{x},\mathbf{x}} = \begin{bmatrix} m_{x_1 x_1} & m_{x_1 x_2} & m_{x_1 x_3} \\ m_{x_2 x_1} & m_{x_2 x_2} & m_{x_2 x_K} \\ m_{x_K x_1} & m_{x_K x_2} & m_{x_K x_K} \end{bmatrix} \equiv \text{ matrix of sample moments}$$

$$\Sigma_{\sigma_v} = \begin{bmatrix} \sigma_{v_1} & & \\ & \sigma_{v_2} & \\ & & \sigma_{v_K} \end{bmatrix} \equiv \text{ diagonal matrix of error standard deviations}$$

$$\Sigma^2_{\sigma_v} = \begin{bmatrix} \sigma^2_{v_1} & & \\ & \sigma^2_{v_2} & \\ & & \sigma^2_{v_K} \end{bmatrix} \equiv \text{ diagonal matrix of error variances}$$

$\mathbf{m}_{\mathbf{yX}} \equiv$ vector of sample moment between $\mathbf{y}$ and $\mathbf{X}$

Then, the general EIV dual least-squares estimator of a multiple regression assumes the following structure:

$$\max NVSI = [\sigma_u \mathbf{y}'\mathbf{u} + \mathbf{s}'\overleftrightarrow{\mathbf{X}}\boldsymbol{\Sigma}_{\sigma_v}\overleftrightarrow{\mathbf{V}}\mathbf{s}]/N \qquad (55)$$
$$-[\mathbf{y} - \alpha\mathbf{s} - \mathbf{X}^*\boldsymbol{\beta}]'[\mathbf{y} - \alpha\mathbf{s} - \mathbf{X}^*\boldsymbol{\beta}]/2N - \mathbf{s}'\overleftrightarrow{\mathbf{V}}\boldsymbol{\Sigma}^2_{\sigma_v}\overleftrightarrow{\mathbf{V}}\mathbf{s}/2N$$

subject to

$$\mathbf{y} = \alpha\mathbf{s} + \mathbf{X}^*\boldsymbol{\beta} + \sigma_u\mathbf{u} \qquad (56)$$
$$\mathbf{X} = \mathbf{X}^* + \boldsymbol{\Sigma}_{\sigma_v}\mathbf{V} \qquad (57)$$

$$\bar{y}^* = \alpha + \bar{\mathbf{x}}^*\boldsymbol{\beta} \tag{58}$$

$$m_{yy} = \boldsymbol{\beta}'\boldsymbol{\Sigma}_{\mathbf{x}^*}\boldsymbol{\beta} + \sigma_u^2 \tag{59}$$

$$\mathbf{M}_{\mathbf{xx}} = \boldsymbol{\Sigma}_{\mathbf{x}^*} + \boldsymbol{\Sigma}_{\sigma_v}^2 \tag{60}$$

$$\mathbf{m}_{y\mathbf{x}} = \boldsymbol{\Sigma}_{\mathbf{x}^*}\boldsymbol{\beta} \tag{61}$$

$$\mathbf{u}'\mathbf{s} / N = 0 \tag{62}$$

$$\mathbf{u}'\mathbf{u} / N = 1 \tag{63}$$

$$\overset{\leftrightarrow}{\mathbf{V}}'\mathbf{s} / N = \mathbf{0}_K \tag{64}$$

$$\overset{\leftrightarrow}{\mathbf{V}}'\overset{\leftrightarrow}{\mathbf{V}} / N = \mathbf{I}_K\,. \tag{65}$$

8. Conclusion

This paper presents an estimator of the EIV structural model using only first and second-order sample moments. The specification of the EIV model follows the lines discussed by Kiefer and Wolfowitz (1956) who derived a consistent estimator of the EIV parameters in the presence of infinitely many "incidental parameters" (latent variables). In order to avoid the difficulties of the ML method in this context (identifiability and estimation problems) we have chosen to optimize a dual least-squares objective function subject to a list of general relations that has been associated for a long time with the EIV problem. This specification – in general – has an interior global maximum. A crucial but neutral novelty of this treatment concerns the parameterization of the error terms of the dependent and "explanatory" latent variables. This "change of variable" allows the introduction of standard deviations of the error terms in all the observations. A direct link is thus established between the second-order moments involving error variances, the detailed EIV observation relations involving the error standard deviations, and the objective function. This intertwining of the moments and variance parameters, without requiring additional (user supplied) information about them, is a fundamental breakthrough in order to achieve the desired goal. In the Monte Carlo experiments, which strictly reflect the structure of the EIV problem presented in sections 2 and 4, it is known that a region of the parameter space exists where all the variances (standard deviations) are positive. If the chosen objective function achieves a global maximum in this region, the EIV parameter estimates are consistent.

We believe that the goal set out at the beginning of this paper – finding an easy method for consistent parameter estimates of the EIV model – has been attained. This method requires only a working knowledge of the least-squares methodology, including its dual specification, and the knowledge of the variance/covariance law for random variables. The estimation procedure presented here includes the naïve least-squares method as a special case, that is, when the "explanatory" variables are measured without error.

References

Adcock, R.J. (1878). A Problem in Least Squares. *The Analyst*, 5, 53-54.

Brooke, A., Kendrick, D. Meeraus, A. (1988). *GAMS, User Guide*, Boyd and Fraser Publishing Company. Danvers, MA.

Cragg, J.G. (1997). Using Higher Moments to Estimate the Simple Errors-In-Variables Model, *The RAND Journal of Economics*, 28, S71-S91.

Dagenais, M.G. and Dagenais, D.L. (1997). Higher Moment Estimators for the Linear Regression Models With Errors in the Variables, Journal of Econometrics, 76, 193-221.

Delaigle, A. and A. Meister (2007). Nonparametric Regression Estimation in the Heteroscedastic Errors-In-Variables Problem, *Journal of the American Statistical Association*, 102, 1416-1426.

Gillard, J. (2010). An Overview of Linear Structural Models in Errors In Variables Regressions, *REVSTAT – Statistical Journal*, 8, 57-80.

Kendall, M. and A. Stuart (1979). *The Advanced Theory of Statistics*. Vol. 2, Fourth Edition. Charles Griffin & Company Limited, London & High Wycombe.

Kiefer, J. and J. Wolfowitz (1956). Consistency of the Maximum Likelihood Estimator in the Presence of Infinitely Many Incidental Parameters. *The Annals of Mathematical Statistics*, 27, 887-906.

Lindley, D.V. (1947). Regression Lines and the Linear Functional Relationship, Suppl. *Journal of the Royal Statistical Society*, 9, 218-244.

Neyman, J. (1951). Existence of Consistent Estimates of the Directional Parameters in a Linear Structural Relation Between Two Variables, *Ann. Math. Stat.*, 22, 497-512.

Pal, M. (1980). Consistent Moment Estimators of Regression Coefficients in the Presence of Errors in Variables, Journal of Econometrics, 14, 349-364.

Paris, Q. (2011). *Economic Foundations of Symmetric Programming*, Cambridge University Press, New York.

Reiersøl, O., 1950. Identifiability of a Linear Relation Between Variables Which Are Subject to Error. *Econometrica*, 18, 375-389.

Van Monfort, K., Mooijaart, A. and De Leeuw, J. (1987). Regression with Errors In

Variables: Estimators Based on Third Order Moments, Statistica Neerlandica, 41, 223-237.

Wolfowitz, J. (1954). Estimation of the Components of Stochastic Structures, *Proc. Nat. Acad. Sci., U.S.A.* 40, 602-606.