Survey Methodology, June 1991
Vol. 17, No. 1, pp. 91–98
Statistics Canada

Estimating a System of Linear Equations with Survey Data

PHILLIP S. KOTT[1]

ABSTRACT

This paper develops a framework for estimating a system of linear equations with survey data. Pure design-based sample survey theory makes little sense in this context, but some of the techniques developed under this theory can be incorporated into robust model-based estimation strategies. Variance estimators with the form of the single equation "linearization" estimator are nearly unbiased under many complex error structures. Moreover, the inclusion of sampling weights in regression estimation can protect against the possibility of missing regressors. In some situations, however, the existence of missing regressors can make the estimation of a system of equations ambiguous.

KEY WORDS: Sampling weights; Putative missing regressor; Robust; Nearly unbiased.

1. INTRODUCTION

Kott (1991) showed that design-based techniques developed for estimating a single linear regression equation could be exploited in a more conventional model-based framework. In particular the use of sample weighted regression was shown to help protect against the possible existence of missing regressors, while the so-called linearization variance estimator was shown to produce nearly unbiased estimators of mean squared error for many complex variance structures.

This paper extends those results to the estimation of a system or "grouping" of linear equations, a topic of considerable interest to econometricians (see, for example, Johnston 1972, pp. 238-241). Two simple examples may shed some light onto the subject for those not already schooled in econometric methods or their equivalent.

Suppose we have a sample of farmers and want to estimate the relationship between the amount of planted soybean acres and the size of the farm. Zellner (1962) showed in effect that even if a simple quadratic equation with independent and identically distributed errors correctly described the universe, a better estimator than the one produced by ordinary least squares (OLS) might exist. This estimator could be found by taking into consideration other linear relationships, say between planted *corn* acres and farm size, that had errors terms correlated with those in the original relationship. Zellner called the system-wide estimation of a group of such equations "seemingly unrelated regression." Oddly, in order for Zellner's generalized least squares (GLS) estimator to produce different results from OLS, it is necessary for some equations to contain regressors not found in other equations. Alternatively, one can think of each equation as containing the same regressors but with certain coefficients constrained to zero.

A second example concerns a sample of firms each producing one output, y, from two inputs, x_1 and x_2, with unit prices, p_1 and p_2. Economists often assume that each firm possesses the same technology (plus or minus an error term). Given p_1, p_2, and y, each firm would choose x_1 and x_2 so as to minimize total cost, $c = p_1x_1 + p_2x_2$. Suppose that the

[1] Phillip S. Kott, Special Assistant for Economic Survey Methods, U.S. Bureau of the Census, Room 3061-3, Washington, DC, 20233, USA.

relation between p_1, p_2, y and the cost minimizing c can be expressed by the following equation (on average):

$$\log(c) = b_0 + b_1\log(p_1) + b_2\log(p_2) + b_3\log(y). \tag{1}$$

Economic theory tells us that a rational firm faced with implicit cost equation (1) would choose its level of x_1 so that

$$x_1 p_1 / c = b_1. \tag{2}$$

Naturally, in order to estimate equations (1) and (2), we need to add a stochastic structure. For simplicity, assume that both equations (1) and (2) fit the behavior of all firms subject to respective independent (across firms) and identically distributed random errors. Observe that in addition to the strong possibility that the error terms in the two equation will be correlated for a particular firm, there is also a coefficient (b_1) shared by both equations.

When faced with a system of linear equations in which the coefficients are known to be constrained, the design-based approaches to linear regression reviewed in Kott (1990a) make little sense. For that reason, although design-based *practice* inspires many of the procedures discussed here, only the extended model-based approach introduced in Kott (1991) will be used to justify them.

Section 2 lays out the theoretical model for the estimation of a system of linear equations based on data from the full population. Section 3 introduces the sample weighted analogues of full population OLS and GLS estimators for a system of linear equations. Section 4 addresses robust mean squared error estimation of both the sample weighted OLS and GLS estimators employing a straigtforward generalization of the linearization variance estimator (see, for example, Shah, Holt, and Folsom 1977). Section 5 discusses a general method for developing test statistics that can be used to evaluate, among other things, whether sample weighted OLS and GLS are actually estimating the same thing. Section 6 explores a simple example. Section 7 sketches an extension of the methodology developed here to what econometricians call "simultaneous equations." In the stochastic version of equation (1), for example, many economists believe that $\log(y)$ should be treated as a random variable and that $\log(c)$ can be assumed to be fixed. This causes a simultaneity bias if not specifically addressed by techniques like two and three stage least squares (see Johnston 1972, pp. 341-420). Finally, section 8 contains a brief discussion.

2. FULL POPULATION ESTIMATION

2.1 The Unconstrained System:

Suppose we have a population containing M data points. Each data point i is associated with $G + \tilde{K}$ observed variables satisfying the following model:

$$Y = \tilde{X}\tilde{\beta} + U + V, \tag{3}$$

where Y is an $M \times G$ matrix of observed dependent variables (the ith row of Y contains the dependent variables associated with the ith data point),

$\tilde{X}$ is an $M \times \tilde{K}$ matrix of observed independent or regressor variables (the ith row of $\tilde{X}$ contains the independent variables associated with the ith data point),

$\tilde{\beta}$ is an $\tilde{K} \times G$ matrix of parameters,

U is an $M \times G$ matrix satisfying the relationship $\lim_{M\to\infty} \tilde{X}'U/M = 0_{\tilde{K}\times G}$ (a matrix of zeroes) - this assumes that there is an underlying process generating the data points which could in principle generate points *ad infinitum* (see Kott 1991), and

V is a $M \times G$ matrix of random variables such that $E(V) = 0_{M\times G}$ and $E(v_{is}v_{it}) = \sigma_{st(i)}$.

It is well known that if $U \equiv 0_{M\times G}$, $E(v_{is}v_{jt}) = 0$ for $i \neq j$, and $\sigma_{st(i)} = \sigma_{st}$ for all i, then

$$\tilde{B}_{\text{OLS}} = (\tilde{X}'\tilde{X})^{-1}\tilde{X}'Y \tag{4}$$

is the best linear unbiased estimator for $\tilde{\beta}$ (see, for example, Johnston 1972, p. 240). This means that the gth column of $\tilde{B}_{\text{OLS}}$, call it $B_{\cdot g}$, is the best linear unbiased estimator of $\beta_{\cdot g}$, where

$$y_{\cdot g} = \tilde{X}\beta_{\cdot g} + u_{\cdot g} + v_{\cdot g}, \tag{5}$$

and $y_{\cdot g}$, $u_{\cdot g}$, and $v_{\cdot g}$ are the gth columns of Y, U, and V, respectively. Equation (5) can be viewed as the gth equation in the system of equations represented by equation (3).

Let us call the matrix U in equation (3) the *putative missing regressor* matrix. Usually, in conventional (*i.e.*, model-based) regression analysis that part of the dependent variable (or variables) not capturable by a linear combination of the independent variables is (are) assumed to be purely random. Here, however, we follow Kott (1991) and allow for the possiblility of non-random missing regressors. Note that even when $U \neq 0_{M\times G}$, $\tilde{B}_{\text{OLS}}$ remains nearly (*i.e.*, asymptotically) unbiased.

2.2 The (Possibly) Constrained System:

Efficient estimation is a more complicated matter when there are constraints on some elements of $\tilde{\beta}$; for example, when $\tilde{\beta}_{kg}$ is known to be zero or when $\tilde{\beta}_{hf}$ is known to equal $\tilde{\beta}_{hj}$.

In this paper, we are interested in a (possibly) constrained systems of equations that can be modelled directly with the following equation:

$$y = X\beta + u + v, \tag{6}$$

where $y = (y_{\cdot 1}', y_{\cdot 2}', \ldots, y_{\cdot G}')'$, u and v are defined in an analogous manner, X is an $MG \times K$ matrix, β is a $K \times 1$ vector, and $K \leq G\tilde{K}$. By definition, $\lim_{M\to\infty} X'u/M = 0_K$.

When the original $\tilde{\beta}$ in equation (3) is unconstrained, $K = G\tilde{K}$, and

$$X = \begin{bmatrix} \tilde{X} & & & \\ & \tilde{X} & & \\ & & \ddots & \\ & & & \tilde{X} \end{bmatrix}.$$

When the original $\tilde{\beta}$ is constrained, however, $K < G\tilde{K}$. For example, when an element of $\tilde{\beta}$ is known to be zero, it can be removed from the β vector in equation (6) along with the column of the X matrix that corresponds to it. When two elements in the same row of $\tilde{\beta}$ are known to be equal, the second can be removed from β, and X can be adjusted accordingly (it will no longer be block diagonal).

When $u \equiv 0_{MG}$ and $\text{Var}(v) = \Sigma \otimes I_M$ (where $\Sigma = \{\sigma_{st}\}$), then $b_{\text{OLS}} = (X'X)^{-1}X'y$ is an unbiased estimator for β, but $b_{\text{GLS}} = (X'[\Sigma^{-1} \otimes I_M]X)^{-1}X'[\Sigma^{-1} \otimes I_M]y$, where I_M is the $M \times M$ identity matrix, is the best linear unbiased estimator. In practice, the elements of Σ have to be estimated from the sample, say by $\hat{\sigma}_{gf} = r_{.g}'r_{.f}/M$, where $r_{.g} = y_{.g} - X_{.g}b_{\text{OLS}}$.

It is well known that b_{OLS} and b_{GLS} are equal when the parameter matrix in (3) is unconstrained (again, see Johnston 1972, p. 240). Turning to the constrained case, if $u \neq 0_{\text{MG}}$, then both b_{OLS} and b_{GLS} are nearly unbiased estimators of β when $\lim_{M\to\infty}\tilde{X}'U/M = 0_{\tilde{K}\times G}$ holds as we originally assumed. Unfortunately, b_{GLS} may not be nearly unbiased under the weaker assumption that $\lim_{M\to\infty}X'u/M = 0_K$, which is more in line with the extended model in Kott (1991) when (6) is viewed as a single equation.

To see why this is, let $X_{.g}$ denote the $M \times K$ matrix formed from the $\{(g-1)M+1\}$th through the $\{gM\}$th row of X and $\Sigma^{-1} = \{\sigma^{fg}\}$, then

$$E(b_{\text{GLS}} - \beta) \propto X'[\Sigma^{-1} \otimes I_M]u/M =$$

$$\sum_g X_{.g}'\left(\sum_f \sigma^{fg}u_{.f}\right)\Big/M = \sum_g \sum_f \sigma^{fg}X_{.g}'u_{.f}\Big/M,$$

which approaches zero as M grows large under the stronger assumption but not necessarily the weaker one.

3. ESTIMATION WITH SURVEY DATA

Suppose now that we observe variables values for only a random sample of the population. Let $P = \text{diag}\{p_i\}$, where p_i is the probability of selection for data point i. Let $S = \text{diag}\{s_i\}$, where $s_i = 1$ if data point i is in the sample and 0 otherwise. Finally, let $W = (m/M)P^{-1}S = \text{diag}\{w_i\}$ be the matrix of sampling weights, where m is the sample size. When all the $p_i = m/M$, note that $W = S$.

It is not difficult to show that for many sample designs and populations (see Kott 1990b and 1991), the *sample weighted OLS estimator*:

$$\hat{\beta}_{W\cdot\text{OLS}} = (X'[I_G \otimes W]X)^{-1}X'[I_G \otimes W]y \tag{7}$$

is a design consistent estimator for b_{OLS}, which means that $\text{plim}_{m\to\infty}(\hat{\beta}_{W\cdot\text{OLS}} - b_{\text{OLS}}) = 0_K$. Under similar conditions, *sample weighted GLS estimator*:

$$\begin{aligned}\hat{\beta}_{W\cdot\text{GLS}} &= (X'[I_G \otimes W][\hat{\Sigma}^{-1} \otimes I_M]X)^{-1}X'[I_G \otimes W][\hat{\Sigma}^{-1} \otimes I_M]y \\ &= (X'[\hat{\Sigma}^{-1} \otimes W]X)^{-1}X'[\hat{\Sigma}^{-1} \otimes W]y,\end{aligned} \tag{8}$$

where

$$\hat{\sigma}_{gf} = r_{.g}'Wr_{.f}\Big/\sum_{i=1}^{M} w_i, \quad \text{and} \quad r = y - X\hat{\beta}_{W\cdot\text{OLS}},$$

is a design consistent estimator for b_{GLS}. Like b_{OLS} and b_{GLS} (and for the same reasons), $\hat{\beta}_{W\cdot\text{OLS}}$ and $\hat{\beta}_{W\cdot\text{GLS}}$ are equal for an unconstrained system of equations.

If $\hat{\beta}_{W\cdot\mathrm{OLS}}$ and $\hat{\beta}_{W\cdot\mathrm{GLS}}$ are design consistent, both are also nearly unbiased estimators of β when $\lim_{M\to\infty}\tilde{X}'U/M = 0_{\tilde{K}\times G}$, because b_{OLS} and b_{GLS} are; however, $\hat{\beta}_{W\cdot\mathrm{GLS}}$ - like b_{GLS} - may not be nearly unbiased under the weaker assumption that $\lim_{M\to\infty}X'u/M = 0_K$. (Unbiasedness here is always defined with respect to the model in equation (6)).

4. MEAN SQUARED ERROR ESTIMATION

Suppose the sample design is such that there are H strata, n_h distinctly sampled PSU's in stratum h, and m_{hj} sampled data points in PSU hj. Both $\hat{\beta}_{W\cdot\mathrm{OLS}}$ and $\hat{\beta}_{W\cdot\mathrm{GLS}}$ have the form $\hat{\beta} = Cy$. Without loss of generality, they can be rewritten as $\hat{\beta} = C^*y^*$, where $y^* = (y_{11}', \ldots, y_{Hn_H}')$ contains only elements corresponding to sampled data points, and y_{hj} is the vector of $G \times m_{hj}$ y-values associated with data points in PSU hj. Define r^* and r_{hj} in an analogous manner to y^* and y_{hj}.

Let D_{hj} be a diagonal matrix of 0's and 1's such that $D_{hj}y^* = (0', \ldots, y_{hj}', \ldots, 0')$, and let $g_{hj} = C^*D_{hj}r^*$. Extending the design-based linearization variance estimator in a straight forward manner, the estimator for the mean squared error of $\hat{\beta} = C^*y^*$ has the form:

$$\mathrm{mse} = \sum_{h=1}^{H} \frac{n_h}{n_h - 1}\left[\sum_{j=1}^{n_h} g_{hj}g_{hj}' - \frac{1}{n_h}\left(\sum_{j=1}^{n_h} g_{hj}\right)\left(\sum_{j=1}^{n_h} g_{hj}'\right)\right]. \tag{9}$$

Under mild restrictions on the sampling design, mse is nearly unbiased when U (from (3)) $\equiv 0_{M\times G}$ and V obeys the following property:

$$|\,E(v_{sg}v_{tf})\,| \begin{cases} = 0 & \text{when } s \text{ and } t \text{ are from different PSU's} \\ < Q & \text{otherwise.} \end{cases}$$

See Kott (1991) for the proof in the $G = 1$ case; the extension to the $G > 1$ case is trivial. The estimator mse remains reasonable when $U \not\equiv 0_{M\times G}$ (see Kott 1990a).

5. TEST STATISTICS

Let $\hat{\beta}_{I\cdot\mathrm{OLS}}$ and $\hat{\beta}_{I\cdot\mathrm{GLS}}$ be the unweighted counterparts of $\hat{\beta}_{W\cdot\mathrm{OLS}}$ and $\hat{\beta}_{W\cdot\mathrm{GLS}}$ derived by replacing the W in (7) and (8) by S. One is often interested in determining whether using the sampling weights really matters. This comes down to testing whether $\hat{\beta}_{I\cdot\mathrm{OLS}}$ and $\hat{\beta}_{W\cdot\mathrm{OLS}}$ are significantly different; that is, whether they are estimating the same thing.

When weights *are* determined to matter, another question of some interest is whether $\hat{\beta}_{W\cdot\mathrm{OLS}}$ and $\hat{\beta}_{W\cdot\mathrm{GLS}}$ are significantly different; that is, does $\lim_{M\to\infty}\tilde{X}'U/M = 0_{\tilde{K}\times G}$ hold so that these two estimators are estimating the same thing?

A general statistic for testing whether:

$$\hat{\beta}_{(1)} = \sum_h \sum_j \{C_{(1)}{}^*D_{hj}y^*\} \quad \text{and} \quad \hat{\beta}_{(2)} = \sum_h \sum_j \{C_{(2)}{}^*D_{hj}y^*\}$$

are equal is

$$T^2 = [\hat{\beta}_{(1)} - \hat{\beta}_{(2)}]'A^{-1}[\hat{\beta}_{(1)} - \hat{\beta}_{(2)}], \tag{10}$$

where

$$A = \sum_{h=1}^{H} \frac{n_h}{n_h - 1}\left[\sum_{j=1}^{n_h} d_{hj} d_{hj}' - \frac{1}{n_h}\left(\sum_{j=1}^{n_h} d_{hj}\right)\left(\sum_{j=1}^{n_h} d_{hj}'\right)\right],$$

$d_{hj} = C_{(1)}{}^{*} D_{hj} r_{hj(1)} - C_{(2)}{}^{*} D_{hj} r_{hj(2)}$, and $r_{hj(f)} = y_{hj} - X_{hj}\hat{\beta}_{(f)}$.

Under the null hypothesis, the test statistic, T^2, is asymptotically a χ^2 random variate with K degrees of freedom. Given our concern for robustness, it seems prudent to question the null hypothesis when prob$(\chi^2_{(K)} > T^2)$ is at considerably less than the standard 0.1 or 0.05 level, but not when T^2 is less than its expected value, K.

6. AN EXAMPLE

Consider the following example synthesized from data from the National Agricultural Statistics Service's June 1989 Agricultural Survey. The data set, previously analyzed in Kott (1990a), is briefly described below.

A sample of 17 primary sampling units was selected from among 4 strata. These PSU's were then subsampled yielding a total sample of 252 farms. Although the sample was random, not all farms had the same probability of selection.

Suppose we are interested in estimating the parameters, β_1 and β_2, of the following equation:

$$y_{1i} = x_{1i}\beta_1 + x_{2i}\beta_2 + u_{1i} + v_{1i}, \tag{11}$$

where i denotes a farm,

y_{1i} is farm i's planted soybeans to cropland ratio when i's cropland is positive, zero otherwise;

x_{1i} is 1 if farm i has positive cropland, zero otherwise; and

x_{2i} is farm i's cropland divided by 10,000.

(Note: dropping all sampled farms with zero cropland from the regression equation will have no effect on the parameter estimation, but it can affect mean squared error estimation.)

Letting $\hat{\beta}_{(1)}$ in equation (10) be the pure OLS estimator for the vector $(\beta_1, \beta_2)'$, and $\hat{\beta}_{(2)}$ be the sample weighted estimator, one computes a T^2 of 4.58. Under the null hypothesis that OLS and sample weighted least squares are estimating the same thing (for which $u_{1i} \equiv 0$ is sufficient but not necessary), T^2 is asymptotically $\chi^2_{(2)}$. We cannot reject this null hypothesis at the 0.1 level. Nevertheless, since T^2 is considerably greater than 2, it seems that the existence of a putative missing regressor is more than likely. Thus, the sample weighted regression estimator should be employed rather than the OLS estimator.

Table 1 displays both the pure OLS and the sample weighted coefficient estimates. Although the sample weighted estimator for β_2 is not significantly different from zero at the 0.1 level, we retain it in the model because it exceeds its estimated root mean squared error. This parallels the reasoning for preferring sample weighted regression over OLS.

Notice the loss of efficiency that results from using the sample weighted estimator in place of pure OLS. The estimated root mean squared error for the β_2 estimator more than doubles (note: both root mean squared errors were estimated using equation (9)).

Table 1

Alternative Estimates for Equation (11)

OLS	$y_{1i} = \underset{(.044)}{0.268x_{1i}} - \underset{(3.95)}{0.92x_{2i}} + u_{1i} + v_{1i}$
sample weighted	$y_{1i} = \underset{(.075)}{0.191x_{1i}} + \underset{(9.95)}{12.15x_{2i}} + u_{1i} + v_{1i}$
sample weighted GLS	$y_{1i} = \underset{(0.71)}{0.197x_{1i}} + \underset{(6.97)}{10.26x_{2i}} + u_{1i} + v_{1i}$

Numbers in parentheses are root mean squared errors.

We can increase the efficiency of the sample weighted estimator by adding a second farm equation and estimating it and (11) as a system. Let

$$y_{2i} = x_{1i}\beta_3 + u_{2i} + v_{2i}, \tag{12}$$

where y_{2i} is farm i's planted corn to cropland ratio when i's cropland is positive, zero otherwise.

The sample weighted estimators in Table 1 and their estimated root mean squared errors are unchanged under system-wide sample weighted OLS. The system approach, however, allows us to calculate sample weighted GLS estimator for β_1 and β_2 which are also displayed in the table. Observe that the estimated root mean squared error for β_2 is reduced by approximately 30% without a loss of robustness, assuming that sample weighted OLS and GLS are estimating the same thing.

The T^2 value for a test comparing the sample weighted OLS and GLS estimators for the vector $(\beta_1, \beta_2)'$ is 0.97. This number is considerably less than 2. Thus, the two estimators do appear to be estimating the same thing. That is to say, there is no additional regressor in one equation related to the putative missing regressor in the other (which is not surprising since when an x_{2i} term was added to the right hand side of equation (12), its estimated coefficient was less than its estimated root mean squared error).

7. SIMULTANEOUS EQUATIONS

In a simultaneous equation framework, some of the columns of the dependent variable matrix, Y, (see (3)) are actually contained on the right hand side of the gth equation (see (5)). Formally, we can write

$$y = Y_{(\cdot)}\alpha + X\beta + u + v \quad \text{or} \quad y = Z\delta + u + v,$$

where

$$Y_{(\cdot)} = \begin{bmatrix} Y_{(1)} \\ Y_{(2)} \\ \cdot \\ \cdot \\ Y_{(G)} \end{bmatrix},$$

$$Z = (Y_{(\cdot)}X), \quad \text{and} \quad \delta = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}.$$

Most of the columns of $Y_{(g)}$ are 0-vectors. The rest (no more than G-1 columns) are columns of Y from equation (3).

Define $\hat{Y}_{(g)}$ as $\tilde{X}(\tilde{X}'W\tilde{X})^{-1}\tilde{X}'WY_{(g)}$. Now replace X in (5) by $\hat{Z} = (\hat{Y}_{(\cdot)}\ X)$ and proceed as before. Equation (7) produces $\hat{\delta}_{W\cdot\mathrm{OLS}}$, akin to two stage least squares, while (8) produces $\hat{\delta}_{W\cdot\mathrm{GLS}}$, akin to three stage least squares. Mean squared error estimation follows along the same line of reasoning that produced equation (9).

8. DISCUSSION

The purpose of this paper was to show how procedures developed in the design-based survey sampling literature - in particular, sample weighted regression and the linearization mean squared error estimator - could be adopted to the estimation of a system of linear equations.

One somewhat unexpected discovery was when estimating the parameters of a constrained linear system, the sample weighted analogues of OLS and GLS might be estimating different things. On further reflection this is not so suprising. If there are missing regressors in our working model, perhaps we don't always know enough about the true model to put constraints on the parameters in the first place.

It is important to realize that mse in equation (9) can be used to estimate the mean squared error of parameter estimators even when there are no missing regressors. The advantages of mse to conventional practice is that it allows for the possibility of heteroscedasticity and complex correlations across data points (but within PSU's).

If there are no missing regressors, however, the following estimator has all the advantages of mse and is generally more efficient:

$$\mathrm{mse}' = \sum_{h=1}^{H} \sum_{j=1}^{n} \frac{n}{n-1} \{g_{hj} g_{hj}' - gg'/n\}, \tag{13}$$

where $n = \sum n_h$ and $g = \sum \sum g_{hj}$ (note: if $Xq = (1, \ldots, 1)'$ for some K-vector q, then $g = 0$).

When there *are* missing regressors the diagonal elements of mse′ may tend to be biased upward. The reasoning here follows that in Wolter (1985) for collapsed strata variance estimators in design-based sampling theory.

REFERENCES

JOHNSTON, J. (1972). *Econometric Methods*, (second edition). New York: McGraw Hill.

KOTT, P.S. (1991). A model-based look at linear regression with survey data. *American Statistician*, forthcoming.

KOTT, P.S. (1990a). What does performing a linear regression on survey data mean? *Proceedings of the Section on Survey Research Methods, American Statistical Association*, forthcoming.

KOTT, P.S. (1990b). The design consistent regression estimator and its conditional variance. *Journal of Statistical Planning and Inference*, 24, 287-296.

SHAH, B.V., HOLT, M.M., and FOLSOM, R.E. (1977). Inference about regression models from sample survey data. *Bulletin of the International Statistical Institute*, 47, 43-57.

WOLTER, K.M. (1985). *Introduction to Variance Estimation*. New York: Springer-Verlag.

ZELLNER, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association*, 57, 348-368.