

Estudios de Economía Aplicada

Nº 14, 1999. Págs. 37-46

Linear Estimation Under Superpopulation Models: Some Results On Robustness

CASAS SÁNCHEZ, J.M.

Universidad de Alcalá de Henares

GUIJARRO GARVI, M.

Universidad de Cantabria

SUMMARY

The present article examines the linear estimation of a finite population mean, under a regression superpopulation model with unknown parameters and correlated residuals.

Asymptotic design unbiasedness and weak robustness are desirable properties of the estimators when the model is misspecified. In this sense, two procedures of choosing among weakly robust and asymptotically design unbiased linear estimators, are suggested.

Keywords: Superpopulation Model; Linear Estimation; Asymptotic Design Unbiasedness; Weakly Robustness.

AMS subjects classification code: 62D05.

RESUMEN

Este trabajo examina la estimación lineal de la media en una población finita, desde el punto de vista de los modelos de superpoblación con parámetros desconocidos y coeficiente de correlación no nulo.

La insesgadez asintótica según el diseño de muestreo y la robustez débil son propiedades deseables del estimador cuando se cometen errores en la especificación del modelo. En este sentido, se presentan dos procedimientos de selección entre estimadores asintóticamente insesgados respecto al diseño y débilmente robustos.

Palabras clave: Modelo de superpoblación; estimación lineal; p -insesgadez asintótica; robustez débil.

Clasificación AMS: 62D05.

1. INTRODUCTION

The problem of estimating means in finite populations assuming a superpopulation model that relates a variable of interest to one or more explanatory variables, leads to the necessity of looking for robust estimators when the model is misspecified.

Since the model parameters are unknown, it is difficult to find design unbiased estimators of the population mean. In this case, as pointed out by Hansen, Madow and Tepping (1988), asymptotic design unbiasedness is the principal guarantee of robustness.

We will examine the robustness of estimators when the covariance matrix of the superpopulation model is misspecified. Following Tam (1988), an estimator is defined as weakly robust if it is robust against this type of error; we shall use this definition in a correlated residuals context, taking Tam results somewhat further.

We will use the expected mean squared error criterion as an indicator of the quality of an estimator (Godambe, 1955; Särndal, 1980; Tam, 1988).

2. PRELIMINARIES

Let U_t denote a population of units labelled $i = 1, \dots, N_t$ and a sample s_t of fixed effective size n_t . Associated with the i th unit is a fixed, known, $q \times 1$ vector, x_i , and an unknown number y_i . We assume $y_t = (y_1, \dots, y_{N_t})'$ to be the realized outcome of a random vector $Y_t = (Y_1, \dots, Y_{N_t})'$ which is related to the matrix $X_t = (x_1, \dots, x_{N_t})'$ through the superpopulation model ξ :

$$\begin{aligned} E_{\xi}(Y_t) &= X_t\beta \\ E_{\xi}[(Y_t - X_t\beta)(Y_t - X_t\beta)'] &= \sigma^2 V_t \end{aligned}$$

where $\beta = (\beta_1, \dots, \beta_q)$ and σ^2 are unknown parameters, and $V_t = (v_{ik})$, a positive definite matrix such that

$$v_{ik} = \begin{cases} v_i & i = k \\ \rho(v_i v_k)^{1/2} & i \neq k \end{cases}$$

with $v_i > 0$ known ($i = 1, \dots, N_t$); ρ unknown constant correlation coefficient such that $-(N_t - 1)^{-1} \leq \rho < 1$.

Let π_{it} the probability that the i th unit is included in the t th sample, and I_{it} the random variable which is equal to 1 if the i th unit is included in the t th sample and 0 otherwise.

Without loss of generality, we list the sampled units first; using the notation s for the sample and r for the remainder of the population, we partition Y_t , X_t , $\Pi_t = \text{diag}(\pi_1, \dots, \pi_t)$ and 1_t , $N_t \times 1$ vector of ones, in the form:

$$\begin{aligned} Y_t &= (Y'_{st}, Y'_{rt})' \\ X_t &= (X'_{st}, X'_{rt})' \\ \Pi_t &= \text{diag}(\pi_{1t}, \dots, \pi_{N_t t}) = \begin{pmatrix} \Pi_{st} & 0 \\ 0 & \Pi_{rt} \end{pmatrix} \\ 1_t &= (1'_{st}, 1'_{rt})' \end{aligned}$$

We shall consider the generalized regression estimator¹ as estimator of the population mean, $\bar{Y}_t = \frac{1}{N_t} \sum_{i=1}^{N_t} Y_{it}$:

$$e_{RG}(Q) = \frac{1}{N_t} 1'_{st} \Pi_{st}^{-1} Y_{st} + \frac{1}{N_t} (1' X_t - 1'_{st} \Pi_{st}^{-1} X_{st}) \hat{\beta}(Q_{st}),$$

where $\hat{\beta}(Q_{st}) = (X'_{st} Q_{st} X_{st})^{-1} X'_{st} Q_{st} Y_{st}$ and Q_{st} , symmetric and positive defined matrix.

In what follows, the asymptotic framework of Isaki and Fuller (1982) is adopted. In particular, a linear homogeneous estimator,

$$e_t = L'_t Y_{st} = N_t^{-1} (l_{1st}, \dots, l_{n_t st}) Y_{st},$$

is asymptotic design unbiased if

$$\lim_{t \rightarrow \infty} [E_p(e_t) - \bar{Y}_t] = 0.$$

We make the following assumptions which will often be appropriate in practice:

$$\text{C.1 } \limsup_{t \rightarrow \infty} N_t^{-1} \sum_{i=1}^{N_t} x_{ij}^2 < \infty \quad j = 1, \dots, q,$$

$$\text{C.2 } \limsup_{t \rightarrow \infty} E_p(\hat{\beta}_{jt})^2 < \infty \quad j = 1, \dots, q,$$

$$\text{C.3 } \liminf_{t \rightarrow \infty} \frac{N_t}{n_t} \min_{1 \leq i \leq N_t} \pi_{it} > 0,$$

$$\text{C.4 } \limsup_{t \rightarrow \infty} \frac{N_t^2}{n_t} \max_{i \neq k} |\pi_{ikt} - \pi_{it} \pi_{kt}| < \infty,$$

$$\text{C.5 } \limsup_{t \rightarrow \infty} \frac{1}{N_t} \sum_{i=1}^{N_t} v_i < \infty,$$

C.6 For a given design p , there exists a constant K such that for all large t ,

$$n_t \sum_{j=1}^q E_{\xi}(\beta_j - \hat{\beta}_{jt})^2 < K < \infty,$$

where $\pi_{ikt} = p(I_{it} = I_{kt} = 1)$, second order inclusion probabilities.

Casas and Guijarro (1993) proved that the generalized regression estimator is asymptotically design unbiased under C.1-C.4.²

In order to simplify expressions, we will omit t .

A linear estimator, e , is said to be robust against covariance matrix misspecification, or weakly robust if, for all s with $p(s) > 0$ holds:

$$Q_s^{-1} \left(L - \frac{1}{N} \Pi_{0s}^{-1} 1_s \right) \in C(X_s)$$

for some symmetric and positive matrix Q_s . $C(X_s)$ denotes column space generated by X_s and $\Pi_{0s} = \text{diag}(\pi_{01}, \dots, \pi_{0n}) = n \left(\sum_{i=1}^n v_i^{1/2} \right)^{-1} \text{diag}(v_1^{1/2}, \dots, v_n^{1/2})$.

3. ASYMPTOTICALLY DESIGN UNBIASED AND WEAKLY ROBUST LINEAR ESTIMATORS

Tam (1988) gave a solution for choosing between design unbiased and weakly robust linear estimators, assuming a superpopulation model with uncorrelated residuals. His result is generalized to a correlated context as shows the following theorem (Casas and Guijarro, 1994).

Theorem 1

Let $L'_m Y_s$ ($m = 1, 2$) two weakly robust and asymptotically design unbiased estimators of $\bar{Y}$, and assume:

$$\text{H.1 } \lim_{t \rightarrow \infty} \frac{n_t}{N_t} \max_{1 \leq i \leq N_t} E_p(l_{m_i s}^2 I_{it}) < \infty,$$

$$\text{H.2 } \lim_{t \rightarrow \infty} \max_{1 \leq i \leq N_t} |E_p(l_{m_i s} I_{it}) - 1| < \infty,$$

$$\text{H.3 } \lim_{t \rightarrow \infty} n_t \max_{i \neq k} |E_p(l_{m_i s} l_{m_k s} I_{it} I_{kt}) - 1| < \infty,$$

where $L'_m = N_t^{-1}(l_{m_1 s}, \dots, l_{m_{n_s} s})$.³ If the model bias of $L'_1 Y_s$ is bigger than that of $L'_2 Y_s$ (both in absolute terms), then, asymptotically,

$$E_p E_\xi (L'_1 Y_s - \bar{Y}_t)^2 \geq E_p E_\xi (L'_2 Y_s - \bar{Y}_t)^2,$$

for all sampling designs p .

Proof. By the weakly robustness condition of the estimators, we can write:

$$L'_m Y_s = e_{RG} + \left(L'_m X_s - \frac{1'_t X_t}{N_t} \right) \hat{\beta},$$

for some symmetric and definite positive matrix, Q .⁴

Therefore,

$$\begin{aligned} n_t E_p E_\xi (L'_m Y_s - \bar{Y}_t)^2 &= n_t E_p E_\xi \left[e_{RG}^* + \left(L'_m X_s - \frac{1'_t X_t}{N_t} \right) \beta - \bar{Y}_t \right]^2 + \\ & n_t E_p E_\xi \left[e_{RG} - e_{RG}^* - \left(L'_m X_s - \frac{1'_t X_t}{N_t} \right) (\beta - \hat{\beta}) \right]^2 + \\ & + 2n_t E_p E_\xi \left[e_{RG}^* + \left(L'_m X_s - \frac{1'_t X_t}{N_t} \right) \beta - \bar{Y}_t \right] \cdot \\ & \cdot \left[e_{RG} - e_{RG}^* - \left(L'_m X_s - \frac{1'_t X_t}{N_t} \right) (\beta - \hat{\beta}) \right], \end{aligned}$$

where e_{RG}^* is the random variable defined from e_{RG} when $\hat{\beta}$ is replaced by β , unknown parameter.

By standard calculations we get that

$$n_t E_p E_\xi \left[e_{RG}^* + \left(L'_m X_s - \frac{1'_t X_t}{N_t} \right) \beta - \bar{Y}_t \right]^2$$

is equal to

$$n_t E_p E_\xi (e_{RG}^* - \bar{Y}_t)^2 + \frac{n_t}{N_t^2} E_p \left[\sum_{j=1}^q \beta_j \sum_{i=1}^{N_t} (l_{is} I_{it} - 1) x_{ij} \right]^2,$$

where β_j is the j th element of the random vector β .

Under conditions C.3-C.5:

$$n_t E_p E_\xi (e_{RG}^* - \bar{Y}_t)^2 < \infty,$$

as $t \rightarrow \infty$ (Casas and Guijarro, 1993).

Also, it can be easily proved that

$$\frac{n_t}{N_t^2} E_p \left(\sum_{j=1}^q \beta_j \sum_{i=1}^{N_t} (l_{is} I_{it} - 1) x_{ij} \right)^2 < \infty,$$

as t tends to ∞ , using conditions H.1-H.3 and C.1.

Conditions H.1, H.3, C.1, C.3, C.4 and C.6 lead to

$$\lim_{t \rightarrow \infty} n_t E_p E_\xi \left[e_{RG} - e_{RG}^* - \left(L'_m X_s - \frac{1'_t X_t}{N_t} \right) (\beta - \hat{\beta}) \right]^2 = 0$$

and, thereby, applying the Schawartz inequality:

$$2n_t E_p E_\xi \left[e_{RG}^* + \left(L'_m X_s - \frac{1'_t X_t}{N_t} \right) \beta - \bar{Y}_t \right] \cdot \left[e_{RG} - e_{RG}^* - \left(L'_m X_s - \frac{1'_t X_t}{N_t} \right) (\beta - \hat{\beta}) \right]$$

tends to 0 as $t \rightarrow \infty$.

It follows that, asymptotically,

$$E_p E_\xi (L'_2 Y_s - \bar{Y}_t)^2 = E_p E_\xi (e_{RG}^* - \bar{Y}_t)^2 + E_p \left[\left(L'_2 X_s - \frac{1'_t X_t}{N_t} \right) \beta \right]^2.$$

Therefore, if

$$\left| \left(L'_1 X_s - \frac{1'_t X_t}{N_t} \right) \beta \right| \geq \left| \left(L'_2 X_s - \frac{1'_t X_t}{N_t} \right) \beta \right|,$$

then,

$$E_p \left[\left(L'_1 X_s - \frac{1'_t X_t}{N_t} \right) \beta \right]^2 \geq E_p \left[\left(L'_2 X_s - \frac{1'_t X_t}{N_t} \right) \beta \right]^2$$

because the design expectation of the square of the model bias is a monotonic function. This completes the proof of the theorem.

The next result provides a way for selecting between linear asymptotically design unbiased estimators.

Theorem 2

Let $L'Y_s$ and $L'_1 Y_s$ two linear asymptotically design unbiased estimators of $\bar{Y}$. Let

$$\lim_{t \rightarrow \infty} n_t \max_{i \neq k} \left| E_p(l_{is} l_{ks} I_{it} I_{kt}) - \frac{\pi_{ikt}}{\pi_{it} \pi_{kt}} \right| = 0$$

and

$$\lim_{t \rightarrow \infty} n_t^2 \max_{1 \leq i \leq N_t} |E_p(l_{is} I_{it}) - 1| < \infty,$$

where

$$L' = N_t^{-1}(l_{1s}, \dots, l_{ns}),$$

and let $L'_1 Y_s$ be a weakly robust estimator verifying H.1-H.3, then:

$$\begin{aligned} n_t E_p E_\xi (L'Y_s - \bar{Y}_t)^2 - n_t E_p E_\xi (L'_1 Y_s - \bar{Y}_t)^2 &= \\ &= \frac{n_t}{N_t^2} \left[\sigma^2 \sum_{i=1}^{N_t} E_p(l_{is}^2 I_{it}) v_i - \sigma^2 \sum_{i=1}^{N_t} \frac{v_i}{\pi_{it}} \right] + \\ &+ E_p[E_\xi(L'Y_s - \bar{Y}_t)]^2 - E_p[E_\xi(L'_1 Y_s - \bar{Y}_t)]^2 + C_t, \end{aligned}$$

for a given sampling design p , and

$$\lim_{t \rightarrow \infty} C_t = 0.$$

Proof. Using the equality (Cassel, Särndal and Wretman, 1977),

$$\begin{aligned} n_t E_p E_\xi (L'Y_s - \bar{Y}_t)^2 &= n_t E_p E_\xi [L'Y_s - E_\xi(L'Y_s)]^2 + n_t E_p [E_\xi(L'Y_s - \bar{Y}_t)]^2 + \\ &+ n_t E_\xi [E_p(L'Y_s) - \bar{Y}_t]^2 - n_t E_\xi \left[E_p(L'Y_s) - \frac{1'_t X_t \beta}{N_t} \right]^2 \end{aligned}$$

and theorem 1, we get

$$\begin{aligned} & n_t E_p E_\xi (L'Y_s - \bar{Y}_t)^2 - n_t E_p E_\xi (L'_1 Y_s - \bar{Y}_t)^2 = \\ & = n_t E_p E_\xi [L'Y_s - E_\xi(L'Y_s)]^2 + n_t E_p [E_\xi(L'Y_s - \bar{Y}_t)]^2 + \\ & + n_t E_\xi [E_p(L'Y_s) - \bar{Y}_t]^2 - n_t E_\xi \left[E_p(L'Y_s) - \frac{1'_t X_t \beta}{N_t} \right]^2 - \\ & n_t E_p E_\xi (e_{RG}^* - \bar{Y}_t)^2 - n_t E_p [E_\xi(L'_1 Y_s - \bar{Y}_t)]^2 - A_t \end{aligned}$$

with

$$\lim_{t \rightarrow \infty} A_t = 0.$$

The expression

$$\begin{aligned} & n_t E_p E_\xi [L'Y_s - E_\xi(L'Y_s)]^2 + n_t E_\xi [E_p(L'Y_s) - \bar{Y}_t]^2 - \\ & - n_t E_\xi \left[E_p(L'Y_s) - \frac{1'_t X_t \beta}{N_t} \right]^2 - n_t E_p E_\xi (e_{RG}^* - \bar{Y}_t)^2 \end{aligned}$$

may be written as

$$\begin{aligned} & 2 \frac{n_t}{N_t^2} \sigma^2 \sum_{i=1}^{N_t} [1 - E_p(l_{is} I_{it})] v_i + \\ & + 2 \frac{n_t}{N_t^2} \sigma^2 \rho \sum_{i=1}^{N_t} \sum_{i \neq k} [1 - E_p(l_{is} I_{it})] (v_i v_k)^{1/2} + \\ & + \frac{n_t}{N_t^2} \sigma^2 \rho \sum_{i=1}^{N_t} \sum_{i \neq k} \left[E_p(l_{is} l_{ks} I_{it} I_{kt}) - \frac{\pi_{ikt}}{\pi_{it} \pi_{kt}} \right] (v_i v_k)^{1/2} + \\ & + \frac{n_t}{N_t^2} \sigma^2 \left[\sum_{i=1}^{N_t} E_p(l_{is}^2 I_{it}) v_i - \sum_{i=1}^{N_t} \frac{v_i}{\pi_{it}} \right]. \end{aligned}$$

The first term of the former expression is dominated by

$$2 \sigma^2 \frac{n_t}{N_t} \max_{1 \leq i \leq N_t} |E_p(l_{is} I_{it}) - 1| \frac{1}{N_t} \sum_{i=1}^{N_t} v_i,$$

that tends to 0 when $t \rightarrow \infty$. The second term is dominated by

$$\frac{1}{n_t} 2 \sigma^2 |\rho| n_t^2 \max_{1 \leq i \leq N_t} |E_p(l_{is} I_{it}) - 1| \frac{1}{N_t} \sum_{i=1}^{N_t} v_i,$$

that, also, tends to 0 as $t \rightarrow \infty$; the same holds for the third term, dominated by

$$\sigma^2 |\rho| n_t \max_{i \neq k} \left| E_p(l_{is} l_{ks} I_{it} I_{kt}) - \frac{\pi_{ikt}}{\pi_{it} \pi_{kt}} \right| \frac{1}{N_t} \sum_{i=1}^{N_t} v_i.$$

Hence,

$$\begin{aligned} & n_t E_p E_\xi (L'Y_s - \bar{Y}_t)^2 - n_t E_p E_\xi (L'_1 Y_s - \bar{Y}_t)^2 = \\ &= \frac{n_t}{N_t^2} \sigma^2 \left[\sum_{i=1}^{N_t} E_p(l_{is}^2 I_{it}) v_i - \sigma^2 \sum_{i=1}^{n_t} \frac{v_i}{\pi_{it}} \right] + n_t E_p [E_\xi (L'Y_s - \bar{Y}_t)]^2 - \\ & - n_t E_p [E_\xi (L'_1 Y_s - \bar{Y}_t)]^2 + C_t \end{aligned} \quad [1]$$

where

$$\lim_{t \rightarrow \infty} C_t = 0.$$

Conditions H.1, C.3 and C.5 lead to the convergence of the following expressions:

$$\frac{n_t}{N_t^2} \sigma^2 \sum_{i=1}^{N_t} E_p(l_{is}^2 I_{it}) v_i \leq \frac{n_t}{N_t} \sigma^2 \max_{1 \leq i \leq N_t} E_p(l_{is}^2 I_{it}) \frac{1}{N_t} \sum_{i=1}^{N_t} v_i < \infty$$

and

$$\frac{n_t}{N_t^2} \sigma^2 \sum_{i=1}^{N_t} \frac{v_i}{\pi_{it}} \leq \frac{n_t}{N_t} \frac{\sigma^2}{\min_{1 \leq i \leq N_t} \pi_{it}} \frac{1}{N_t} \sum_{i=1}^{N_t} v_i < \infty.$$

Since $E_p(l_{is} I_{it}) = 1$ asymptotically for $i = 1, \dots, N_t$:

$$E_p(l_{is} I_{it}) = \pi_{it} E_p(l_{is}/I_{it} = 1) \quad i = 1, \dots, N_t.$$

This implies that, asymptotically,

$$E_p(l_{is}/I_{it} = 1) = \frac{1}{\pi_{it}} \quad i = 1, \dots, N_t.$$

So,

$$E_p(l_{is}^2 I_{it}) = \pi_{it} E_p(l_{is}^2/I_{it} = 1) \geq \pi_{it} \frac{1}{\pi_{it}^2} = \frac{1}{\pi_{it}},$$

asymptotically for all i .

And, then,

$$\frac{n_t}{N_t^2} \sigma^2 \sum_{i=1}^{N_t} E_p(I_{it} l_{is}^2) v_i \geq \frac{n_t}{N_t^2} \sigma^2 \sum_{i=1}^{N_t} \frac{v_i}{\pi_{it}}.$$

Note that if

$$n_t E_p [E_\xi (L'Y_s - \bar{Y}_t)]^2 - n_t E_p [E_\xi (L'_1 Y_s - \bar{Y}_t)]^2 \geq 0,$$

that is, if the $L'Y_s$ model bias is bigger than L'_1Y_s model bias, the linear asymptotically design unbiased and weakly robust estimator verifying H.1-H.3 conditions, is preferable to the asymptotically design unbiased estimator $L'Y_s$.

However, when

$$n_t E_p [E_\xi(L'Y_s - \bar{Y}_t)]^2 - n_t E_p [E_\xi(L'_1Y_s - \bar{Y}_t)]^2 < 0,$$

the choice of estimator depends on $[1]$ sign. As pointed out by Tam (1988) in his uncorrelated context, unless we were sure that the sign of $[1]$ is negative, it seems better to use an asymptotically design unbiased and weakly robust estimator with H.1-H.3 conditions that provides a guarantee of robustness against model misspecified.

Although the results presented in this paper, extend those by Tam (1988), there is, however, a restriction an the estimators to be linear.

¹ This estimator belongs to the *QR estimators class* (Wright, 1983):

$$e_{QR} = \frac{1}{N} 1'_s R_s Y_s + \frac{1}{N} (1'X - 1'_s R_s X_s) \hat{\beta}(Q_s)$$

making $R_s = \Pi_s^{-1}$.

² In fact, Casas and Guijarro proved the asymptotic design unbiasedness of the generalized regression estimator under conditions C.1, C.2 and

$$\liminf_{t \rightarrow \infty} N_t \min_{1 \leq i \leq N_t} \pi_{it} = \infty$$

$$\limsup_{t \rightarrow \infty} \left| \frac{\pi_{ikt}}{\pi_{it}\pi_{kt}} - 1 \right| = 0$$

less restrictive than C.3 and C.4.

³ Although, $L'_{m_t}Y_{s_t} = N_t^{-1}(l_{m_{1s_t}}, \dots, l_{m_{n_t s_t}})Y_{s_t}$, would have been formally more correct, we shall for simplicity write $L'_mY_s = N_t^{-1}(l_{m_{1s}}, \dots, l_{m_{n_s}})Y_s$.

⁴ For simplicity we shall write e_{RG} and $\hat{\beta}$ instead of $e_{RG}(Q)$ and $\hat{\beta}(Q_s)$.

REFERENCES

- Brewer, K. R. W. (1979): "A class of robust sampling designs for large-scale surveys", *Journal of American Statistical Association* **74**, pp. 911-915.
- Cassel, C., Särndal, C. and Wretman, J. H. (1977): *Foundations of Inference in Survey Sampling*, New York: John Wiley.
- Casas, J. M. and Guijarro, M. (1993): "El estimador de regresión generalizado en el modelo de superpoblación: p -insesgadez asintótica y robustez", *Estadística Española* **35**, pp. 425-437.
- Casas, J. M. and Guijarro, M. (1994): " P -insesgadez asintótica y robustez de la estimación con modelos de superpoblaciones: un criterio de selección", *Qüestio* **18** (2), pp. 173-188.
- Godambe, V. P. (1982): "Estimation in survey sampling: robustness and optimality", *Journal of American Statistical Association* **77** pp. 393-406.
- Godambe, V. P. and Thompson, M. E. (1977): "Robust near optimal estimation in survey practice", *Bulletin of the International Statistical Institute*. **47**, pp. 129-146.
- Hansen, M. H., Madow, W. G. and Tepping, B. J. (1983): "An evaluation of model-dependent and probability-sampling inferences in sample surveys", *Journal of American Statistical Association* **78**, pp. 776-807.
- Herson, J. and Royall, R. M. (1973): "Robust estimation in finite populations. *Journal of American Statistical Association* **68**, pp. 880-893.
- Tam, S. M. (1988b): "Some results on robust estimation in finite population sampling", *Journal of American Statistical Association* **83**, pp. 242-248.