

INFORMATION LOST IN AGGREGATION:

A BAYESIAN APPROACH

H. H. Kelejian

Econometric Research Program  
Research Paper No. 23  
August 1969

The research described in this paper  
was supported by the National Science  
Foundation (GS 1840)

Princeton University  
Econometric Research Program  
207 Dickinson Hall  
Princeton, New Jersey

# INFORMATION LOST IN AGGREGATION:

## A BAYESIAN APPROACH

H. H. Kelejian

### 1. Introduction<sup>1</sup>

In a recent article Chetty (1968) demonstrated that cross sectional and time series data may be combined via Bayes' theorem to obtain sharper posterior distributions of the parameters which are common to both types of regression models. The purpose of this note is to correct and extend some of Chetty's results in order to analyze the nature of the information lost in aggregating cross sectional into time series data. Among other things, it will be shown that such information is important only in a small sample framework. For instance, it turns out that if there are a large number of cross sectional units, and if cross sectional observations are available for a large number of years, the information lost in aggregating cross sectional into time series data is negligible. Ironically, the information lost in aggregation will not be negligible if the number of cross-sectional units is small. Thus, the results presented below refute just such an assumption implicitly made by Orcutt, Watts and Edwards (1968, p. 777) in their Monte Carlo study of aggregation loss.

### 2. Model and Assumptions

We begin by assuming that we have  $T$  observations on a time series regression model. In addition, we assume that observations on the corresponding cross sectional regression model are available for  $T_1 \leq T$  of these time periods. Thus, the full

---

<sup>1</sup>The research described in this paper was supported by the National Science Foundation (GS 1840).

information posterior distribution of the parameters common to both types of models would be based on the  $T_1$  sets of cross sectional observations, and the complementary set of  $T-T_1$  joint time series observations.<sup>2</sup> The difference between the variance-covariance matrix of the full information posterior distribution and that obtained for the corresponding posterior distribution based only on the  $T$  joint time series observations can be taken as a measure of the information lost in aggregation.

Consider the regression model

$$(1) \quad y_{it} = a + x_t b_1 + w_i b_2 + z_{it} b_3 + \epsilon_{it}, \quad \begin{array}{l} i = 1, \dots, n; \\ t = 1, \dots, T, \end{array}$$

where  $y_{it}$  is the value of the dependent variable corresponding to the  $i^{\text{th}}$  cross sectional unit at time  $t$ ;  $a$  is a constant;  $b_1$  is a  $K_1 \times 1$  vector of parameters;  $x_t$  is a  $1 \times K_1$  vector of observations at time  $t$  on  $K_1$  independent variables;<sup>3</sup>  $b_2$  is a  $K_2 \times 1$  vector of parameters;  $w_i$  is a  $1 \times K_2$  vector of observations on  $K_2$  independent variables which are assumed to vary only with respect to the cross sectional units;<sup>4</sup>  $b_3$  is a  $K_3 \times 1$  vector of parameters;  $z_{it}$  is  $1 \times K_3$  vector of observations on  $K_3$  independent variables which relate at time  $t$  to the  $i^{\text{th}}$  cross sectional unit;  $\epsilon_{it}$  is a disturbance term.

---

<sup>2</sup>Chetty (1968) implicitly assumes that the cross sectional observations do not correspond to a time period covered by the time series data. See Chetty (1968, p. 281).

<sup>3</sup>Examples of variables which may only vary with respect to time are prices and interest rates.

<sup>4</sup>Examples of such variables are those related to demographic, occupational, or social class considerations.

We assume that  $\epsilon_{it}$  is normally distributed with  $E\epsilon_{it} = 0$ ,  $E\epsilon_{it}^2 = \sigma^2$ , and  $E\epsilon_{it}\epsilon_{jt} = E\epsilon_{it}\epsilon_{is} = 0$  for  $i \neq j$  and  $s \neq t$ . Finally, we assume that  $\epsilon_{it}$  is independent of all of the regressors in (1), and the distribution of the regressors in (1) is independent of  $a, b_1, b_2, b_3$ .

### 3. Aggregation Loss

In this section we first derive the posterior distribution of  $b_3$  which is based only on the  $T$  joint time series observations and then compare its variance-covariance matrix to that of the full information posterior distribution. We begin by expressing in matrix formulation the time series counterpart of (1):

$$(2) \quad \begin{aligned} Y &= XC + Zb_3 + \epsilon \\ &= HB + \epsilon \end{aligned}$$

where  $Y$  is the  $T \times 1$  vector whose  $t^{\text{th}}$  element is the cross sectional average of  $y_{it}$ ;  $C' = (A \ b_1')$  where  $A = a + \bar{w} \cdot b_2$  and where  $\bar{w}$  is the  $1 \times K_2$  vector whose elements are the cross sectional averages of the elements of  $w_i$ ;  $X$  is the  $T \times K_1 + 1$  matrix whose  $t^{\text{th}}$  row is  $(1 \ x_t)$ ;  $Z$  and  $\epsilon$  are  $T \times K_3$  and  $T \times 1$  matrices defined comparably to  $Y$ . It follows that  $H = (X \ Z)$  and  $B' = (C' \ b_3')$ . We note that the elements of  $\epsilon$  are normally distributed with  $E\epsilon = 0$  and  $E\epsilon\epsilon' = N^{-1}\sigma^2I$ .

By Bayes' theorem, the posterior distribution of  $C, b_3$ , and  $\sigma$  is

$$(3) \quad P(C, b_3, \sigma | \text{data}) = K p(C, b_3, \sigma)$$

where  $K$  is the normalizing constant,  $p(C, b_3, \sigma)$  is the

prior density, and  $L(C, b_3, \sigma | \text{data})$  is the likelihood function corresponding to (2). Henceforth, we eliminate reference to normalizing constants by using the symbol  $\propto$  to denote factor of proportionality.

Following Chetty (1968), Tiao and Zellner (1964B), and Jeffreys (1961) we assume that we have no prior information and that the elements of  $C$  and  $b_3$  are independent of  $\sigma$ . We therefore take as our prior

$$(4) \quad p(C, b_3, \sigma) \propto \sigma^{-1} .$$

It is not difficult to show--see Tiao and Zellner (1964B)--that the posterior distribution defined by (2), (3), and (4) is

$$(5) \quad P(C, b_3, \sigma | \text{data}) \propto \sigma^{-(T+1)} \exp\left[-\frac{n}{2\sigma^2} (vs^2 + Q(B, \hat{B}, H'H))\right]$$

where  $s^2 = v^{-1} (Y - H\hat{B})'(Y - H\hat{B})$ ,  $v = T - 1 - K_1 - K_3$ ,  $\hat{B} = (H'H)^{-1} H'Y$ , and  $Q(B, \hat{B}, H'H)$  is a quadratic form in  $B$ , which is centered at  $\hat{B}$  with matrix  $H'H$ :  $Q = (B - \hat{B})' H'H (B - \hat{B})$ .

Integrating out  $\sigma$  from (5) we obtain the posterior distribution of  $C$  and  $b_3$

$$(6) \quad P(C, b_3 | \text{data}) \propto \left[1 + \frac{Q(B, \hat{B}, H'H)}{vs^2}\right]^{-\frac{(v+1+K_1+K_3)}{2}} .$$

It is clear from (6) that  $B$  has a multivariate  $t$  distribution which is centered at  $\hat{B}$  with variance-covariance matrix

$s^2 \left(\frac{v}{v-2}\right) (H'H)^{-1}$ ,<sup>5</sup> It follows that the marginal distribution of

---

<sup>5</sup>A term corresponding to  $\left(\frac{v}{v-2}\right)$  is mistakenly omitted from Chetty's (1968) analysis. For an excellent discussion of the properties of the multivariate  $t$  as they relate to a multivariate normal compounded with a gamma-2 variable see Raiffa and Schlaifer (1961, pp. 256-59).

$b_3$  is also a multivariate  $t$  with mean vector given by the corresponding elements of  $\hat{B}$ , and variance-covariance matrix

$$(7) \quad V = s^2 \left( \frac{v}{v-2} \right) (Z'_{.X} Z_{.X})^{-1}$$

where  $Z_{.X} = Z - X(X'X)^{-1} X'Z$ . Finally, before the sample is drawn  $s^2$  can be considered as a random variable and hence

$$(8) \quad E[V|X, Z] = \left( \frac{\sigma^2}{n} \right) \left( \frac{v}{v-2} \right) (Z'_{.X} Z_{.X})^{-1} .$$

It is noted that with the exception of the second factor, the result given in (8) is identical to that which would be obtained from a classical regression analysis of (2).

We turn to derive the full information posterior distribution of  $b_3$ . Without loss of generality, we assume that the available  $T_1$  sets of cross sectional observations relate to the first  $T_1$  time periods. Thus our time series and cross sectional observations may be expressed in terms of the models

$$(9) \quad y_{it} = a + x_t b_1 + w_i b_2 + z_{it} b_3 + \epsilon_{it}, \quad i = 1, \dots, n; \quad t = 1, \dots, T_1;$$

$$(10) \quad y_t = A + x_t b_1 + z_t b_3 + \epsilon_t, \quad t = T_1 + 1, \dots, T,$$

where  $y_t$ ,  $z_t$ , and  $\epsilon_t$  are the cross sectional averages of  $y_{it}$ ,  $z_{it}$ , and  $\epsilon_{it}$ , and  $A$  has already been defined in (2).

The joint likelihood function of (9) and (10) is

$$(11) \quad L(a, b_1, b_2, b_3, \sigma | \text{data}) \propto \frac{1}{\sigma^{nT_1 + T - T_1}} \times$$

$$\exp \left[ - \frac{1}{2\sigma^2} \sum_{t=1}^{T_1} \sum_{i=1}^n (y_{it} - a - x_t b_1 - w_i b_2 - z_{it} b_3)^2 - \right.$$

$$\left. \frac{n}{\sigma^2} \sum_{t=T_1+1}^T (y_t - A - x_t b_1 - z_t b_3)^2 \right] .$$

Let  $G_{it} = (y_{it} - a - x_t b_1 - w_i b_2 - z_{it} b_3)$  and let  $G_t$  be the cross sectional average of  $G_{it}$ . Then  $\sum_i G_{it}^2 = nG_t^2 + \sum_i (G_{it} - G_t)^2$ . Using this, it is clear that the likelihood function in (11) can be rewritten as

$$(12) \quad L(a, b_1, b_2, b_3, \sigma | \text{data}) \propto \frac{1}{\sigma^T} \exp\left[-\frac{n}{2\sigma^2} \sum_{t=1}^T (y_t - A - x_t b_1 - z_t b_3)^2\right] \times \\ \frac{1}{\sigma^{(n-1)T_1}} \exp\left[-\frac{1}{2\sigma^2} \sum_{t=1}^T \sum_{i=1}^n (y_{it}^* - w_i b_2 - z_{it} b_3)^2\right]$$

where  $y_{it}^*$ ,  $w_i$  and  $z_{it}^*$  are, respectively, the deviations of  $y_{it}$ ,  $w_i$ , and  $z_{it}$  from their cross sectional averages.

The form of equation (12) permits an observation. For instance, the likelihood function defined by the first two factors of (12) is identical to that defined by the pure time series model (2). Therefore, the information contained in the cross sectional observations that is not accounted for by the time series data (i.e., that information lost in aggregation) is given by the likelihood function defined by the third and fourth factors of (12).<sup>6</sup> Since this likelihood is least peaked when  $y_{it}^* = 0$ ,  $w_i = 0$ , and  $z_{it}^* = 0$ , it follows that the information lost in aggregation is at a minimum when the cross sectional variances of the variables involved are zero.

---

<sup>6</sup>The factor  $\sigma^{-(n-1)T_1}$  reflects the fact that  $\sum_{i=1}^n (y_{it}^* - w_i b_2 - z_{it}^* b_3)^2$  contains only  $n-1$  linearly independent terms.

In determining our prior density, we again assume that we have no prior information and that the elements of  $a$ ,  $b_1$ ,  $b_2$ , and  $b_3$  are independent of  $\sigma$ . Thus, analogous to (4), our prior is

$$(13) \quad p(a, b_1, b_2, b_3, \sigma) \propto \sigma^{-1} .$$

Combining (12) and (13) in accordance with Bayes' theorem and using (5) we obtain the full information posterior distribution

$$(14) \quad P_F(a, b_1, b_2, b_3, \sigma | \text{data}) \propto \frac{1}{\sigma^{T+1+nT_1-T_1}} \times \\ \exp \left[ - \frac{n}{2\sigma^2} (vs^2 + Q(B, \hat{B}, H'H) - \frac{1}{2\sigma^2} \sum_{t=1}^{T_1} \sum_{i=1}^n (y_{it}^* - w_i b_2 - z_{it}^* b_3)^2) \right] ,$$

We now express the  $nT_1$  observations on  $y_{it}^*$ ,  $z_{it}^*$ , and  $w_{it}^*$  in a sequential fashion. In particular, let  $y_{it}^* = Y_{(t-1)n+i}^*$ ,  $z_{it}^* = Z_{(t-1)n+i}^*$ , and let the observations on  $w_i^*$  appearing in the double summation in (14) be numbered correspondingly. Then, denoting the double summation in (14) by  $\zeta$ , we have

$$(15) \quad \zeta = \sum_{j=1}^{nT_1} (Y_j^* - w_j^* b_2 - z_j^* b_3)^2 .$$

Now let the  $nT_1$  observations on  $y_j^*$ ,  $w_j^*$  and  $z_j^*$  be given by the vector and matrices  $\mathbf{y}^*$ ,  $\mathbf{w}^*$ ,  $\mathbf{z}^*$ . Further, let  $H_1 = (\mathbf{w}^* \mathbf{z}^*)$ ,  $D' = (b_2' \ b_3')$ , and  $\hat{D} = (H_1' H_1)^{-1} H_1' \mathbf{y}^*$ . It is clear, then, that  $\zeta$  can be expressed as

$$(16) \quad \zeta = (v_1 s_1^2 + Q(D, \hat{D}, H_1' H_1))$$



where  $v_1 = (n-1)T_1 - K_2 - K_3$ , and  $s_1^2 = v_1^{-1}(\hat{Y} - H_1 \hat{D})'(\hat{Y} - H_1 \hat{D})$ . It follows from (15) and (16) that the posterior distribution in (14) can be rewritten as

$$(17) \quad P_F(a, b_1, b_2, b_3, \sigma | \text{data}) \propto \frac{1}{\sigma^{T+1+nT_1-T_1}} \times \\ \exp\left[-\frac{1}{2\sigma^2} (nvs^2 + v_1 s_1^2 + nQ(B, \hat{B}, H'H) + Q(D, \hat{D}, H_1'H_1))\right].$$

Because we are interested in the full information posterior distribution of  $b_3$ , the other parameters appearing in (17), namely  $a, b_1, b_2$ , and  $\sigma$  must be integrated out. Since, for practical purposes, the elements of  $b_3$  are the only parameters common to both  $B$  and  $D$ , the necessary integrations can be performed using the properties of the multivariate normal. Thus, integrating (17) first with respect to  $a$ , and then with respect to  $b_1, b_2$ , and  $\sigma$  we obtain

$$(18) \quad P_F(b_3 | \text{data}) \propto [ (nvs^2 + v_1 s_1^2 + Q(b_3, \tilde{b}_3, \tilde{z}_{.w}^* \tilde{z}_{.w}^*) \\ + Q(b_3, \hat{b}_3, nZ_{.x}' Z_{.x}) ]^{-(T+nT_1-T_1-K_1-K_2-1)/2}$$

where  $\tilde{b}_3$  and  $\hat{b}_3$  are the  $K_3 \times 1$  subvectors, respectively, of  $\hat{D}$  and  $\hat{B}$  which correspond to  $b_3$ ,  $\tilde{z}_{.w}^* = z - w(w'w)^{-1}w'z$ , and  $Z_{.x}$  has been defined in (7). The final form of the posterior distribution is obtained by completing the quadratic in (18) <sup>7</sup>

---

<sup>7</sup>Both Chetty (1968) and Tiao and Zellner (1964B) mistakenly omit a term analogous to  $\Lambda$  from their results.

$$(19) P_F(b_3 | \text{data}) \propto \left[ 1 + \frac{(b_3 - b_3^F)' (Z'_{.w} Z'_{.w} + nZ'_{.x} Z'_{.x})(b_3 - b_3^F)}{(vns^2 + v_1s_1^2 + \Lambda)} \right]^{-m},$$

$$m = \frac{(T + nT_1 - T_1 - K_1 - K_2 - 1)}{2},$$

$$\text{where } b_3^F = [Z'_{.w} Z'_{.w} + nZ'_{.x} Z'_{.x}]^{-1} [Z'_{.w} Z'_{.w} \tilde{b}_3 + nZ'_{.x} Z'_{.x} \hat{b}_3],$$

$$\text{and } \Lambda = \tilde{b}_3' Z'_{.w} Z'_{.w} \tilde{b}_3 + \hat{b}_3' nZ'_{.x} Z'_{.x} \hat{b}_3 - b_3^F (Z'_{.w} Z'_{.w} + nZ'_{.x} Z'_{.x}) b_3^F.$$

It is clear from (19) that the full information posterior distribution of  $b_3$  has a multivariate  $t$  distribution with  $\Theta = T + (n-1)T_1 - K_1 - K_2 - K_3 - 1 = v + (n-1)T_1 - K_2$  degrees of freedom. The distribution is centered at  $b_3^F$  and has variance-covariance matrix

$$(20) \quad V_F = (n^{-1}S^2) \left( \frac{\Theta}{\Theta - 2} \right) (n^{-1} Z'_{.w} Z'_{.w} + Z'_{.x} Z'_{.x})^{-1},$$

where  $S^2 = (vns^2 + v_1s_1^2 + \Lambda) / \Theta$ . From the sampling theory point of view  $b_3^F$  can be interpreted as a weighted average of the least squares estimate of  $b_3$  based only on the time series data,  $\hat{b}_3$ , and the least squares estimate of  $b_3$  based only on the cross sectional data after the mean has been eliminated within each of the  $T_1$  cross sections.

In comparing the "time series" posterior distribution in (6) to the full information posterior distribution in (19) we note that the former is centered at  $\hat{b}_3$  while the latter is centered at  $b_3^F$ . Before the sample is taken  $\tilde{b}_3$ ,  $\hat{b}_3$ , and so  $b_3^F$

can be regarded as random variables. Within this sampling theory framework it is not difficult to show that  $E\hat{b}_3 = Eb_3^F = b_3$  and that the conditional variance-covariance matrix of  $\hat{b}_3$  is  $\Phi_1 = \left(\frac{\sigma}{n}\right)^2 [Z'_{.x} Z_{.x}]^{-1}$  while that of  $b_3^F$  is  $\Phi_2 = \left(\frac{\sigma}{n}\right)^2 [Z'_{.x} Z_{.x} + n^{-1} z'_{.w} z_{.w}]^{-1}$ . Thus, one would not expect the means of the posterior distributions in (6) and (19) to be different; however, if these means are used as point estimates of  $\hat{b}_3$ ,  $b_3^F$  is to be preferred to  $\hat{b}_3$  since  $\Phi_1 - \Phi_2$  is positive semi definite.<sup>8</sup>

Consider now the variance-covariance matrices in (7) and (20). Again, before the sample is taken  $S^2$  in (20) is a random variable, and using procedures similar to those used by Goldberger (1964, p. 166), it is not difficult to show that  $ES^2 = \sigma^2$ . Thus, the first factor in (7) and (20) have the same expectation. However, since  $\theta = \nu + (n-1)T_1 > \nu$ , the second factor in (20) is less than the second factor in (7). Finally, in light of footnote (8), each diagonal element of the third factor in (20) is less than or equal to the corresponding element in the third factor of (7). Therefore, the full information posterior distribution in (19) is sharper than the "time series" distribution given in (6).

---

<sup>8</sup>This follows from a theorem described in Goldberger (1964, p. 38) which states that if  $A = B + C$  where  $B$  and  $C$  are positive definite matrices, then  $B^{-1} - A^{-1}$  is positive semi definite.

A final point concerning the asymptotic counterparts of the variance-covariance matrices in (7) and (20) from the sampling theory point of view should be noted. In particular replacing  $s^2$  in (7) and  $n^{-1}s^2$  in (20) by their expected value  $(\frac{\sigma^2}{n})$  we see that

$$(21) \quad V = \left(\frac{\sigma}{n}\right)^2 \left(\frac{v}{v-2}\right) (Z'_{.x} Z_{.x})^{-1} = o((nT)^{-1})$$

and

$$(22) \quad V_F = \left(\frac{\sigma}{n}\right)^2 \left(\frac{\theta}{\theta-2}\right) [Z'_{.x} Z_{.x} + n^{-1} Z'^*_{.w} Z^*_{.w}]^{-1} = o((nT)^{-1})$$

where, in general,  $o(n^j)$  is a term of at most  $n^j$  in probability. In brief, both  $V$  and  $V_F$  are of the order  $(nT)^{-1}$ . Now note that<sup>9</sup>

$$(23) \quad [Z'_{.x} Z_{.x} + n^{-1} Z'^*_{.w} Z^*_{.w}]^{-1} = (Z'_{.x} Z_{.x})^{-1} \times \\ [I - n^{-1} Z'^*_{.w} Z^*_{.w} (Z'_{.x} Z_{.x} + Z'^*_{.w} Z^*_{.w})^{-1}] .$$

It follows from (21) - (23) that

$$(24) \quad V - V_F = \frac{\sigma^2}{n} [(Z'_{.x} Z_{.x})^{-1} \left[\frac{v}{v-2} - \frac{\theta}{\theta-2}\right] + \\ + \left(\frac{\theta}{\theta-2}\right) (Z'_{.x} Z_{.x})^{-1} (n^{-1} Z'^*_{.w} Z^*_{.w}) [Z'_{.x} Z_{.x} + Z'^*_{.w} Z^*_{.w}] ] \\ = \frac{\sigma^2}{n} [o(T^{-1}) o(T^{-1}) + o(T^{-1}) o(T_1) o((T + nT_1)^{-1})] .$$

---

<sup>9</sup>The formula used in deriving (23) is one described in Tiao and Zellner (1964B) which states that if  $Q_1$  is an  $m \times n$  matrix and  $Q_2$  is an  $n \times m$  matrix, then

$$(I - Q_1 Q_2)^{-1} = I + Q_1 (I - Q_2 Q_1)^{-1} Q_2 .$$

From (24) we see that if  $(T/T_1) \rightarrow \infty$ , the difference between  $V$  and  $V_F$  is of a lower order in probability than  $(nT)^{-1}$ ; thus  $V_F$  and  $V$  converge, asymptotically, regardless of the magnitude of  $n$ . Indeed, from (24) we see that even if  $n$ ,  $T$ , and  $T_1 \rightarrow \infty$ ,  $V$  and  $V_F$  still converge to the same matrix. These results suggest that information lost in aggregating cross sectional into time series data is actually only a small sample problem.

It is interesting to note that in the unrealistic case  $T$  and  $T_1 \rightarrow \infty$  but  $n$  remains finite and small,  $V$  and  $V_F$  do not converge to the same matrix. This, however, approximates one of the two cases considered by Orcutt, Watts and Edwards (1968) in their Monte Carlo study of aggregation loss. In particular, they considered the cases  $T = T_1 = 20$ ,  $n = 16$ , and  $T_1 = T = 80$ ,  $n = 16$ . As expected from (24), their results show that the standard errors of the regression parameter estimates increase, in both cases, when only the time series data are used. However, it is also clear from (24) that the relative increase should be less when  $T = 80$  than when  $T = 20$ . The results given in their Monte Carlo study confirm this expectation - see Orcutt, Watts, and Edwards (1968, Table 1).

BIBLIOGRAPHY

- Chetty, V. (1968), "Pooling of Time Series and Cross Section Data," Econometrica, Vol. 36, April, 1968, pp. 279-290.
- Christ, C. (1966), Econometric Models and Methods, New York: John Wiley and Sons, 1966.
- Goldberger, A. (1964), Econometric Theory, New York: John Wiley and Sons, 1964.
- Jeffreys, H. (1961), Theory of Probability (3rd Edition), Oxford: Clarendon Press, 1961.
- Lindley, D. (1965), Introduction to Probability and Statistics From a Bayesian Viewpoint (Part II), Cambridge: University Press, 1965.
- Orcutt, G., H. Watts, and J. Edwards, (1968), American Economic Review, Vol. LVIII, September 1968, pp. 773-787.
- Raiffa, H. and R. Schlaifer (1961), Applied Statistical Decision Theory, Boston: Graduate School of Business Administration, Harvard University, 1961.
- Tiao, G. and A. Zellner (1964A), "On the Bayesian Estimation of Multivariate Regression," Journal of the Royal Statistical Society, Series B, 26, pp. 277-285.
- Tiao, G. and A. Zellner (1964B), "Bayes' Theorem and the use of Prior Knowledge in Regression Analysis," Biometrika, 51, pp. 219-230.