

Prediction of rates of inbreeding in populations undergoing index selection

N. R. Wray¹, J. A. Woolliams², R. Thompson²

¹ Livestock Improvement Unit, Victorian Institute of Animal Science, 475 Mickleham Road, Attwood, Victoria 3049, Australia

² AFRC Roslin Institute (Edinburgh), Roslin, Midlothian, EH25 9PS, Scotland, UK

Received: 22 February 1993 / Accepted: 16 June 1993

Abstract. For populations undergoing mass selection, previous studies have shown that the rate of inbreeding is directly related to the mean and variance of long-term contributions from ancestors to descendants, and thus prediction of the rate of inbreeding can be achieved via the prediction of long-term contributions. In this paper, it is shown that the same relationship between the rate of inbreeding and long-term contributions is found when selection is based on an index of individual and sib records (index selection) and where sib records may be influenced by a common environment. In these situations, rates of inbreeding may be considerably higher than under mass selection. An expression for the rate of inbreeding is derived for populations undergoing index selection based on variances of (one-generation) family size and incorporating the concept of long-term selective advantage. When the mating structure is hierarchical, and when half-sib records are included in the index, the correlation between parental breeding values and the index values of their offspring is higher for male parents than female parents. This introduces an important asymmetry between the contributions of male and female ancestors to the evolution of inbreeding which is not present when selection is based on individual and/or full-sib records alone. The prediction equation for index selection accounts for this asymmetry. The prediction is compared to rates of inbreeding calculated from simulation. The prediction is good when family size is small relative to the number selected. The reasons for over-prediction in other situations are discussed.

Key words: Inbreeding rate – Effective population size – Index selection

Introduction

The mean and variance of long-term genetic contributions from ancestors (in a closed population) can be related to the rate of inbreeding (Wray and Thompson 1990). After several generations, the long-term contributions from an ancestor stabilise and are the same to all individuals born into the population, with the values differing between ancestors. The mean simply reflects the constraint of the number of ancestors and the parents used in each generation and is the same whether or not selection is practised. The variance, however, is increased by selection.

Offspring of parents who are genetically superior for the trait under selection are more likely to be selected than the offspring of genetically-average or inferior parents. The parents are said to confer a selective advantage to their offspring. When parents are selected at random, the sampling of parents is independent of the sampling process of the previous generation. However, when parents are selected on a heritable trait, the selective advantage is inherited, and is conveyed, in part, from parent to offspring. Thus the breeding value of an ancestor has influences on selection decisions in all subsequent generations. Wray et al. (1990) introduced the terminology of one-generation, two-generation and long-term selective advantage, referring to that conveyed from parent to offspring, grandparent to grandoffspring, and ancestor to (distant) descendant. They reviewed different methods to predict the rate of inbreeding which can be

classified in the same way, according to the number of generations of selective advantage they attempt to incorporate. Only the methods of Robertson (1961), Wray and Thompson (1990) and Woolliams et al. (1993) are long-term methods.

For populations undergoing mass selection Wray and Thompson (1990) presented a recursive algorithm to predict the mean and variance of long-term contributions and hence to predict the rate of inbreeding. Woolliams et al. (1993) modified components of the prediction and presented an explicit expression for the long-term selective advantage and the rate of inbreeding. Further, they showed that the important terms of the prediction can be related back to the equations of Latter (1959) and Hill (1979), with the addition of a term describing the contribution of the expected long-term selective advantage.

The present paper is concerned primarily with the prediction of the rate of inbreeding when selection is based on an index of records of an individual and its collateral relatives and where sib records may be influenced by a common environment. In these situations, the rate of inbreeding may be considerably higher than under mass selection. Justified by the formal derivation of Woolliams et al. (1993), this paper presents a more intuitive derivation which highlights more explicitly the relationship between the long-term contribution method of Wray and Thompson (1990) with that of Robertson (1961), on the one hand, and Latter (1959) and Hill (1979), on the other.

Methods

Definition of population structure and index parameters

Throughout, conventions on notation follow as closely as possible those of Woolliams et al., (1993). The population structure considered is one of hierarchical random mating of F females with M males ($M \leq F$) with discrete generations. Generation 1 is produced by the mating structure from an unrelated, unselected base population. The term 'ancestors' is used to refer to individuals born and selected in generation 1. Each female produces a family of n offspring comprising n_f males and n_f females ($n = 2n_f$). Each male has n_m offspring of each sex ($n_m = n_f F/M$). X , Y or W and subscripts x , y or w may be used to specify a single sex, either male or female e.g., $X = M$ or F , $n_x = n_m$ or n_f . T is used to denote the total number of offspring of each sex born each generation, $T = Xn_x$. The proportion selected is p after truncation at the standardised normal deviate v , $p = \Phi(v)$ with the corresponding normal ordinate $z = \phi(v)$, where $\Phi(\cdot)$ and $\phi(\cdot)$ represent the cumulative and probability density functions of the normal distribution. The standardised selection intensity is $i = z/p$ and the variance reduction factor is $k = i(i - v)$. When these terms have subscripts they are the values for the sex of animals indicated by the subscript, and without subscripts they are the average of the sexes.

Assuming an infinitesimal model of gene effects, the total genetic variance of individuals born in generation t is $\sigma_{A,t}^2$, which

can be decomposed as,

$$\sigma_{A,t}^2 = \sigma_{Am,t-1}^2 + \sigma_{Af,t-1}^2 + \sigma_{Aw}^2$$

where $\sigma_{Am,t-1}^2$ and $\sigma_{Af,t-1}^2$ are $\frac{1}{4}$ of the genetic variances between sires and dams born and selected in generation $t-1$, σ_{Aw}^2 is the within-family genetic variance, and $\sigma_{Am,0}^2 = \sigma_{Af,0}^2 = \frac{1}{2}\sigma_{Aw}^2 = \frac{1}{4}\sigma_{A,0}^2$. These parameters are used only for $t \leq 2$, and so reductions in genetic variance due to inbreeding are ignored. The phenotypic variance in generation t is,

$$\sigma_{P,t}^2 = \sigma_{A,t}^2 + \sigma_C^2 + \sigma_E^2$$

where σ_C^2 and σ_E^2 are the common environmental variance of full-sibs and the error variance respectively. Heritability in generation t is defined as $h_t^2 = \sigma_{A,t}^2 / \sigma_{P,t}^2$.

Selection is assumed to be based on an index (I_H) of individual record (P), the mean of n full-sib records (including individual) (\bar{P}_D), and the mean of $(F/M)n$ half-sib (including the individual and its full-sibs) records (\bar{P}_H),

$$I_H = \beta_1(P - \bar{P}_D) + \beta_2(\bar{P}_D - \bar{P}_H) + \beta_3\bar{P}_H,$$

where β_1 , β_2 , β_3 are selection index weights. The index is written in this way because $\text{Cov}(P - \bar{P}_D, \bar{P}_D - \bar{P}_H) = \text{Cov}(P - \bar{P}_D, \bar{P}_H) = \text{Cov}(\bar{P}_F - \bar{P}_H, \bar{P}_H) = 0$. P is defined so that the mean of P each generation is zero. Selection index weights are assumed to be constant throughout and are derived so that the index is optimum in the first generation: $\text{Cov}(A, I) = V(I)$, where A is the individual's breeding value. This assumption is adopted for simplicity and for comparison of prediction results with simulation results later; the theory can be developed analogously without this assumption (Woolliams and Wray, in preparation). The selection index weights each generation are,

$$\beta_1 = \frac{\sigma_{Aw}^2}{\sigma_{Aw}^2 + \sigma_E^2}, \quad \beta_2 = \frac{\sigma_{Af,0}^2 + \frac{1}{n}\sigma_{Aw}^2}{\sigma_{Af,0}^2 + \sigma_C^2 + \frac{1}{n}(\sigma_{Aw}^2 + \sigma_E^2)}$$

and

$$\beta_3 = \frac{\sigma_{Am,0}^2 + \frac{M}{F}\left(\sigma_{Af,0}^2 + \frac{1}{n}\sigma_{Aw}^2\right)}{\sigma_{Am,0}^2 + \frac{M}{F}\left[\sigma_{Af,0}^2 + \sigma_C^2 + \frac{1}{n}(\sigma_{Aw}^2 + \sigma_E^2)\right]}$$

Before selection the variance of the indices of individuals born in generation t is,

$$\sigma_{I,t}^2 = \beta_1^2(\sigma_{Aw}^2 + \sigma_E^2)\left(1 - \frac{1}{n}\right) + \beta_2^2\left[\sigma_{Af,t-1}^2 + \sigma_C^2 + \frac{1}{n}(\sigma_{Aw}^2 + \sigma_E^2)\right]\left(1 - \frac{M}{F}\right) + \beta_3^2\left\{\sigma_{Am,t-1}^2 + \frac{M}{F}\left[\sigma_{Af,t-1}^2 + \sigma_C^2 + \frac{1}{n}(\sigma_{Aw}^2 + \sigma_E^2)\right]\right\}. \quad (1)$$

Before selection $\text{Cov}(A$ of sex x parent, I_H of offspring) = $2\tau_x\sigma_{Ax,t-1}^2$ for offspring born in generation t , where

$$\tau_m = \beta_3 \quad \text{and} \quad \tau_f = \beta_2(1 - M/F) + \beta_3(M/F). \quad (2)$$

In the absence of selection $\text{Cov}(I_H$ of sex x parent, I_H of offspring) = $\frac{1}{2}\tau_x\sigma_{I,0}^2$.

The correlations between indices of two full-sibs ($\rho_{D,t}$) and two half-sibs ($\rho_{H,t}$) are

$$\rho_{D,t} = \left\{ -\beta_1^2 \frac{1}{n} (\sigma_{A_w}^2 + \sigma_E^2) + \beta_2^2 \left(1 - \frac{M}{F} \right) \left[\sigma_{A_{f,t-1}}^2 + \sigma_C^2 + \frac{1}{n} (\sigma_{A_w}^2 + \sigma_E^2) \right] + \beta_3^2 \left[\sigma_{A_{m,t-1}}^2 + \frac{M}{F} \left[\sigma_{A_{f,t-1}}^2 + \sigma_C^2 + \frac{1}{n} (\sigma_{A_w}^2 + \sigma_E^2) \right] \right] \right\} \sigma_{I,t}^{-2}, \quad (3)$$

and

$$\rho_{H,t} = \left\{ -\beta_2^2 \frac{M}{F} \left[\sigma_{A_{f,t-1}}^2 + \sigma_C^2 + \frac{1}{n} (\sigma_{A_w}^2 + \sigma_E^2) \right] + \beta_3^2 \left[\sigma_{A_{m,t-1}}^2 + \frac{M}{F} \left[\sigma_{A_{f,t-1}}^2 + \sigma_C^2 + \frac{1}{n} (\sigma_{A_w}^2 + \sigma_E^2) \right] \right] \right\} \sigma_{I,t}^{-2}. \quad (4)$$

The correlation between indices of two full-sibs due to the breeding values of the dam is $\rho_{f,t}$ and the correlation between indices of two full-sibs or two half-sibs due to the breeding value of their sire is $\rho_{m,t}$ where

$$\rho_{x,t} = \tau_x^2 \sigma_{A_{x,t-1}}^2 \sigma_{I,t}^{-2}. \quad (5)$$

Two other selection indices are considered: an index of individual record and full-sib mean, $I_D = \beta_1(P - \bar{P}_D) + \beta_2 \bar{P}_D$, where β_1 is the same as in I_H and

$$\beta_2 = \frac{\sigma_{A_{m,0}}^2 + \sigma_{A_{f,0}}^2 + \frac{1}{n} \sigma_{A_w}^2}{\sigma_{A_{m,0}}^2 + \sigma_{A_{f,0}}^2 + \sigma_C^2 + \frac{1}{n} (\sigma_{A_w}^2 + \sigma_E^2)},$$

and an index of individual record only (mass selection), $I_P = \beta_1 P$, where $\beta_1 = h^2$. In the methodology that follows, derivations are made for index I_H . The results are also appropriate for I_D if in I_H , β_3 is set equal to β_2 of I_D and for I_P if in I_H , β_3 and β_2 are set equal to β_1 of I_P . For each index $\sigma_{I,t}^2$ must be calculated appropriately.

Simplification of the method of Wray and Thompson (1990)

Rate of inbreeding from long-term contributions. Under the assumption of constant rate of inbreeding each year, Wray and Thompson (1990) presented an expression for the rate of inbreeding (ΔF) appropriate for selected populations,

$$\Delta F \approx \frac{1}{4(M+F)^2} \sum_{j=1}^{M+F} r_{j,t}^2 \quad (6)$$

where $r_{j,t}$ is the total additive genetic contribution of ancestor j born in generation 1 to its descendants born in generation t . Alternatively, $r_{j,t}/(M+F)$ is the additive genetic relationship between the Mendelian sampling term that ancestor j received and each descendant, i.e., the genetic relationship between ancestor and descendant which cannot be traced to the base generation (the parents of the ancestors). Appendix 1 shows that this expression can be partitioned by sex of ancestor and by sex of descendant to give,

$$\Delta F \approx \frac{1}{16} \left\{ \frac{1}{M^2} \left[\sum_{j=1}^M r_{j(mm),\infty}^2 + \sum_{j=1}^F r_{j(fm),\infty}^2 \right] \right.$$

$$\left. + \frac{2}{MF} \left[\sum_{j=1}^M r_{j(mm),\infty} r_{j(mf),\infty} + \sum_{j=1}^F r_{j(fm),\infty} r_{j(ff),\infty} \right] + \frac{1}{F^2} \left[\sum_{j=1}^M r_{j(mf),\infty}^2 + \sum_{j=1}^F r_{j(ff),\infty}^2 \right] \right\} \quad (7)$$

where $r_{j(xy),t}$ is the long-term contribution of ancestor j of sex x to its descendants of sex y . Terms $r_{j(xy),t}$ have mean $\mu_{r(xy),t}$, variance $\sigma_{r(xy),t}^2$ and covariance between male and female descendants of $\sigma_{r(xm,xf),t}$. Therefore, an equivalent expression can be written,

$$E[\Delta F] \approx \frac{1}{16M} \left\{ [\mu_{r(mm),\infty}^2 + \sigma_{r(mm),\infty}^2] + 2 \left[\frac{M}{F} \right] [\mu_{r(mm),\infty} \mu_{r(mf),\infty} + \sigma_{r(mm,mf),\infty}] + \left[\frac{M}{F} \right]^2 [\mu_{r(mf),\infty}^2 + \sigma_{r(mf),\infty}^2] \right\} + \frac{1}{16F} \left\{ [\mu_{r(ff),\infty}^2 + \sigma_{r(ff),\infty}^2] + 2 \left[\frac{F}{M} \right] [\mu_{r(fm),\infty} \mu_{r(ff),\infty} + \sigma_{r(fm,ff),\infty}] + \left[\frac{F}{M} \right]^2 [\mu_{r(fm),\infty}^2 + \sigma_{r(fm),\infty}^2] \right\}. \quad (8)$$

Mean of long term contributions. The mean $\mu_{r(xy),t} = E[r_{j(xy),t}]$ is an expectation conditional on the deviation of the breeding value of ancestor j over that of its selected contemporaries of $A_{j(x)}$, such that $E[A_{j(x)}] = 0$ and $V[A_{j(x)}] = 4\sigma_{A_{x,1}}^2$ (which is the variance about the mean of all selected ancestors of sex x , evaluated in Appendix 2). Assuming a linear model, the mean can be expressed as

$$\mu_{r(xy),t} = E[r_{j(xy),t}] = 2^{t-2} E \left[\frac{Y}{X} + b_{xy,t} A_{j(x)} \right] \left(\frac{1}{2} \right)^{t-1} = \frac{1}{2} E \left[\frac{Y}{X} + b_{xy,t} A_{j(x)} \right] \quad (9)$$

(Wray and Thompson 1990). The term $2^{t-2} b_{xy,t}$ can be interpreted as the regression coefficient of the number of distinct pedigree pathways to descendants of sex y in generation t on the breeding values of their ancestors of sex x . The term $(\frac{1}{2})^{t-1}$ represents the relationship between ancestor and descendant along a single pathway. When selection is at random $b_{xy,t}$ is zero. Under selection, $b_{xy,2}$ is the one-generation selective advantage and $b_{xy,\infty}$ is the long-term selective advantage.

Under mass selection Wray and Thompson (1990) showed that,

$$b_{xy,2} \approx \frac{1}{2\sigma_{P,2}} n_x z_y = \frac{1}{2\sigma_{P,2}} \frac{Y}{X} i_y,$$

where $\sigma_{P,2}$ is the phenotypic standard deviation in generation 2. They presented a recursion to calculate $b_{xy,t}$ and ultimately $b_{xy,\infty}$. However, Woolliams et al. (1993) derived a direct expression for $b_{xy,t}$ and showed that,

$$b_{xy,\infty} \approx \frac{1}{2\sigma_{P,2}} \frac{Y}{X} i S_\infty = \frac{1}{\sigma_{P,2}} \frac{Y}{X} \frac{i}{(1+kh_2^2)}, \quad (10)$$

where S_∞ represents the sum of an infinite series and $S_\infty = (1-c)^{-1}$ with $c = 0.5(1-kh_2^2)$ defined as the 'coefficient of com-

petitiveness'. For selection indices I_H or I_D (or I_P) the expressions for $b_{xy,2}$ and $b_{xy,\infty}$ are derived in Appendix 2 resulting in,

$$b_{xy,2} \approx \frac{1}{2\sigma_{I,2}} n_x \tau_x z_y = \frac{1}{2\sigma_{I,2}} \frac{Y}{X} \tau_x i_y$$

$$b_{xy,\infty} \approx \frac{1}{2\sigma_{I,2}} \frac{Y}{X} i S_\infty \frac{(\tau + \tau_x)}{2} (1 - \psi) = \frac{1}{2\sigma_{I,2}} \frac{Y}{X} i \frac{(\tau + \tau_x)}{1 + k\tau} (1 - \psi) \quad (11)$$

where S_∞ is defined as for mass selection as $(1 - c)^{-1}$ but the coefficient of competitiveness is more generally defined as $c = \frac{1}{2}(1 - k\tau)$ with $\tau = \frac{1}{2}(\tau_m + \tau_f)$ and τ_x is defined in equation (2) and where $\psi = (k_m - k_f)(\tau_m - \tau_f)/8$. For selection on index I_P (mass selection) where $\tau_m = \tau_f = h^2$ and $\sigma_{I,2} = h^2 \sigma_{P,2}$, equation (11) reduces to equation (10) except that h_2^2 is replaced by $\tau = h^2$ here, where h_2^2 arose from a more-accurate approximation to equation (A2.2) than the one used here for reasons of complexity with index selection.

If the increase in selective advantage is defined as B_{xy} where

$$B_{xy} = b_{xy,\infty}/b_{xy,2}, \quad (12)$$

then by examination of equations (10) and (11) it can be seen that B_{xy} is independent of the sex of the ancestor (x) when $\tau_m = \tau_f$ as is the case for mass selection or selection using index I_D , whilst for selection using I_H , B_{xy} is dependent on x .

Evaluation of equation (8) with $t = 2$ rather than $t = \infty$. Let us now examine equation (8) but using, in the first instance, $t = 2$ rather than $t = \infty$. Firstly, by noting that $n_x = T/X = Fn_f/X$, it is found that

$$\mu_{r(xy),2}^2 = \frac{1}{4} \left\{ \left(\frac{Y}{X} \right)^2 + b_{xy,2}^2 V[A_{i(x)}] \right\}$$

$$\approx \frac{1}{4} \left[\left(\frac{Y}{X} \right)^2 + \frac{1}{4\sigma_{I,2}^2} n_x^2 \tau_x^2 z_y^2 4\sigma_{Ax,1}^2 \right]$$

$$= \frac{1}{4} \left[\left(\frac{Y}{X} \right)^2 + \left(\frac{F}{X} \right)^2 n_f^2 \rho_{x,2} z_y^2 \right] \quad (13)$$

where $\rho_{x,2}$ is the correlation between full-sibs due to the breeding value of the parent of sex x , as defined in equation (5).

An extended form of equation (22) of Wray and Thompson (1990) approximates $\sigma_{r(xy),2}^2$ as,

$$\sigma_{r(xy),2}^2 \approx \frac{1}{4} E \left\{ n_x \left[p_y + \frac{b_{xy,2}}{n_x} A_{j(x)} \right] \left[1 - p_y - \frac{b_{xy,2}}{n_x} A_{j(x)} \right] \right. \\ \left. + fs_{xy,2} + hs_{xy,2} \right\} \quad (14)$$

where the first term is the binomial sampling variance of the number selected from the n_x offspring of sex y born to parents of sex x , each of which is selected with probability $p_y + (b_{xy,2}/n_x)A_{j(x)}$ which depends on the genetic merit of the parent. For the population structure considered here, this sampling should be hypergeometric because family sizes before selection are constant and the sampling is without replacement. Woolliams et al. (1993) approximated this by multiplying the term by $(X - 1)/X$. For random selection ($b_{xy,2} = 0$), this results in $n_x p_y (1 - p_y) (X - 1) X^{-1} = YX^{-1} (T - Y) T^{-1} (X - 1) X^{-1}$, this is an approximation to the exact hypergeometric variance $YX^{-1} (T - Y) (T - 1)^{-1} (X - 1) X^{-1}$, which shall be used here. (The correction to the terms involving $A_{j(x)}$ will be ignored until the section 'More accurate prediction of coselection of sibs'.) The term $fs_{xy,2}$ [equation (14)] is the probability of coselection of full-sibs not attributed to the parents of sex x [which has already

been accounted for via the $b_{xy,2}^2 V[A_{j(x)}]$ in equation (13)], thus

$$fs_{xy,2} = \frac{F}{X} n_f (n_f - 1) (\rho_{D,2} - \rho_{x,2}) z_y^2$$

where $\rho_{D,2}$ is the correlation between indices of full-sibs [equation (3)] and $\rho_{x,2}$ is the correlation between indices of full-sibs attributed to the breeding value of the parents of sex x [equation (5)]. For a general correlation between sib indices ρ , ρz_y^2 is an approximation to the additional probability of coselection of two sibs of sex y (Robertson 1961). The coefficient $(F/X)n_f(n_f - 1)$ reflects that each of the n_f offspring of sex y has $(n_f - 1)$ opportunities for coselection with a sib and that a parent of sex x contributes of F/X full-sib families. The term $hs_{xy,2}$ [in equation (14)] is the additional probability of coselection of half-sibs not attributed to the parent of sex x ; $hs_{fy,2} = 0$ since in the hierarchal population structure there are no maternal half-sibs; and analogously to the full-sib coselection, if $\rho_{H,2}$ is the correlation between indices of half-sibs [equation (4)], then,

$$hs_{xy,2} = \frac{F}{X} \left(\frac{F}{X} - 1 \right) n_f^2 (\rho_{H,2} - \rho_{m,2}) z_y^2$$

since there are $(F/M)n_f$ offspring of a sire which have probabilities of coselection with their $((F/M) - 1)n_f$ half-sibs. This term was not introduced by Wray and Thompson (1990) or Woolliams et al. (1993) because under mass selection the covariance between indices of half-sibs is completely accounted for by the sire, i.e., $\rho_{H,2} = \rho_{m,2}$. Whilst this is true also under selection on an index of (individual and) full-sib records, it is not true for selection using an index of (individual full- and) half-sib records.

Following from the above we can write,

$$\mu_{r(xy),2}^2 + \sigma_{e(xy),2}^2 \approx \frac{1}{4} \left[\left(\frac{Y}{X} \right)^2 + \sigma_{e(xy)}^2 + \sigma_{g(xy)}^2 \right] \quad (15)$$

where,

$$\sigma_{e(xy)}^2 \approx \frac{Y}{X} \frac{T - Y}{T - 1} \frac{X - 1}{X} + \frac{F}{X} n_f (n_f - 1) (\rho_{D,2} - \rho_{x,2}) z_y^2$$

$$+ \frac{F}{X} \left(\frac{F}{X} - 1 \right) n_f^2 (\rho_{H,2} - \rho_{x,2}) z_y^2,$$

and

$$\sigma_{g(xy)}^2 = \left[\frac{F}{X} n_f (n_f - 1) + \frac{F}{X} \left(\frac{F}{X} - 1 \right) n_f^2 \right] \rho_{x,2} z_y^2$$

$$= \left[1 - \frac{X}{Fn_f} \right] b_{xy,2}^2 V[A_{j(x)}].$$

The first term of equation (15) is the mean squared under random selection. The variance $\sigma_{e(xy)}^2$ is hypergeometric sampling variance appropriate under random selection plus covariances due to coselection of sibs which are attributed to correlations arising from the mate of the parent of sex x or to shared estimation errors of family means. The variance $\sigma_{g(xy)}^2$ represents covariances of selection between sibs which are attributable to the parent of sex x . The form of $\sigma_{g(xy)}^2$ in terms of $b_{xy,2}$ will be used later.

The covariances of long-term contributions between male and female descendants [in equation (8) with $t = 2$ instead of ∞] can be written similarly except that there is no hypergeometric sampling term. By noting that each offspring has n_f full-sibs of the opposite sex, it is found that,

$$\mu_{r(xm),2} \mu_{r(xf),2} + \sigma_{r(xm,xf),2} \approx \frac{1}{4} \left[\frac{M}{X} \frac{F}{X} + \sigma_{e(xm,xf)} + \sigma_{g(xm,xf)} \right] \quad (17)$$

where

$$\sigma_{e(xm,xf)} = \frac{F}{X} n_f^2 (\rho_{D,2} - \rho_{x,2}) z_m z_f + \frac{F}{X} \left(\frac{F}{X} - 1 \right) n_f^2 (\rho_{H,2} - \rho_{x,2}) z_m z_f$$

and

$$\sigma_{g(xm,xf)} = \left(\frac{F}{X} \right)^2 n_f^2 \rho_{x,2} z_m z_f = b_{xf,2} b_{xm,2} V(A_{j(x)}).$$

Substituting equation (15) and (17) into (8) [ignoring for the moment that in (8) $t = \infty$] results in

$$\frac{1}{64M} \left\{ 4 + [\sigma_{e(mm)}^2 + \sigma_{g(mm)}^2] + 2 \left[\frac{M}{F} \right] [\sigma_{e(mm,mf)} + \sigma_{g(mm,mf)}] + \left[\frac{M}{F} \right]^2 [\sigma_{e(mf)}^2 + \sigma_{g(mf)}^2] \right\} + \frac{1}{64F} \left\{ 4 + [\sigma_{e(ff)}^2 + \sigma_{g(ff)}^2] + 2 \left[\frac{F}{M} \right] [\sigma_{e(fm,ff)} + \sigma_{g(fm,ff)}] + \left[\frac{F}{M} \right]^2 [\sigma_{e(fm)}^2 + \sigma_{g(fm)}^2] \right\}. \quad (19)$$

Evaluation of equation (8) with $t = \infty$. Finally let us consider the case when $t = \infty$. There are three important aspects to take into account:

- (1) Under random selection (i.e., $b_{xy,t} = 0$) the mean of the long-term contributions of the ancestors is the same for $t = 2$ as for $t = \infty$ [see equation (9)] and so the first terms of equations (15) and (17) remain unchanged from $t = 2$ to $t = \infty$. The method of Woolliams et al. (1993) shows that this holds also under selection. [This is term A3.1 of Woolliams et al. (1993)]
- (2) Under random selection, the collective contributions of the variances of the long-term relationships to rate of inbreeding is increased by a factor of 2 from $t = 2$ to $t = \infty$. This does not imply that each $\sigma_{e(xy)}^2$ doubles, but rather via dispersion of genes the total contribution of $\lambda_{x,t}$ for $t = \infty$ is twice that of $t = 2$, where

$$\lambda_{x,t} = \frac{1}{16X} \left[\left(\frac{X}{M} \right)^2 \sigma_{r(xm),t}^2 + 2 \frac{X}{M} \frac{X}{F} \sigma_{r(xm,xf),t} + \left(\frac{X}{F} \right)^2 \sigma_{r(xf),t}^2 \right] \quad (20)$$

which are the variance and covariance terms in equation (8). This doubling has been observed in simulation when selection is at random. It is also as intrinsic to the predictions of Wray and Thompson (1990) and Woolliams et al. (1993) (in the method of the latter it arises from the summation of terms that occur each generation from $t = 2$ onwards, but weighted by 2^{t-2}). In selected populations, simulation results show that $\lambda_{x,\infty} > 2\lambda_{x,2}$ where $\lambda_{x,2}$ takes the form,

$$\lambda_{x,2} = \frac{1}{64X} \left\{ \left(\frac{X}{M} \right)^2 [\sigma_{g(xm)}^2 + \sigma_{e(xm)}^2] + 2 \frac{X}{M} \frac{X}{F} [\sigma_{g(xm,xf)} + \sigma_{e(xm,xf)}] + \left(\frac{X}{F} \right)^2 [\sigma_{g(xf)}^2 + \sigma_{e(xf)}^2] \right\}. \quad (21)$$

Woolliams et al. (1993) show that prediction of $\lambda_{x,\infty}$ can be achieved by $2\lambda_{x,2}$ plus extra terms considered in point 3 below. Under random selection ($b_{xy,2} = 0$) equations (21) and (20) are identical but, under selection, equation (21) also contains the $b_{xy,2}^2$ terms from equation (15). From Appendix 3 of Woolliams et al. (1993) terms A3.2, A3.6, A3.13, A3.16 and A3.17 sum to $2(\lambda_{m,2} + \lambda_{f,2})$ for mass selection.

(3) The doubling of (21) is insufficient to account for the cumulative selective advantage and extra terms must be included. The

selective advantage from parent to offspring is included in the $\sigma_{g(xy)}^2$ and $\sigma_{g(xm,xf)}$ terms which are functions of $b_{xy,2}$ [equations (16) and (18)]. Investigation of the method of Woolliams et al. (1993) suggests that (by making some assumptions discussed below) the increase in selective advantage from ancestor to descendant can be accounted for by replacing $b_{xy,2}$ in $\sigma_{g(xy)}^2$ and $\sigma_{g(xm,xf)}$ by $b_{xy,\infty}$. Equivalently, this can be achieved by multiplying $\sigma_{g(xy)}^2$ by B_{xy}^2 and $\sigma_{g(xm,xf)}$ by $B_{xm} B_{xf}$ where B_{xy} was defined in equation (12).

Accounting for these points results in the prediction of rate of inbreeding,

$$\Delta F \approx \frac{1}{32M} \left\{ 2 + \sigma_{e(mm)}^2 + B_{mm}^2 \sigma_{g(mm)}^2 + 2 \left(\frac{M}{F} \right) \cdot [\sigma_{e(mm,mf)} + B_{mm} B_{mf} \sigma_{g(mm,mf)}] + \left(\frac{M}{F} \right)^2 [\sigma_{e(mf)}^2 + B_{mf}^2 \sigma_{g(mf)}^2] \right\} + \frac{1}{32F} \left\{ 2 + \sigma_{e(ff)}^2 + B_{ff}^2 \sigma_{g(ff)}^2 + 2 \left(\frac{F}{M} \right) \cdot [\sigma_{e(fm,ff)} + B_{fm} B_{ff} \sigma_{g(fm,ff)}] + \left(\frac{F}{M} \right)^2 [\sigma_{e(fm)}^2 + B_{fm}^2 \sigma_{g(fm)}^2] \right\}.$$

Woolliams et al. (1993) also showed that the rate of inbreeding predicted from long-term contributions should be corrected for contributions from the base population. This correction also applies to equation (22) resulting in a final prediction ΔF where

$$\Delta F = \Delta F(1 + 2\Delta F). \quad (23)$$

Relationship to equation (4) of Woolliams et al. (1993)

Equivalent terms to those in Appendix 3 of Woolliams et al. (1993) have been derived for index selection (Wray and Woolliams, unpublished notes), but their form is complex. The complexity can be traced to the inequality between Cov(breeding value of sire, index or offspring) and Cov(breeding value of dam, index or offspring) i.e., $\tau_m \neq \tau_f$ for I_H . Under mass selection (and I_D) this asymmetry between sexes does not exist. When terms involving τ_m and τ_f are multiplied and accumulated over generations many more types of terms result than in the analogous derivation for mass selection. This is illustrated by derivation of $b_{xy,\infty}$ for index selection in Appendix 2. Approximations invoked for index selection in point 3 above, involve using $\tau_m = \tau_f$ for some product terms. If the equations for $\sigma_{g(xy)}^2$, $\sigma_{g(xm,xf)}$, $\sigma_{e(xy)}$, $\sigma_{e(xm,xf)}$ and B_{xy} [equations (16), but approximating $YX^{-1}(T-Y) \cdot (T-1)^{-1}(X-1)X^{-1}$ to $YX^{-1}(T-Y)T^{-1}$, (18) and (12) are substituted into equation (22) then the following equality results

$$\Delta F \approx \frac{1}{8M} \left\{ 1 + i^2 \left[\frac{M}{F} (\rho_{D,2} - \rho_{m,2}) + \left(1 - \frac{M}{F} \right) \cdot (\rho_{H,2} - \rho_{m,2}) + \rho_{m,2} Q_m^2 \right] \right\} + \frac{1}{8F} [1 + i^2 ((\rho_{D,2} - \rho_{f,2}) + \rho_{f,2} Q_f^2)] \quad (24) - \frac{1}{32T} [4 + (i_m^2 + i_f^2) ((\rho_{D,2} - \rho_{f,2}) + (\rho_{D,2} - \rho_{m,2}) + 2i^2 (\rho_{m,2} Q_m^2 + \rho_{f,2} Q_f^2))],$$

where $Q_x = B_{xy} i_y / i$. Under mass selection (where $\rho_{D,2} = \rho_{m,2} + \rho_{f,2}$, $\rho_{H,2} = \rho_{m,2}$, $Q_x = S_\infty$), equation (24) reduces to equation (4)

of Woolliams et al. (1993) with their $K = i^2(S_{\infty}^2 - 1) + 2i[\frac{1}{2}(i_m c_m + i_f c_f) + i^2 S_{\infty}] (1 - c^2)^{-1}$ approximated to $2i^2(S_{\infty}^2 - 1)$ and where their equation (4) has ignored terms in T except for $4/32T$.

Relationship to the equation of Latter (1959) and Hill (1979)

The form of equation (22) has been chosen for its similarity (and equality when B_{xy} is set to unity) to the discrete generation equation for the prediction of rate of inbreeding of Latter (1959) and Hill (1979) which is based on variance of family size in one generation for random selection. The variance of family size of selected offspring of sex y from parents of sex x , σ_{xy}^2 , is equal to $\sigma_{e(xy)}^2 + \sigma_{g(xy)}^2$ here (and similarly for covariances). They derived their equation from a genetic drift argument, where effective population size is defined by the variance in change of gene frequency. Their expression was derived to account for non-genetic differences in fecundity and viability of offspring, $\sigma_{g(xy)}^2 = 0$, rather than for selection on a heritable trait, although it has been used as such (e.g., de Vries et al. 1990; Wray et al. 1990). The two-generation Latter-Hill equation proposed by Wray et al. (1990) is expected to be approximately equal to equation (22) but with $B_{xy} = b_{xy,3}/b_{xy,2}$. The Latter-Hill equation ignores some higher-order terms which may be approximately incorporated through the correction of equation (23). Like the equation (4) of Woolliams et al. (1993), equation (22) could be rewritten in the form of the (one-generation) Latter-Hill prediction of inbreeding plus a term describing the proliferation of lines from superior ancestors at the expense of their inferior contemporaries.

Relationship with the equation of Robertson (1961)

Equation (22) can also be related to the prediction of Robertson (1961) for populations of full-sib families ($M = F$). Understanding of his method has been hindered by an anomaly in the derivation whereby the interpretation of the N used changes from $N =$ number of full-sib families (therefore the number of parents is $2N$) to N is the number of parents (Felsenstein 1989). However, the (one-generation) result can be derived using the method presented in Latter (1959). Robertson's prediction for one-generation can be obtained by setting $M = F = N$ in the Latter-Hill equation in which Poisson distribution of family size and sampling with replacement are assumed. (Also there is assumed to be no environmental correlation between full-sibs so that $\rho = \rho_D = \rho_m + \rho_f$). Robertson argued for a two-fold increase in selective advantage from generation 2 to infinity ($B = 2$), but which Wray and Thompson (1990) argued should be $B = B_{xy}$ in the notation of this paper. Robertson's prediction is,

$$\Delta F = \frac{1}{4N}(1 + B^2 i^2 \rho),$$

which based on a more thorough theoretical derivation equation (22) with $M = F$ reduces to,

$$\Delta F = \frac{1}{4N} \left[1 + \frac{1}{2}(1 + B^2) i^2 \rho \right].$$

More accurate prediction of coselection sibs

The use of $z_x z_y \rho$ [e.g., in $f_{s_{xy,2}}$ and $h_{s_{xy,2}}$ in equation (14)] is a first-order approximation to the probability of coselection of a pair of sibs of sexes x and y over and above that due to chance alone. This can be more-accurately predicted using the approximation of Mendell and Elston (1974),

$$\left\{ \Phi \left[\frac{i_x \rho - v_y}{(1 - k_x)^{1/2}} \right] p_x - p_x p_y \right\}.$$

This expression is well-defined for $x = y$. However, when $x \neq y$, whilst the probability is symmetric in x and y , the expression is not and both forms are approximations to it. Mendell and Elston (1974) show that accuracy decreases with i so the preferred form has $x = m$ and $y = f$. Under index selection when correlations between sib indices can become very high, the use of this more-accurate prediction of coselection of sibs is important.

When $x = y$, both first-order and second-order approximations to the probability should be multiplied by $((Y-1)/Y)(T/(T-1))$ in an attempt to account for selection without replacement.

Populations in which family size is large relative to the number selected

In the predictions of variance of long-term contributions (or variance of one-generation family size) discussed above, variances have been increased over and above random selection by considering coselection of sibs. This has been calculated as a probability of selection of a pair of sibs multiplied by the possible number of pairs available for selection, without imposition of a constraint of total number selected. In general, this approximation is good (see Results section, Table 2), but when family size available for selection is greater than the number selected, for example $(F/M)n_f > M$, then highly-inflated probabilities of coselection and variances can arise, particularly when the correlation between selection criterion of sibs is high. At the extreme, if $\rho_H = 1$, then all M males will be selected from a single half-sib family. Wray et al. (1990) discussed this problem and for these situations proposed the use of p'_m instead of p_m , where $p'_m = (1 - \rho_H)p_m + \rho_H(F/M)n_f/T$. The full impact of this approximation affects several of the equations presented in this paper and their adapted form is given in Appendix 3.

Simulation

Predictions from equation (23) are compared to rates of inbreeding observed from simulation. Simulations for mass selection are those presented in Wray and Thompson (1990) based on 100 replicates. For index selection, simulations are similar except that selection is based on either I_H or I_D . Populations have $M = 20$ males, $F = 20, 40, 200$ females with $n_f = 3, 6$ offspring of each sex per dam. Heritabilities considered are $h^2 = 10^{-6}, 0.1, 0.2, 0.4, 0.6, 0.99, \sigma_{P,0}^2 = 1$ and common environment variance $\sigma_c^2/\sigma_p^2 = 0$. Heritability values close to zero and close to unity have been investigated so that the predictions can be tested at the extremes where it is possible to postulate the way in which selection is operating. When heritability is exactly zero, index weights are null and selection is at random. But when heritability is close to zero ($h^2 = 10^{-6}$) selection on I_H (or I_D) is close to selection on the family mean since the correlation between sib indices is high. When $h^2 = 0.99$, selection using any of the three indices will result in selection of the same individuals. Other populations simulated have $M = F = 20$, with $n_f = 3, 6, 12, 20$ under mass selection for traits with $h^2 = 10^{-6}, \sigma_{P,0}^2 = 1$ but with $\sigma_c^2 = 0.00, 0.20, 0.60, 1-10^{-6}$. Within the simulations many statistics are calculated which are checked with predictions. These include variances of breeding values of selected ancestors (about the mean of the selected group), correlations between selection criteria of sibs (born in generation 2, calculated by analysis of variance), probabilities of coselection of sibs, variance of family size from parents (born in generation 1) to offspring, $b_{xy,t}$, total sums of squares of long-term contributions and rates of inbreeding. Rates of inbreeding presented are the average of those observed from generations 5 to 14. Simulation results are the average of 1000 replicates for $F = 20, 40$ and 500 replicates for $F = 200$. An example calculation is given in Appendix 4.

Results

Variances and correlations

Predicted values and prediction errors of variances of true breeding values of selected individuals born in generation 1 about the mean of the selected group [$V(A_{j(x)})$] are presented in Table 1. Predictions agree well with simulated values (maximum error 10%). In comparison, calculation of the variance of breeding values about the unconditional mean [$V(A_{j(x)}^*)$] can lead to overestimation by as much as 20% (data not shown). Correlations between full and half-sib index values ($\rho_{D,2}$ and $\rho_{H,2}$) are also presented in Table 1; predicted values are those described in the notation section using $\sigma_{Ax,1}^2$ and calculated by Appendix 2. Predictions of correlations are also accurate (maximum standard error of simulations is 0.007) although correlations of 0.8 or greater tend to be underpredicted. Correlations shown are between sibs born in generation 2; these correlations may be substantially lower (particularly for high h^2) than correlations between sibs born in generation 1 (before selection).

Predicted values and prediction errors of variances of family size (of offspring born in generation 2 from male parents) are presented in Table 2. Predictions of (co)variances of family size from female parents show smaller prediction errors (data not shown). Probabilities of coselection were also examined but these show a similar pattern to the variances of family size. There is a tendency to overpredict (co)variances of family size from male parents when selection uses I_H and in situations where n_f and ρ_H are high (i.e., h^2 low). The overprediction becomes particularly acute in situations where $(F/M)n_f > M$ and ρ_H is high. Predictions using p'_m and the equations of Appendix 3 are also presented, which do remarkably well given that the adjustments are based on heuristic arguments.

Expected long-term contributions

Predicted values and prediction errors of $b_{mm,2}$ and B_{xy} [equations (11) and (12)] are presented in Table 3. When selection uses I_D (or I_P , data not shown) the simulation results are in good agreement with the theoretical result that B_{xy} is independent of x , the sex of the ancestor (since $\tau_m = \tau_f$). This is not true for selection using I_H , where for hierarchical populations, breeding values of male ancestors are more highly correlated to their offspring's index values than are female ancestors ($\tau_m > \tau_f$). If this difference is ignored and an average τ is used when selection is on I_H then serious errors in the prediction of rate of inbreeding and its components arise (data not shown). Simulation results show that the increase in long-term contributions from generation 2 to ∞ is greater for female ancestors ($B_{fy} > B_{my}$),

this is expected from evaluation of equation (12) which can be shown to be a function of τ/τ_x . For two populations with the same structure and heritability, but where selection has used different indices, it is found that in the population where $b_{xy,2}$ is higher, then B_{xy} is lower. The predictions of $b_{xy,2}$ are generally good (shown only for $b_{mm,2}$), although they tend to underpredict when $h^2 = 0.99$. For populations where $(F/M)n_f > M$, predictions of $b_{mm,2}$ are too high. Prediction of $b_{mm,2}$ using the results which depend on p'_m give satisfactory predictions. Prediction errors in B_{xy} are found to be robust compared to prediction errors in $b_{xy,2}$ and $b_{xy,\infty}$.

Rates of inbreeding

Predicted values and prediction errors for rates of inbreeding are presented in Table 4. In some simulations, particularly for high h^2 , rates of inbreeding were observed to be somewhat higher in generation 2 (and sometimes 3). For example, when $h^2 = 0.99$, the first round of selection (where selected individuals tend to come from a few good families) results in a high initial rate of inbreeding. In subsequent generations, the rate of inbreeding is less as a result of selection of the best individuals across families that are genetically less variable. However, in all cases investigated, inbreeding reached an approximately steady rate by generation 4 and over the generations included in the average. For selection on I_H , the prediction error of the rate of inbreeding, as calculated from the total sum of squares of long-term contributions, accumulated within the simulation, [equation (6) with the correction for base contributions, equation (23)] is also presented. This demonstrates that under index selection where rates of inbreeding can be much higher than under mass selection, the prediction of the rate of inbreeding via long-term contributions remains appropriate.

Predictions of rate of inbreeding for mass selection are accurate when $F \geq 100$ (maximum error of prediction 7%), but have a tendency to underpredict for $F \leq 40$ (maximum error 6%). This is the same pattern as found by the approximation equation (4) of Woolliams et al. (1993). (Mass selection predictions presented here use the Mendell and Elston (1974) probabilities of coselection, whereas Woolliams et al. used only the first order approximation for their tabulated results.) Predictions are also accurate for selection on I_D but with a tendency to overpredict at low h^2 (high ρ_F) (which partially reflects overprediction of one-generation variance of family size, data not shown), with maximum errors of 6% for $n_f = 3$ and 13% for $n_f = 6$. The predictions for I_H are accurate providing that h^2 is greater than 0.2 and $M > (F/M)n_f$, but may overestimate otherwise. The overprediction can be as much as 114% ($M = 20$, $F = 200$, $n_f = 6$,

Table 1. Variance of breeding values of selected sires, $V[A_{I(m)}]$, and dams, $V[A_{I(f)}]$, and correlations between index values of full sibs, $\rho_{F,2}$ and half-sibs, $\rho_{H,2}$ (born generation 2). Predicted (P), predicted-simulated (P-S) and percentage error [(100*(P-S)/S)] values for populations with $M = 20$ mates, F females and n_f offspring of each sex born/female with selection using indices I_D or I_H

I	F	n_f	h^2	$V[A_{I(f)}]$			$V[A_{I(m)}]$			$\rho_{F,2}$			$\rho_{H,2}$						
				P	P-S	% Error	P	P-S	% Error	P	P-S	% Error	P	P-S	% Error				
I_D	40	3	0.00 ^a																
			0.10	0.083	0.000	0.0	0.078	+0.001	+1.3	0.889	-0.038	-4.1	0.000	0.000	0.000	0.000	0.000	0.0	
			0.40	0.242	0.000	0.0	0.219	-0.002	-0.9	0.852	+0.010	+1.2	0.086	0.086	-0.001	-10.4			
I_H	40	3	0.99	0.273	0.000	0.0	0.195	0.000	0.0	0.197	+0.002	+1.0	0.164	0.164	-0.005	-3.0	0.080	0.080	0.0
			0.00 ^a	0.081	0.000	0.0	0.076	0.000	0.0	0.905	-0.002	-0.2	0.667	0.667	-0.002	-0.3			
			0.10	0.240	+0.013	+5.7	0.217	+0.002	+0.9	0.861	+0.002	+0.2	0.573	0.573	+0.007	+1.2			
		6	0.99	0.273	0.000	0.0	0.195	+0.005	+2.6	0.197	+0.005	+2.6	0.355	0.355	+0.002	+0.6	0.081	0.081	+2.5
			0.00 ^a	0.073	0.000	0.0	0.067	0.000	0.0	0.944	-0.010	-1.1	0.722	0.722	-0.006	-0.8			
			0.10	0.211	+0.006	+2.9	0.194	+0.004	+2.1	0.888	-0.001	-0.1	0.573	0.573	+0.006	+1.1			
	200	3	0.99	0.199	0.000	0.0	0.155	+0.006	+4.0	0.156	+0.003	+2.0	0.330	0.330	-0.012	-3.5	0.068	0.068	+1.5
			0.00 ^a	0.080	+0.001	+1.3	0.067	+0.002	+3.0	0.956	-0.032	-3.2	0.844	0.844	-0.023	-2.7			
			0.10	0.242	-0.002	-0.8	0.198	+0.004	+2.1	0.882	-0.020	-2.2	0.637	0.637	-0.010	-1.5			
		6	0.99	0.280	+0.003	+1.1	0.128	+0.013	+10.1	0.176	+0.008	+4.8	0.330	0.330	-0.006	-1.8	0.054	0.054	+5.9
			0.00 ^a	0.073	+0.001	+1.4	0.060	-0.002	-3.2	0.976	-0.023	-2.3	0.881	0.881	-0.029	-3.1			
			0.10	0.216	+0.003	+1.4	0.182	+0.001	+0.5	0.893	-0.021	-2.1	0.591	0.591	-0.015	-2.5			
		0.99	0.205	0.000	0.0	0.113	+0.009	+8.0	0.143	0.003	+2.1	0.308	0.308	-0.003	-1.0	0.050	0.050	+6.4	

^a $h^2 = 10^{-6}$

Table 2. Variances and covariances of family sizes of selected offspring from male parents. Predicted (P), predicted-simulated (P-S) and percentage error [(100*(P-S)/S)] values for populations with $M = 20$ males, F females and n_f offspring of each sex per female with selection using indices I_D or I_H

I	F	n_f	h^2	σ_{mm}^2 ^a			σ_{mf}^2 ^a			σ_{mmmf}^2 ^a			
				P	P-S	% Error	(^b)	P	P-S	% Error	(^b)	P	P-S
I_D	40	3	0.00 ^c	1.87	+0.02	+1.1	3.09	+0.05	+1.6	1.79	+0.05	+2.9	
				1.88	+0.03	+1.6	3.18	+0.06	+1.9	1.86	+0.07	+3.9	
				1.70	+0.04	+2.4	2.94	+0.10	+3.5	1.62	+0.09	+5.9	
I_H	40	3	0.00 ^c	1.04	-0.07	-6.3	1.78	-0.04	-2.2	0.47	-0.06	-11.1	
				2.91	+0.07	+2.5	4.96	+0.08	+1.6	3.11	+0.09	+3.0	
				2.62	+0.03	+1.2	4.51	-0.10	+2.3	2.81	+0.10	+3.7	
			0.40	1.96	+0.02	+1.0	3.43	0.01	+0.3	1.97	+0.07	+3.7	
				1.04	-0.09	-8.0	1.78	-0.05	-2.7	0.47	-0.07	-13.0	
				6.32	+0.57	+9.9	12.39	+0.63	+5.4	7.73	+0.55	+7.7	
			0.10	5.18	+0.46	+9.7	10.38	+0.53	+5.4	6.48	+0.53	+8.9	
				3.25	+0.18	+5.9	6.79	+0.18	+2.7	3.96	+0.29	+7.9	
				1.23	-0.12	-8.9	2.51	-0.21	-7.8	0.67	-0.18	-21.1	
200	3	0.00 ^c	15.00	+4.62	+44.5	131.34	+5.43	+4.3	19.74	+0.60	+3.1	(-2.1)	
			8.89	+1.84	+26.1	93.57	+2.61	+2.9	17.03	+0.91	+5.6	(-0.6)	
			3.83	+0.38	+11.0	49.67	-0.09	-0.2	9.65	+0.47	+5.1	(+1.2)	
			0.40	1.25	-0.14	-10.0	13.67	-0.62	-4.3	1.67	-0.34	-16.9	(-17.4)
				31.56	+18.83	+147.9	337.30	+23.03	+7.3	49.22	+4.86	+11.0	(-13.3)
				13.67	+5.49	+67.1	194.87	+9.43	+5.1	35.20	+5.20	+17.3	(-7.5)
			0.40	5.28	+1.18	28.8	97.02	+2.83	+3.0	17.41	+2.16	+14.2	(+0.8)
				1.33	-0.14	-9.5	21.10	-1.29	-5.8	2.38	+0.34	+16.7	(+14.2)

^a σ_{xy}^2 and $\sigma_{xmx,f}$ are predicted as $\sigma_{g(xy)}^2 + \sigma_{g(x,y)}^2$ and $\sigma_{g(xm,x,f)}^2 + \sigma_{g(xm,x,f)}$ respectively

^b Percentage errors in parenthesis are achieved when using p_m instead of $p_{m'}$ see section 'Population in which family size is large relative to number selected'

^c $h^2 = 10^{-6}$

Table 3. One-generation selective advantage of male ancestors to their male offspring born in generation 2, $b_{mm,2}$ [equation (11)] and increase in selective advantage from one-generation to long-term for ancestors of each sex to descendants of each sex, B_{xy} [equation (12)]. Predicted (P), predicted-simulated (P-S) and percentage error [(100*(P-S)/S)] values for populations with $M = 20$ males, F females and n_f offspring of each sex/female and with selection using indices I_D or I_H

I	F	n_f	h^2	$b_{mm,2}$			B_{mm}			B_{jm}			B_{mf}			B_{jf}		
				P	P-S	%Error	(^b)	P	P-S	%Error	(^b)	P	P-S	%Error	(^b)	P	P-S	%Error
I_D	40	3	0.10	1.58	+0.03	+1.9	1.42	-0.06	-4.4	1.43	+0.03	+2.1	1.96	+0.09	+4.8	1.96	+0.03	+1.6
				1.30	-0.02	-1.5	1.13	-0.02	-1.7	1.13	-0.04	-3.4	1.55	-0.04	-2.5	1.55	-0.02	-1.3
				0.96	-0.08	-7.6	0.98	-0.01	-1.0	0.98	-0.02	-2.0	1.35	-0.07	-4.9	1.35	-0.15	-10.0
I_H	40	3	0.10	1.94	0.00	0.0	1.31	+0.08	+6.5	1.51	+0.06	+4.1	1.80	+0.12	+7.1	2.07	+0.15	+7.8
				1.43	-0.02	-1.3	1.08	-0.02	-1.8	1.16	-0.05	-4.1	1.49	-0.04	-2.6	1.59	-0.06	-3.6
				1.03	-0.07	-6.8	0.98	-0.04	-3.9	0.98	-0.06	-5.7	1.35	-0.11	-7.5	1.35	-0.14	-9.3
			0.10	3.08	+0.05	+1.6	1.25	+0.05	+4.2	1.42	+0.07	+5.2	1.53	+0.09	+6.3	1.74	+0.09	+5.5
				2.01	+0.04	+2.0	1.06	-0.02	-1.9	1.11	-0.04	-3.4	1.30	+0.01	+0.8	1.37	-0.04	-2.8
				0.99	-0.16	-11.7	1.00	+0.02	+2.0	1.00	0.00	0.0	1.22	0.00	0.0	1.23	-0.07	-5.4
200	3	0.10	4.87	+0.45	+10.2	0.92	+0.01	+1.1	1.60	+0.07	+4.2	1.87	+0.19	+11.3	3.27	+0.14	+4.5	
			2.58	+0.06	+2.4	0.84	-0.01	-1.1	0.84	-0.06	-5.4	1.72	0.00	0.0	2.12	-0.16	-7.0	
			1.45	-0.30	-17.1	0.83	+0.01	+1.2	0.83	0.00	0.0	1.69	-0.30	-15.1	1.70	-0.39	-18.7	
			0.10	6.25	+0.88	+16.4	0.93	-0.03	-3.1	1.46	+0.12	+9.0	1.54	+0.26	+18.8	2.42	+0.05	+2.1
				3.05	+0.20	+7.0	0.88	-0.03	-3.2	1.01	-0.01	-1.0	1.46	+0.07	+5.0	1.68	-0.06	-3.4
				1.65	-0.27	-14.1	0.87	+0.01	+1.2	0.87	0.00	0.0	1.44	-0.10	-6.5	1.45	-0.27	-15.7

^a Percentage errors in parenthesis are achieved when using p_m instead of $p_{m'}$ see section 'Populations in which family size is large relative to number selected'

Table 4. Rate of inbreeding $\times 100$. Predicted (P), predicted-simulated (P-S) and percentage errors [$100*(P-S)/S$] values for populations with $M = 20$ males and F females and n_f offspring of each sex/female and with selection using I_p , I_D or I_H . For selection using I_H , (S*-S) represents the difference between rate of inbreeding calculated from equation (23) using ΔF from equation (6) (where $\sum r^2$ is calculated in the simulation) and rate of inbreeding from inbreeding coefficients (calculated in the simulation)

F	n_f	h^2	I_p			I_D			I_H					
			P	P-S	%Error	P	P-S	%Error	P	S*-S	P-S	%Error	() ^a	
20	3	0.00 ^b	1.05	-0.02	-1.9	2.57	+0.14	+5.8	2.57	-0.03	+0.14	+5.8		
		0.10	1.18	-0.05	-4.0	2.68	+0.07	+2.7	2.68	+0.01	+0.07	+2.7		
		0.20	1.26	-0.07	-5.2	2.58	+0.03	+1.2	2.58	-0.01	+0.03	+1.2		
		0.40	1.36	-0.06	-4.2	2.29	0.00	0.0	2.29	+0.03	0.00	0.0		
		0.60	1.41	-0.09	-6.0	1.99	-0.02	-1.0	1.99	-0.01	-0.02	-1.0		
		0.99	1.37	0.00	0.0	1.38	+0.01	+0.7	1.38	+0.04	+0.01	+0.7		
40	3	0.00 ^b	0.84	+0.01	+1.2	1.92	+0.07	+3.8	2.42	-0.04	+0.14	+6.1		
		0.10	0.97	-0.01	-1.0	2.14	+0.03	+1.4	2.52	0.00	+0.11	+4.6		
		0.40	1.14	-0.04	-3.3	1.89	-0.03	-1.6	2.01	0.00	+0.02	+1.0		
	6	0.00 ^b	0.89	+0.01	+1.1	4.02	+0.46	+12.9	5.28	-0.08	+0.82	+18.4		
		0.10	1.11	-0.06	-5.1	4.37	+0.36	+9.0	5.11	-0.08	+0.62	+13.8		
		0.40	1.42	-0.08	-5.3	3.20	+0.07	+2.2	3.36	+0.04	+0.18	+5.7		
	200	3	0.00 ^b	0.67	+0.01	+1.5	1.16	+0.03	+2.7	4.27	-0.03	+1.34	+45.7	(+20.5)
			0.10	0.81	-0.01	-1.2	1.56	+0.03	+2.0	3.31	-0.01	+0.74	+28.8	(+16.3)
			0.20	0.90	+0.04	+4.7	1.60	-0.01	-0.6	2.58	+0.03	+0.34	+15.2	(+8.0)
6		0.40	0.99	-0.01	-1.0	1.48	-0.03	-2.0	1.84	+0.02	+0.08	+4.5	(+1.1)	
		0.60	0.99	0.00	0.0	1.29	-0.01	-0.8	1.41	+0.01	+0.01	+0.7	(-0.7)	
		0.99	0.83	+0.02	+2.5	0.84	+0.02	+2.4	0.84	+0.30	0.00	0.0	(0.0)	
200	6	0.00 ^b	0.68	0.00	0.0	2.11	+0.18	+9.3	9.65	-0.07	+5.13	+113.5	(+22.6)	
		0.10	0.89	0.02	+2.2	2.81	+0.18	+6.8	5.58	+0.08	+1.90	+51.6	(+16.8)	
		0.20	1.03	0.00	0.0	2.71	+0.10	+3.8	4.04	-0.06	+0.84	+26.3	(+6.9)	
		0.40	1.15	-0.02	-1.7	2.26	+0.07	+3.2	2.66	+0.01	+0.30	+12.7	(+3.4)	
		0.60	1.14	-0.05	-4.2	1.77	+0.02	+1.2	1.89	-0.05	+0.09	+5.0	(0.0)	
		0.99	0.89	+0.02	+2.3	0.90	+0.02	+2.3	0.90	+0.06	+0.02	+2.3	(+2.3)	

^a Percentage errors in parentheses are achieved when using p'_m instead of p_m , see section 'Populations in which family size is large relative to number selected'

^b $h^2 = 10^{-6}$

$h^2 = 10^{-6}$). The use of p'_m leads to improved predictions with maximum errors of 21% for $n_f = 3$ and 23% for $n_f = 6$.

Errors in the prediction of the rate of inbreeding

Errors in the prediction of the rate of inbreeding for index selection are greatest when h^2 is close to zero. In this case equation (22) reduces to the Latter-Hill equation. As such it is independent of any errors in prediction of selective advantage and depends only on the variance of one-generation family size. Prediction of the variance of family size is fairly good when the equations of Appendix 3 are used for populations of large family size relative to the number selected. Indeed, if variances of family size from the simulation are substituted into the Latter-Hill equation then overprediction of observed rate of inbreeding is found; for example for I_H , $F = 200$, $n_f = 6$ the predicted rate of inbreeding using variances from the simulation is 0.0558, which is close to the prediction using predicted variances of 0.0560 and which both overpredict the

observed rate of inbreeding of 0.0452. (Even without the correction used here, equation (23), the Latter-Hill equation still overpredicts at 0.0508.) When $M = F$, index selection for a trait with near zero heritability, is equivalent to mass selection for a trait with near zero heritability and non-zero σ_C^2/σ_p^2 . Further investigation of the problem was conducted with simulations of populations of this type; $\sigma_{(r)xy,t}^2$ and $\sigma_{(r)xm,x,t}^2$ from these are presented in Table 5. According to the theory used in this paper $\lambda_{x,\infty}/\lambda_{x,2} = 2$ where $\lambda_{x,t}$ is defined in equation (20). However, it is found that when σ_C^2/σ_p^2 is high, the ratio is considerably less than 2. In this case selected parents are all chosen from a minimum of families. The problem is found to be more apparent when family size is large relative to the number selected. Indeed, in the most extreme case considered, when $h^2 = 10^{-6}$, $\sigma_C^2/\sigma_p^2 = 1-10^{-6}$ and $M = F = n_f = 20$, all offspring are chosen from a single family and the variance of contributions from ancestors to descendants cannot increase over and above that from parents to offspring. The ratio in Table 5 is shown to be unity. In this example, the observed rate of inbreeding is 0.191

Table 5. Variances of long-term contributions (from simulation) when $h^2 = 10^{-6}$, $\sigma_P^2 = 1$ and σ_C^2 is non-zero from simulation, $M = F = 20$ and with selection using I_P

σ_C^2	$n_f = 3$			$n_f = 6$			$n_f = 12$			$n_f = 20$		
	$\sigma^2_{r(xp),2}$	$\sigma_{r(xm,x)f,2}$	Ratio ^a	$\sigma^2_{r(xp),2}$	$\sigma_{r(xm,x)f,2}$	Ratio ^a	$\sigma^2_{r(xp),2}$	$\sigma_{r(xm,x)f,2}$	Ratio ^a	$\sigma^2_{r(xp),2}$	$\sigma_{r(xm,x)f,2}$	Ratio ^a
0.00	0.16	0.00	1.99	0.20	0.00	2.00	0.22	0.00	2.01	0.23	0.00	2.01
0.20	0.20	0.06	2.00	0.29	0.11	2.00	0.39	0.19	1.96	0.44	0.24	2.05
0.60	0.29	0.19	1.97	0.54	0.42	1.88	0.89	0.75	1.83	1.18	1.02	1.86
1.00 ^c	0.48	0.48	1.81	1.15	1.15	1.60	2.35	2.35	1.34	4.74	4.74	1.00

^a $\sigma_{r(xp),\infty} = \sigma_{r(xm,x)f,\infty}$ when $M = F$.
^b Ratio = $\lambda_{x,\infty}/\lambda_{x,2}$ where $\lambda_{x,t}$ is defined in equation [20]
^c $\sigma_C^2 = 1-10^{-6}$

(after an initial rate of 0.250), identical to the classical result for full-sib mating (Wright 1931). The rate of inbreeding calculated from the observed long-term contributions [equation (6) with correction (23)] is 0.161. Whilst the rate of inbreeding calculated from observed or predicted (assuming family selection) variances of family size using the equation of Hill (1979) [which is equivalent to the prediction using equation (22)] is 0.243 [or 0.361 after correction by equation (23)].

Discussion

The prediction of rate of inbreeding presented here can be divided into three steps:

- (1) Equality of rate of inbreeding calculated by identity of descent to equation (6) plus (23).
- (2) Equality of equation (22) to (6) which includes (i) the assumption (used directly in this paper, but also implicit in Wray and Thompson 1990 and Woolliams et al. 1993) that the contribution of $\lambda_{x,t}$ [equation (21)] to rate of inbreeding increases by a factor of 2 or more from $t = 2$ to $t = \infty$, and (ii) the approximation of terms which describe the contributions of the expected long-term selective advantage when selection is on a heritable trait.
- (3) Prediction of the components of equation (22), i.e., variances of one-generation family size and the long-term selective advantage terms.

Errors in prediction of the rate of inbreeding defined by identity by descent can occur in any of the three steps. The extreme example of $M = F = n_f = 20$, $h^2 = 10^{-6}$, $c^2 = 1-10^{-6}$ with mass selection highlights the first of these errors. In this example, selection is for the best full-sib family and the rate of inbreeding is that appropriate to repeated full-sib mating. It shows that the errors in the prediction of rate of inbreeding are not confined to the situation of selection on a heritable trait but can be concerned with constraints of population structure. The rates of inbreeding defined by squared contributions and by identity by descent may not be identical in all circumstances since approximations are invoked in the proof of equivalence given by Wray and Thompson (1990). However, for the range of breeding programmes investigated in Table 4, which includes some population structures which could be considered extreme for livestock, the equality between identity by descent rate of inbreeding and equations (6) plus (23) is good and the prediction of long-term contributions as a means of predicting the rate of inbreeding remains an appropriate goal.

In some situations, errors in the prediction of rate of inbreeding occur at stage 2, where the ratio $\lambda_{x,\infty}/\lambda_{x,2}$ [equation (21)] is less than 2. In the extreme example

above, of selection of a single full-sib family, this ratio is 1, as the variance of long-term contributions from ancestors to descendants achieves its maximum after a single generation. Since it is at stage 2 that the equality of equation (22) and the Latter-Hill equation is found (when selection is for a non-heritable trait), then errors arising at stage 2 are equally applicable to the Latter-Hill equation. Such errors have already been discussed by Wray et al. (1990) in comparison of the Latter-Hill prediction of inbreeding with the approach using the maximum eigen value from transition matrices (Woolliams 1989). Errors in stage 2 are less easy to detect explicitly when selection is on a heritable trait, but are likely to contribute to the prediction errors observed in Table 4, particularly when the family size available for selection is large relative to the number of parents and when the correlation between selection criteria of sibs is high.

Errors also occur at stage 3, in the prediction of variances of one-generation family size (Table 2) and long-term selective advantage (Table 3) (errors in the latter reflect errors in the former). The prediction of variances of family size assumes selection across families, but in examples when family size prior to selection is large and correlation between sib records is high, selection tends to be for selection of the best families, in which case the variance of family size after selection is less than predicted. A method has been presented to account for selection of half-sib families, but selection of the best full-sib families within half-sib families has been ignored.

For the populations investigated in this paper (Table 4), rates of inbreeding are underpredicted when selection is on phenotype alone, this is expected as the terms presented here are an approximation to the derivation of Woolliams et al. (1993). Under mass selection, the errors in steps 1–3 discussed above, are unlikely to occur for population structures relevant to livestock breeding. However, these errors can become important when selection is on an index which includes records of collateral relatives, ensuring that the correlation between the selection criteria of sibs is high. In these cases, there is a tendency for overprediction of ΔF , which is found despite the fact that (small positive) terms are ignored in the derivation. The correction to the rate of inbreeding for base contributions (equation (23)) sometimes causes a good prediction of ΔF to become an overprediction, the good prediction before the correction can be attributed to compensatory errors.

In summary, whilst the goal of prediction of rate of inbreeding via long-term contributions remains valid, the method presented here to achieve this prediction does not fully account for constraints upon the variance of long-term contributions arising from the population structure. Under most circumstances the absence of

such a constraint is not an issue, but it becomes important when correlations between sib indices are high (> 0.8) and family size is large relative to the number selected.

In many practical breeding programmes, a restriction is placed on the number of offspring selected per full-sib and half-sib family. Such a restriction would influence the prediction of $\sigma_{r(xy),2}^2$, $\sigma_{g(xy)}^2$, $\sigma_{e(xy),2}^2$ and the equivalent covariances, and the linear prediction of number of offspring selected per parent. However, the problems in the assumption of the increase of long-term contributions from $t = 2$ to $t = \infty$ are unlikely to arise.

When selection in a hierarchical population is on an index which includes half-sib records, a sire is more highly correlated to the index values of his offspring than are his mates. This results in an asymmetry between the contributions from male and female ancestors to the evolution of the rate of inbreeding which does not arise under mass selection or when full-sib records alone are included in an index. It is important to account for this asymmetry in the prediction of inbreeding when selection is on I_H . One-generation selective advantage ($b_{xy,2}$) is greater for I_H vs I_D and I_D vs I_P , whereas the increase in selective advantage (B_{xy}) shows the reverse pattern. Thus, one-generation predictions of rates of inbreeding are expected to underpredict the observed rate of inbreeding to a lesser extent when selection uses sib records compared to mass selection (examples are given in Wray 1989). Under mass selection and with no common environmental effects, the correlation between the selection criteria of sibs is entirely of genetic origin, whilst under selection on family indices, the correlation is partly of environmental origin. Thus, the relative contribution to the total inbreeding of the long-term increase of the genetic component must be smaller for selection using family indices (Wray et al. 1990).

Prediction of rates of inbreeding when selection is based on estimated breeding values calculated by BLUP (best linear unbiased prediction) could be achieved using a selection index, an extension of I_H but including estimated breeding values of the sire, dam and other mates of the sire (Wray and Hill 1989). However, overprediction of rates of inbreeding will be expected as BLUP induces even higher correlations between sib indices than I_H . Use of the predictions obtained here for I_H , although not exact, would be an improvement on what has been used in the past.

This paper represents the completion of the second stage towards the joint description of progress and inbreeding in terms of the same predictable parameters. Woolliams et al. (1993) derived terms which describe the expected proliferation of lines and show how inbreeding can be related concisely to these terms in addition to the variance of family size in mass

selection. This paper extends this theory to the important case of index selection and could be adapted to other situations, such as sex-limited traits, non-hierarchical population structures, and variances in physical family size prior to selection (the subject of a later paper). It is now possible for a wide range of circumstances to assess quanta of information for their value in promoting progress and their value in promoting inbreeding.

Appendix 1. Expression for ΔF partitioned for male and female ancestors

Under the assumption of constant rate of inbreeding each generation, Wray and Thompson (1990) showed that the rate of inbreeding can be expressed as,

$$\Delta F \approx \frac{1}{4(M+F)^2} \mathbf{1}^T \mathbf{C}_{1(1,t)}^T \mathbf{C}_{1(1,t)} \mathbf{1} = \frac{1}{4(M+F)^2} \sum_{i=1}^{M+F} r_i^2$$

where $\mathbf{C}_{1(1,t)}$ is a square matrix of order $M \times F$ of relationships between ancestors born in generation 1 and their descendants born in generation t , which result from the Mendelian samplings received by the ancestors. The elements of $\mathbf{C}_{1(1,t)}^T \mathbf{C}_{1(1,t)}$ are all the same when t is large, thus the pre- and post-multiplication by the unity vector $\mathbf{1}$ sums all the elements and division by $(M+F)^2$ gives the average. If the order of ancestors is males then females and the order of descendants is males then females, then $\mathbf{C}_{1(1,t)}$ can be partitioned as,

$$\mathbf{C}_{1(1,t)} = \begin{bmatrix} \mathbf{C}_{mm} & \mathbf{C}_{mf} \\ \mathbf{C}_{fm} & \mathbf{C}_{ff} \end{bmatrix}$$

and $\mathbf{C}_{1(1,t)}^T \mathbf{C}_{1(1,t)}$ can be written as,

$$\begin{bmatrix} \mathbf{C}_{mm}^T \mathbf{C}_{mm} + \mathbf{C}_{fm}^T \mathbf{C}_{fm} & \mathbf{C}_{mm}^T \mathbf{C}_{mf} + \mathbf{C}_{fm}^T \mathbf{C}_{ff} \\ \mathbf{C}_{mf}^T \mathbf{C}_{mm} + \mathbf{C}_{ff}^T \mathbf{C}_{fm} & \mathbf{C}_{mf}^T \mathbf{C}_{mf} + \mathbf{C}_{ff}^T \mathbf{C}_{ff} \end{bmatrix}$$

Since all elements are equal, the four blocks can be averaged separately,

$$\begin{aligned} & \frac{1}{4(M+F)^2} \mathbf{1}^T \mathbf{C}_{1(1,t)}^T \mathbf{C}_{1(1,t)} \mathbf{1} \\ &= \frac{1}{4} \left[\frac{1}{M^2} \mathbf{1}^T (\mathbf{C}_{mm}^T \mathbf{C}_{mm} + \mathbf{C}_{fm}^T \mathbf{C}_{fm}) \mathbf{1} \right. \\ & \quad + \frac{2}{MF} \mathbf{1}^T (\mathbf{C}_{mm}^T \mathbf{C}_{mf} + \mathbf{C}_{fm}^T \mathbf{C}_{ff}) \mathbf{1} \\ & \quad \left. + \frac{1}{F^2} \mathbf{1}^T (\mathbf{C}_{mf}^T \mathbf{C}_{mf} + \mathbf{C}_{ff}^T \mathbf{C}_{ff}) \mathbf{1} \right]. \end{aligned}$$

One reason that all elements of $\mathbf{C}_{1(1,t)}^T \mathbf{C}_{1(1,t)}$ are identical is because the rows within each \mathbf{C}_{xy} are identical (Wray and Thompson 1990), thus \mathbf{C}_{xy} can be represented by X identical column vectors. The i^{th} element of a column vector of \mathbf{C}_{xy} is $\frac{1}{2} r_{i(xy),t}$ where $r_{i(xy),t}$ is the additive genetic long-term contribution between (the Mendelian samplings of) ancestor i of sex x and its descendants of sex y and the $\frac{1}{2}$ is the value of the Mendelian sampling of the ancestors. Thus, equation (7) follows.

Appendix 2

Derivation of $V(A_{j(x)}) = 4\sigma_{A_{x,1}}^2$. $V[A_{j(x)}]$ is the variance of selected individuals born in generation 1 of sex x ($A_{j(x)}^*$) about the mean of all the selected individuals of sex x ($\bar{A}_{j(x)}^*$).

$$\begin{aligned} V[A_{j(x)}] &= V[A_{j(x)}^* - \bar{A}_{j(x)}^*] = V[A_{j(x)}^*] - V[\bar{A}_{j(x)}^*] \\ V[A_{j(x)}^*] &= (1 - k_x \rho^2) \text{ (Pearson 1903), } \rho^2 = \sigma_{t,0}^2 / \sigma_{A,0}^2 \text{ and} \\ V[\bar{A}_{j(x)}^*] &= \frac{1}{X} [V[A_{j(x)}^*] + (X-1) \text{Cov}[A_{j(x)}^*, A_{j'(x)}^*]] \end{aligned}$$

(Woolliams et al. 1993) where

$$\text{Cov}[A_{j(x)}^*, A_{j'(x)}^*] = P_{D(x)}^* C_{D(x)}^* + P_{H(x)}^* C_{H(x)}^*$$

$P_{D(x)}^*$ ($P_{H(x)}^*$) is the probability of two selected individuals being full (half) sibs calculated as the probability of coselection of a pair of sibs (Mendell and Elston 1974) multiplied by the number of sibs,

$$\begin{aligned} P_{D(x)}^* &\approx (n_f - 1) \Phi \left[\frac{\rho_{D,1} i_x - v_x}{(1 - \rho_{D,1}^2 k_x)^{1/2}} \right], \\ P_{H(x)}^* &\approx \left(\frac{F}{M} - 1 \right) n_f \Phi \left[\frac{\rho_{H,1} i_x - v_x}{(1 - \rho_{H,1}^2 k_x)^{1/2}} \right]. \end{aligned}$$

$C_{D(x)}^*$ ($C_{H(x)}^*$) is the covariance between breeding values of full (half) sibs after selection (see Tallis 1964, p228),

$$\begin{aligned} C_{D(x)}^* &= \sigma_{A,0}^2 \left[a_D + \phi[v_x, v_x, \rho_{D,1}] [\rho(\rho - \rho_{D,1} u_D) \right. \\ & \quad \left. + u_D(u_D - \rho_{D,1} \rho)] / (p_x P_{D(x)}^*) + \rho^2 i_x^2 + 2\rho i_x \right. \\ & \quad \left. \cdot \Phi \left[\frac{-v_x(1 - \rho_{D,1})}{(1 - \rho_{D,1}^2)} \right] [u_D v_x - (\rho + u_D) i_x] / P_{D(x)}^* \right] \end{aligned}$$

and $\Phi(v_x, v_x, \rho_{D,1})$ is the frequency of a bivariate normal with correlation $\rho_{D,1}$ where both truncation deviates have value v_x (so $\Phi(v_x, v_x, \rho_{D,1}) = e^{-v_x^2(1 - \rho_{D,1}^2)} / [2\pi(1 - \rho_{D,1}^2)^{1/2}]$). a_D is the correlation between the breeding values of generation 1 full-sibs before selection ($a_D = \frac{1}{2}$) and u_F is the correlation between the breeding value of an individual with the index value of its full-sib (before selection). For index I_H ,

$$\begin{aligned} u_D &= \left\{ \frac{1}{2n} \left[-\beta_1 + \beta_2 \left(1 - \frac{M}{F} \right) + \beta_3 \frac{M}{F} \right] \right. \\ & \quad \left. + \frac{1}{4} \left[\beta_2 \left(1 - \frac{M}{F} \right) + \beta_3 \left(1 + \frac{M}{F} \right) \right] \right\} \sigma_{A,0} / \sigma_{I,0}. \end{aligned}$$

To calculate $C_{H(x)}^*$ replace all subscripts of D for full-sibs by H for half-sibs, $a_H = \frac{1}{4}$ and

$$\begin{aligned} u_H &= \left\{ \frac{1}{2n} \frac{M}{F} (\beta_2 - \beta_2) \right. \\ & \quad \left. + \frac{1}{4} \left[\frac{M}{F} (\beta_3 - \beta_2) + \beta_3 \right] \right\} \sigma_{A,0} / \sigma_{I,0}. \end{aligned}$$

Derivation of $b_{xy,2}$ and $b_{xwy,t}$.

Wray and Thompson (1990) showed that $b_{xy,2}$ can be represented as $b_{xy,2} = n_x \beta_{s2, A_{i(x)}} = n_x \beta_{s2, 12} \beta_{I2, A_{i(x)}}$ where st and It are the selection scores (1 if selected, 0 otherwise) and index scores of individuals of sex y born in generation t , β represent regression coefficients and $\beta_{st, It} \approx z_y / \sigma_{I,t}$. The coefficient $\beta_{I2, A_{i(x)}} = \frac{1}{2} \tau_x$ and thus $b_{xy,2} \approx \frac{1}{2} n_x \tau_x z_y / \sigma_{I,2}$.

Further, Wray and Thompson (1990) showed that,

$$b_{xy,t} = \frac{1}{2} \left(\frac{Y}{M} b_{xm,t-1} + \frac{Y}{F} b_{xf,t-1} + \frac{M}{X} b_{xmy,t} + \frac{F}{X} b_{xfy,t} \right) \quad (\text{A2.1})$$

where $b_{xwy,t}$ is the regression coefficient which accounts for the additional selective advantage of the ancestor to the descendant over and above the selective advantage from the ancestor to the parent of sex w in generation $t-1$.

Similar to the case for $b_{xy,2}$, $b_{xwy,3}$ can be represented as

$$b_{xwy,3} \approx n_w \beta_{s3,13} \beta_{I3,Ai(x)}$$

where the regression coefficients represent regression over and above that already accounted for between x and w , i.e., after accounting for the selection of w (as well as selection of x). The covariance between grandparent of sex x and grandoffspring before accounting for selection of w is $\frac{1}{4} \tau_w V_x$ and

$$\beta_{I3,Ai(x)} \approx \left[\frac{\frac{1}{4} \tau_w V_x - \frac{\frac{1}{2} \tau_x V_x \frac{1}{2} \tau_w \sigma_I^2 (1 - \frac{1}{4} \tau_m k_m - \frac{1}{4} \tau_f k_f) k_w}{\sigma_I^2 (1 - \frac{1}{4} \tau_m k_m - \frac{1}{4} \tau_f k_f)} k_w}{V_x - \frac{(\frac{1}{2} \tau_x V_x)^2}{\sigma_I^2 (1 - \frac{1}{4} \tau_m k_m - \frac{1}{4} \tau_f k_f)} k_w} \right] \quad (\text{A2.2})$$

where $V_x = V(A_{i(x)})$, the variance of the breeding value of the ancestors. This can be approximated as $\beta_{I3,Ai(x)} \approx \frac{1}{4} \tau_w (1 - \tau_x k_w)$. Thus $b_{xwy,3} \approx \frac{1}{4} n_w \tau_w (1 - \tau_x k_w) z_y / \sigma_{I,3}$. Using similar derivations and approximations it can be shown that for $t \geq 4$,

$$b_{xwy,t} \approx \frac{1}{2} n_w \tau_w \frac{1}{2} (1 - \tau_x k) \left[\frac{1}{2} (1 - k\tau) \right]^{t-4} \cdot \frac{1}{2} (1 - \tau k_w) z_y / \sigma_{I,t}$$

In this equation $\frac{1}{2}(1 - \tau_x k)$ is the reduction accounting for the selection of descendants born in generation 2 averaged over males and females (hence the average k) but dependent on the sex of the ancestor. The $\frac{1}{2}(1 - \tau k)$ terms are reductions due to intermediate generations of selection, averaged over both sexes of descendants; hence the term involves average k and average τ . $\frac{1}{2} \tau_w$ is the coefficient of relationship between w and y and $\frac{1}{2}(1 - \tau k_w)$ is the reduction due to selection of sex w averaged over male and females in generation $t-2$.

Thus in summary we can write,

$$b_{xwy,3} \approx \frac{1}{2} n_w \tau_w \frac{1}{2} (1 - \tau_x k_w) z_y / \sigma_{I,3},$$

$$b_{xwy,t} \approx \frac{1}{2} n_w \tau_w c'_x c_w c^{t-4} z_y / \sigma_{I,t}$$

where $c_w = \frac{1}{2}(1 - \tau k_w)$, $c'_x = \frac{1}{2}(1 - \tau_x k)$, $c = \frac{1}{2}(c_m + c_f)$. As in Woolliams et al. (1993) it will be assumed that $\sigma_{I,2}$ will be close to its equilibrium value.

Following Woolliams et al. (1993), equation (A2.1) can be represented in matrix notation by,

$$\mathbf{b}_{x,t} = \mathbf{D} \mathbf{b}_{x,t-1} + \frac{M}{2X} (b_{xmm,t} \mathbf{b}_{xmf,t})^T + \frac{F}{2X} (b_{xfm,t} \mathbf{b}_{xfy,t})^T,$$

where $\mathbf{b}_{x,t} = [b_{xm,t} \ b_{xf,t}]^T$ and \mathbf{D} is $\begin{bmatrix} 1/2 & 1/2M/F \\ 1/2F/M & 1/2 \end{bmatrix}$. Matrix \mathbf{D} describes the dispersion of genes through the population from generation to generation in the absence of selection and has the property of idempotency ($\mathbf{D}^2 = \mathbf{D}$). By analogous derivation to that presented in Appendix 1 of Woolliams et al. (1993) it can be shown that,

$$\mathbf{b}_{x,3} = \frac{1}{2} \sigma_I^{-1} \left[i \tau_x \begin{pmatrix} M/X \\ F/X \end{pmatrix} + d_x \begin{pmatrix} i_m M/X \\ i_f F/X \end{pmatrix} \right]$$

$$\mathbf{b}_{x,t} = \frac{1}{2} \sigma_I^{-1} \left[i (\tau_x + d_x + c'_x d S_{t-4}) \begin{pmatrix} M/X \\ F/X \end{pmatrix} + d_x c'_x c^{t-4} \begin{pmatrix} i_m M/X \\ i_f F/X \end{pmatrix} \right] \quad \text{for } t \geq 4$$

where $S_t = \sum_{j=0}^t c^j$, $d = \frac{1}{2}(\tau_m c_m + \tau_f c_f)$ and $d_x = \frac{1}{2}[\tau_m \frac{1}{2}(1 - \tau_x k_m) + \tau_f \frac{1}{2}(1 - \tau_x k_f)]$. Following from this we can write in summary,

$$b_{xy,2} \approx \frac{1}{2} \frac{Y}{\sigma_{I,2} X} \tau_x i_y$$

$$b_{xy,3} \approx \frac{1}{2} \frac{Y}{\sigma_{I,2} X} (i \tau_x + i_y d_x)$$

$$b_{xy,t} \approx \frac{1}{2} \frac{Y}{\sigma_{I,2} X} (i \tau_x + i d_x + i d c'_x S_{t-4} + i_y c'_x d c^{t-4})$$

$$b_{xy,\infty} \approx \frac{1}{2} \frac{Y}{\sigma_{I,2} X} (i \tau_x + i d_x + i d c'_x S_\infty)$$

where $S_\infty = (1 - c)^{-1} = 2(1 + k\tau)^{-1}$. Substitution of expressions of d , d_x , c'_x and S_∞ into $b_{xy,\infty}$ results in equation (11).

Appendix 3. Adaptation to equations when $(F/M)n_f > M$

The adaptation of equations in this situation affects only males selected from male parents. Therefore, in generation 2, only regression coefficient $b_{mm,2}$ is affected, resulting in $b_{mm,2} \approx \frac{1}{2\sigma_{I,2}} \tau_m i'_m$. However, in subsequent generations all $b_{xy,t}$ are affected via $b_{xmm,t}$ terms [see equation (A2.1)], and it can be shown that $b_{xy,\infty} \approx \frac{1}{2\sigma_{I,2}} \frac{Y}{X} (i_{(x)} \tau_m + G_x + G'_x \eta_\infty)$ where i'_m is the selection intensity appropriate to p'_m , $i_{(x)} = i$ if $x = f$ and $\frac{1}{2}(i'_m + i_f)$ if $x = m$. $G = \frac{1}{2}(\varepsilon i_m + d i_f)$, $G_x = \frac{1}{2}(\varepsilon_x i_m + d_x i_f)$, $\varepsilon = \frac{1}{2}(\tau_m c_m i'_m / i_m + \tau_f c_f)$ and $\varepsilon_x = \frac{1}{2}(\tau_m \frac{1}{2}(1 - k_m \tau_x) i'_m / i_m + \tau_f \frac{1}{2}(1 - k_f \tau_x))$.

Changes to the variances required for equation (22) are

$$\sigma_{(e)mm}^2 = \left(1 - \frac{1}{M}\right) \left(\frac{T-M}{T-1}\right) + n_m(n_f - 1) U_{(e)fm} + \frac{1}{p_m} \left(\frac{1}{p'_m} - n_f\right) U'_{(e)mm}$$

$$\sigma_{(g)mm}^2 = \left[\frac{1}{p'_m} \left(\frac{1}{p'_m} - n_f\right) + n_m(n_f - 1)\right] U'_{(g)mm}$$

$$\sigma_{(e)mm,mf} = n_m n_f U_{(e)fm,ff} + n_m \left(\frac{1}{p'_m} - n_f\right) U'_{(s)mm,mf}$$

$$\sigma_{(g)mm,mf} = \frac{n_m}{p'_m} U'_{(g)mm,mf}$$

where the U terms represent the Mendell and Elston (1974) probabilities of coselection of sibs, the ' implying the use of p'_m , i'_m , k'_m and v'_m , and the use of $p_f + p_{Fe}$, p_{He} or p_m are implied in $U_{(s)fm}$ (or $U_{(s)fm,ff}$), $U_{(s)mm}$ (or $U_{(s)mm,mf}$) of $U_{(t)mm}$ (or $U_{(t)mm,mf}$) respectively. These terms are equivalent to those in the main text if p'_m is replaced by p_m .

Appendix 4. Example

For an example population $M = 20$, $F = 200$, $n_f = 6$, $h_0^2 = 0.4$ and $\sigma_c^2/\sigma_p^2 = 0$ with selection using I_H , then terms described in the first section of the methods are $n_m = 60$, $T = 1200$, $p_m = 0.0167$, $p_f = 0.1667$, $v_m = 2.129$, $v_f = 0.967$, $z_m = 0.041$, $z_f = 0.250$, $i_m = 2.485$, $i_f = 1.499$, $i = 1.992$, $k_m = 0.885$, $k_f = 0.797$ and $k = 0.841$. If $\sigma_{P,0}^2 = 1$, then $\sigma_{A,0}^2 = 0.4$, $\sigma_{Am,0}^2 = \sigma_{Af,0}^2 = 0.1$ and $\sigma_{Aw}^2 = 0.2$. Index parameters are $\beta_1 = 0.250$, $\beta_2 = 0.700$, $\beta_3 = 0.957$, $\sigma_{T,0}^2 = 0.226$, $\tau_m = 0.957$, $\tau_f = 0.726$ and $\tau = 0.841$. Calculation of $V[A_{j(x)}]$ described in Appendix 2 requires the following: $\rho = 0.752$, $\rho_{H,1} = 0.436$, $\rho_{D,1} = 0.779$, $P_{H(m)}^* = 0.126$, $P_{H(f)}^* = 0.367$, $P_{D(m)}^* = 0.388$, $P_{D(f)}^* = 0.610$, $u_H = 0.328$, $u_D = 0.586$, $C_{H(m)}^* = 0.007$, $C_{H(f)}^* = 0.013$, $C_{D(m)}^* = 0.058$ and $C_{D(f)}^* = 0.060$. From these $V[A_{j(m)}^*] = 0.200$, $V[A_{j(f)}^*] = 0.220$, $V[A_{j(m)}^*] = 0.018$, $V[\bar{A}_{j(f)}^*] = 0.003$ and $V[A_{j(m)}] = 0.182$, $V[A_{j(f)}] = 0.217$. Then, since $\sigma_{Ax,1}^2 = \frac{1}{4} V[A_{j(x)}]$, $\sigma_{Am,1}^2 = 0.046$, $\sigma_{Af,1}^2 = 0.054$ and $\sigma_{A,2}^2 = 0.300$. From equation (1) $\sigma_{T,2}^2 = 0.152$ and from equation (11) $b_{mm,2} = 3.05$, $b_{mf,2} = 18.41$, $b_{ff,2} = 0.23$, $b_{f,2} = 1.40$, $b_{mm,\infty} = 2.69$, $b_{mf,\infty} = 26.85$, $b_{ff,\infty} = 0.23$, $b_{f,\infty} = 2.34$, B_{xy} are reported in Table 3. For offspring born in generation 2, correlations are: $\rho_{H,2} = 0.308$, $\rho_{D,2} = 0.671$, $\rho_{m,2} = 0.275$, $\rho_{f,2} = 0.188$ from equations (3)–(5). Variances and covariances of contributions from parents of sex x which are not attributable to the selective advantage conferred by the parent of sex x are: $\sigma_{(e)mm}^2 = 2.45$, $\sigma_{(e)mf}^2 = 28.52$, $\sigma_{(e)mm,mf}^2 = 4.37$, $\sigma_{(e)fm}^2 = 0.21$, $\sigma_{(e)ff}^2 = 2.20$ and $\sigma_{(e)fm,ff}^2 = 0.30$ [equations (16) and (18)]. Variances and covariances of contributions from parent to offspring which are attributable to the selective advantage of the parent are: $\sigma_{(g)mm}^2 = 2.84$, $\sigma_{(g)mf}^2 = 68.52$, $\sigma_{(g)mm,mf}^2 = 13.04$, $\sigma_{(g)fm}^2 = 0.01$, $\sigma_{(g)ff}^2 = 0.38$ and $\sigma_{(g)fm,ff}^2 = 0.08$ [equations (16) and (18)]. These variances and covariances when summed, e.g., $\sigma_{(e)xy}^2 + \sigma_{(g)xy}^2$, give the one generation variances of family size presented in Table 2. The rate of inbreeding as evaluated by equation (22) is $\Delta F = 0.0254$ and with correction [equation (23)] $\Delta F = 0.0267$. In this example $M < n_m$ and therefore the adaptations presented in Appendix 3 apply. Thus, $p'_m = 0.0269$, $i'_m = 2.307$, $i_{(m)} = 1.903$, $i_{(f)} = 1.992$, $v'_m = 1.928$, $k'_m = 0.875$. For $b_{xy,t}$, $G_m = 0.077$, $G_f = 0.156$, $G = 0.116$, $\varepsilon_m = 0.155$, $\varepsilon_f = 0.315$ and $\varepsilon = 0.235$ resulting in $b_{mm,2} = 2.83$, $b_{mm,\infty} = 2.57$, $b_{mf,\infty} = 25.71$, $b_{fm,\infty} = 0.23$, $b_{ff,\infty} = 2.33$. Variances and covariances of contributions, which differ from those above, are $\sigma_{(e)mm}^2 = 2.21$, $\sigma_{(e)mm,mf}^2 = 3.40$, $\sigma_{(g)mm}^2 = 3.47$ and $\sigma_{(g)mm,mf}^2 = 11.90$. Finally, $\Delta F = 0.0233$ and with correction (23) $\Delta F = 0.0244$.

Acknowledgements. The authors gratefully acknowledge funding from the Australian Wool Corporation (N.R.W.) and the Ministry of Agriculture, Fisheries and Food, U.K. (J.A.W. and R.T.). Drs. P. Visscher and J. van Arendonk and an anonymous referee are thanked for their comments on the manuscript.

References

- Felsenstein J (1989) Alan Robertson's contributions to population genetics. In: (eds) Hill WG, Mackay TFC Evolution and animal breeding. C.A.B International, Wallingford, Oxon, England, pp 3–11
- Hill WG (1979) A note on effective population size with overlapping generations. *Genetics* 92: 317–322
- Latter BDH (1959) Genetic sampling in a random mating population of constant size and sex ratio. *Aust J Biol Sci* 12: 500–505
- Mendell NC, Elston RC (1974) Multifactorial traits: genetic analysis and prediction of recurrence risk. *Biometrics* 30: 41–57
- Robertson A (1961) Inbreeding in artificial selection programmes. *Genet Res* 2: 189–194
- Tallis GM (1961) The moment generating function of the truncated multi-normal distribution. *J Roy Stat Soc B* 23: 223–229
- Vries AG de, Van der Steen HAM, De Roo G (1990) Effects of family size in selection and testing in a closed dam line of pigs. *Livest Prod Sci* 24: 47–63
- Woolliams JA (1989) Modification of MOET nucleus breeding schemes to improve rates of genetic progress and decrease rates of inbreeding in dairy cattle. *Anim Prod* 49: 1–14
- Woolliams JA, Wray NR, Thompson R (1993) Prediction of long-term contributions and inbreeding in populations undergoing mass selection. *Genet Res* 62: 231–242
- Wray NR (1989) Consequences of selection in finite populations with particular reference to closed nucleus herds of pigs. PhD thesis, University of Edinburgh
- Wray NR, Thompson R (1990) Prediction of rates of inbreeding in selected populations. *Genet Res* 55: 41–54
- Wray NR, Woolliams JA, Thompson R (1990) Methods for prediction rates of inbreeding in selected populations. *Theor Appl Genet* 80: 503–512
- Wright S (1931) Evolution in Mendelian populations. *Genetics* 16: 97–159