

# Age-specificity in conditional ring-recovery models

CHIARA MAZZETTA

We consider the case of age-specific ring-recovery data obtained only from recovered individual birds and modelled by conditioning a multinomial distribution on the recovery. These models may be appealing when the information about the numbers of marked individuals is missing but they are often analysed by ignoring a large set of nuisance parameters, the recovery rates. We investigate what are the consequences of this type of conditioning by relating the age-time specific structure of recovery rates to the estimation of survival.

**Key words:** Bayesian estimation; Conditional multinomial; MCMC; Ring-recovery; Survival rates.

---

Chiara Mazzetta is a Research Associate, Institute of Mathematics and Statistics, University of Kent, Canterbury, Kent CT2 7NF, England. email: C.Mazzetta@kent.ac.uk

# 1 INTRODUCTION

Ring-recovery data are an important source of information for estimating and monitoring survival rates that, either alone or combined with other demographic parameters such as productivity rates, provide an insight into the population dynamics of species. The survey method consists in capturing a bird, placing a lightweight metal ring on its leg, for unique identification, and then releasing the individual; when some of the marked birds are recovered dead, in the following years, the number on the ring provides access to data that is used to estimate age-specific and time-specific survival rates, as well as recovery rates. In the UK, ring-recovery data for birds are maintained by the British Trust for Ornithology (BTO) through a ringing scheme that is part of the Integrated Population Monitoring Program and involves almost 2000 trained ringers who are mainly volunteers, as described in Baillie et al. (2006). In this paper we consider birds ringed in the UK during the breeding season and recovered during the whole year. We focus on a particular ringing period to better define the length of time for which birds are 'at risk' of recovery and to exclude those birds ringed in winter which may breed outside the UK. If we denote by  $R_i^d$  the number of birds ringed at age  $d$  in year  $i$  and with  $m_{ij}^d$  those among  $R_i^d$  that were recovered in year  $j$ , then for  $i = 1, \dots, I$  and  $j = 1, \dots, J$ , with  $J \geq I$ , we can build a matrix of data like the one in Table 1, where  $m_{iJ+1}^d = R_i^d - m_{ij}^d$  denotes the number of birds ringed in year  $i$  and never recovered, while  $d = 1$  or  $a$  denotes the age at ringing for juveniles and adults, respectively. Recovery data are typically modelled by assuming that, for each year  $i$ , the counts  $(m_{i1}^d, \dots, m_{iJ+1}^d)$  follow a multinomial distribution  $\text{Mult}(R_i^d; \psi_{i1}^d, \dots, \psi_{iJ+1}^d)$ , where  $\psi_{ij}^d$  is the probability that a bird ringed at age  $d$  in year  $i$  is recovered in year  $j$ , while  $\psi_{iJ+1}^d = 1 - \sum_{j=1}^J \psi_{ij}^d$ . The availability of age-specific ringing data is very important if we want to implement multiple-age models such as the one described above. For example, Williams et al. (2002) discuss the difficulty of separating out survival and recovery rates estimates

for juveniles from those of adults, when only juvenile birds are ringed. Another important issue is related to the study design and to the fact that sometimes the number of birds ringed may be unknown or the available figures may be unreliable. For example, the BTO started only recently to computerise ring-recovery data so that the quantities  $R_i^d$  are readily available for all species for the last two or three years. Previously, the information had been collected only for those species for which other indicators suggested an increased concern for their conservation status. Thus, sometimes we may know the total numbers of birds ringed - as nestlings, juveniles or adults - but not the proportion of those ringed during the breeding season or how that proportion was split between the age classes we are considering. In these cases, the data consist only of records of recovered individuals and are analysed by conditioning the multinomial distribution on the recovery of birds and often by assuming that the recovery rate is constant across age groups and time periods. These are very strong assumptions that ignore the structure of the recovery rates and are likely to produce biased survival estimates. Burnham (1990) considers the case where only juveniles are ringed and investigates the potential for a loss of power in goodness of fit tests and the potential for a loss of efficiency of survival estimates when the ringing numbers are not known. In his comparisons between conditional and unconditional models, the author assumes that recovery rates are not varying with age and mentions in the discussion that if this is not the case then the conditional survival estimates can be very biased. Freeman & Morgan (1992) consider again the case where only juveniles are ringed but with the numbers ringed available. The authors do not fit directly a model with age-varying reporting rates but use simulations to investigate what happens to survival rates when juvenile and adult recovery rates are different but are assumed to be equal. Catchpole et al. (2004) analyse a similar type of bias arising in survival rates estimates when using conditional capture-recapture models that do not account for a different structure in recovery and recapture rates. The authors do not consider the general case but

show how the bias occurs after specifying a particular type of relationship between the rates and some covariates. In this work we study what happens to survival rates when the information  $R_i^d$  is missing and it is assumed that the recovery rates are constant across age and time, thus explaining the findings of Freeman & Morgan (1992) and Catchpole et al. (2004) in more general terms. We compare the behaviour of survival estimates obtained from the conditional model described above with that of survival estimates obtained from a general unconditional model, where recovery rates are allowed to vary with both age and time. In order to avoid the difficult task of comparing directly the two likelihoods, we compare the behaviour of the survival estimates by looking at their acceptance functions within a Metropolis-Hastings algorithm with single components updates. In §2 and §3 we describe the structure of the unconditional and conditional model, respectively; in §4 we show how the bias is related to age and time variation in recovery rates; in §5 we present results from an application where we use field data; and in §6 we conclude with a discussion.

[Table 1 about here.]

## 2 MODEL WITH KNOWN NUMBERS OF MARKED INDIVIDUALS

As mentioned in the introduction, when the numbers  $R_i^d$  of birds that have been ringed are known the counts  $(m_{i1}^d, \dots, m_{iJ+1}^d)$  can be modelled with a multinomial distribution  $\text{Mult}(R_i^d; \psi_{i1}^d, \dots, \psi_{iJ+1}^d)$ . Each probability  $\psi_{ij}^d$  depends on the age of the bird, on its survival history over time and on the probability of being recovered in a specific year. We allow both survival and recovery to vary with age and time and we denote by  $\phi_{1i}$  the probability that a juvenile survives year  $i$ ; by  $\phi_{aj}$  the probability that an adult survives year  $j$ ; and by  $\lambda_{1i}$  and  $\lambda_{aj}$  the corresponding probabilities of being recovered, for  $i = 1, \dots, I$  and  $j = 1, \dots, J$ . Throughout this work we assume

that an individual can only be recovered in the same year of its death. These assumptions are appropriate to describe most of the species studied by the BTO, and are also commonly used for other taxa, but can be extended to accommodate different biological characteristics, for example when more than two age classes are required. The multinomial probabilities are modelled as:

$$\psi_{ij}^d = \begin{cases} (1 - \phi_{di})\lambda_{di} & i = 1, \dots, I \quad j = i \quad d = 1, a \\ \phi_{di} \prod_{k=i+1}^{j-1} \phi_{ak}(1 - \phi_{aj})\lambda_{aj} & i = 1, \dots, L^* \quad j = i+1, \dots, J \quad d = 1, a \end{cases} \quad (2.1)$$

with  $\prod_{k=i+1}^{j-1} \phi_{ak} = 1$  when  $j = i+1$  and  $L^* = I - I_{[J=I]}$ , where  $I_{[J=I]}$  is an indicator function for the case  $J = I$  of square matrices. From (2.1) we can build the likelihood for the ringing data:

$$L(\mathbf{m}^d \mid \boldsymbol{\psi}) \propto \Delta^d(\boldsymbol{\psi}) \prod_{i=1}^I \prod_{j=i}^J (\psi_{ij}^d)^{m_{ij}^d} = \Delta^d(\boldsymbol{\psi}) \prod_{i=1}^I (\psi_{ii}^d)^{m_{ii}^d} \prod_{i=1}^{L^*} \prod_{j>i}^J (\psi_{ij}^d)^{m_{ij}^d} \quad (2.2)$$

where  $\Delta^d(\boldsymbol{\psi}) = \prod_{i=1}^I (1 - \psi_{ii}^d)^{R_i^d - m_{ii}^d}$  and a dot denotes summation over all values of an index. If we express this likelihood in terms of the survival and recovery parameters then, for example, by considering the juveniles we obtain:

$$\begin{aligned} L(\mathbf{m}^1 \mid \boldsymbol{\phi}, \boldsymbol{\lambda}) &\propto \Delta^1(\boldsymbol{\phi}, \boldsymbol{\lambda}) \prod_{i=1}^I [(1 - \phi_{1i})\lambda_{1i}]^{m_{ii}^1} \prod_{i=1}^{L^*} \prod_{j>i}^J \left[ \phi_{1i} \left( \prod_{k=i+1}^{j-1} \phi_{ak} \right) (1 - \phi_{aj})\lambda_{aj} \right]^{m_{ij}^1} \\ &\propto \Delta^1(\boldsymbol{\phi}, \boldsymbol{\lambda}) \left[ \prod_{i=1}^{L^*} \phi_{1i}^{S_{11}^{(i)}} \prod_{i=1}^I (1 - \phi_{1i})^{m_{ii}^1} \right] \left[ \prod_{j=2}^{J-1} \phi_{aj}^{S_{11}^{(j)}} \prod_{j=2}^J (1 - \phi_{aj})^{S_{21}^{(j)}} \right] \\ &\quad \times \left[ \prod_{i=1}^I \lambda_{1i}^{m_{ii}^1} \right] \left[ \prod_{j=2}^J \lambda_{aj}^{S_{21}^{(j)}} \right] \end{aligned} \quad (2.3)$$

with statistics:

$$S_1(i) = \sum_{j=i+1}^J m_{ij}^1, \quad S_{11}(j) = \sum_{i=1}^{h^*} \sum_{l=j+1}^J m_{il}^1 \quad \text{and} \quad S_{21}(j) = \sum_{i=1}^{h^*} m_{ij}^1 \quad (2.4)$$

and  $h^*$  equal to  $\min\{I, j-1\}$ . If the matrix  $\mathbf{m}^1$  is square then we have  $n = I(I+1)/2$  cells,  $4I$  parameters for juvenile birds and  $2I$  parameters for adult birds. At the end of the series, parameters  $(\phi_{1I}, \lambda_{1I})$  and  $(\phi_{aJ}, \lambda_{aJ})$  appear only as a product and are not individually identifiable; this can be assessed by implementing, for example with Mathematica, the method proposed by Catchpole & Morgan (1997), see also Catchpole et al. (1998). When considering birds ringed as adults, the likelihood can be re-arranged in a similar way:

$$\begin{aligned}
L(\mathbf{m}^a \mid \boldsymbol{\phi}, \boldsymbol{\lambda}) &\propto \Delta^a(\boldsymbol{\phi}, \boldsymbol{\lambda}) \prod_{i=1}^I [(1 - \phi_{ai}) \lambda_{ai}]^{m_{ii}^a} \prod_{i=1}^{L^*} \prod_{j>i} \left[ \phi_{ai} \left( \prod_{k=i+1}^{j-1} \phi_{ak} \right) (1 - \phi_{aj}) \lambda_{aj} \right]^{m_{ij}^a} \\
&\propto \Delta^a(\boldsymbol{\phi}, \boldsymbol{\lambda}) \left[ \phi_{a1}^{S_{a1}} (1 - \phi_{a1})^{m_{11}^a} \right] \left[ \prod_{j=2}^{J-1} \phi_{aj}^{S_{1a}(j)} \prod_{j=2}^J (1 - \phi_{aj})^{S_{2a}(j)} \right] \\
&\quad \times \lambda_{a1}^{m_{11}^a} \left[ \prod_{j=1}^J \lambda_{aj}^{S_{2a}(j)} \right]
\end{aligned} \tag{2.5}$$

with slightly different statistics:

$$S_{a1} = \sum_{j=2}^J m_{1j}^a, \quad S_{1a}(j) = \sum_{i=1}^{k^*} \sum_{l=j+1}^J m_{il}^a \quad \text{and} \quad S_{2a}(j) = \sum_{i=1}^{k^*} m_{ij}^a \tag{2.6}$$

and  $k^* = \min\{I, j\}$ . The two expressions (2.3) and (2.5) are useful to understand how the age-specific ringing data,  $m^1$  and  $m^a$ , contribute to the estimation of survival and recovery probabilities.

### 3 MODEL WITH UNKNOWN NUMBERS OF MARKED INDIVIDUALS

When the numbers  $R_i^d$  are not available, we only have individual records of recovered birds and we can build a matrix of data as in Table 1, but without the last column. We still need the rows of the corresponding matrix of multinomial proba-

bilities to sum to one, for each  $d$ , and we can, for example: denote by  $p_{ij}^d$  the conditional probability that a bird released at age  $d$  in year  $i$  is recovered in year  $j$ , given that it is recovered between  $i$  and  $J$ ; normalise the parameters  $p_{ij}^{*d} = p_{ij}^d / \sum_{k=i}^J p_{ik}^d$ ; and model the observed counts as  $(m_{i1}^d, \dots, m_{iJ}^d) \sim \text{Mult}(m_i^d; p_{i1}^*, \dots, p_{iJ}^*)$ . We are using a different notation, ( $p$  instead of  $\psi$ ), because we illustrate what happens to the survival estimates when in addition to this conditioning it is assumed that the recovery rates are constant across age and time, i.e that  $\lambda_j^d = \lambda, \forall d$  and  $\forall j$ . This assumption is used for example in the software MARK. In this case, the recovery parameter  $\lambda$  cancels out as it appears both in the numerator and in the denominator of the normalised probabilities which become:

$$p_{ij}^{*d} = \begin{cases} (1 - \phi_{di})/p_i^d & i = 1, \dots, I \quad j = i \quad d = 1, a \\ \phi_{di} \left[ \prod_{k=i+1}^{j-1} \phi_{ak} \right] (1 - \phi_{aj})/p_i^d & i = 1, \dots, L^* \quad j = i+1, \dots, J \quad d = 1, a \end{cases} \quad (3.1)$$

where  $p_i^d = (1 - \phi_{di}) + \sum_{k=i+1}^J \phi_{di} \left[ \prod_{k=i+1}^{j-1} \phi_{ak} \right] (1 - \phi_{aj})$ . Proceeding as in the previous section, we obtain a conditional likelihood for juveniles:

$$L(\mathbf{m}^1 | \boldsymbol{\phi}) \propto \frac{1}{K^1(\boldsymbol{\phi})} \left[ \prod_{i=1}^{L^*} \phi_{1i}^{S_{1i}^{(i)}} \prod_{i=1}^I (1 - \phi_{1i})^{m_{1i}^1} \right] \left[ \prod_{j=2}^{J-1} \phi_{aj}^{S_{11}^{(j)}} \prod_{j=2}^J (1 - \phi_{aj})^{S_{21}^{(j)}} \right] \quad (3.2)$$

where  $K^1(\boldsymbol{\phi}) = \prod_{i=1}^I (p_i^1)^{m_i^1}$  and the statistics are those described in the previous section, and a conditional likelihood for adult data:

$$L(\mathbf{m}^a | \boldsymbol{\phi}) \propto \frac{1}{K^a(\boldsymbol{\phi})} \left[ \prod_{i=1}^{L^*} \phi_{ai}^{S_a^{(i)}} \prod_{i=1}^I (1 - \phi_{1i})^{m_{ii}^a} \right] \left[ \prod_{j=2}^{J-1} \phi_{aj}^{S_{1a}^{(j)}} \prod_{j=2}^J (1 - \phi_{aj})^{S_{2a}^{(j)}} \right] \quad (3.3)$$

where  $K^a(\boldsymbol{\phi}) = \prod_{i=1}^I (p_i^a)^{m_i^a}$ . We still have an identifiability problem as in the probability matrix  $\mathbf{P}^{*1}$  we have  $2I - 1$  parameters of which  $\phi_{1I}$  is not estimable as it only appears in the bottom right cell and disappears from the likelihood since  $P_{II}^{*1} = \frac{1 - \phi_{1I}}{1 - \phi_{1I}}$ . In the next section we propose a way of comparing the role of the unconditional likelihood of the previous section with that of the conditional one obtained above in the estimation of the survival probabilities

## 4 COMPARING LIKELIHOODS FROM THE TWO MODELS

The comparison between the two likelihoods for juveniles and adults, (2.3) and (2.5), and the corresponding conditional ones is complicated by the terms  $K^1(\Phi)$ ,  $K^a(\Phi)$ ,  $\Delta^1(\Phi, \lambda)$  and  $\Delta^a(\Phi, \lambda)$ . Given the identifiability issues discussed in §2 and §3 only the first  $I - 1$  elements composing these quantities are relevant for a comparison of the survival estimates. Here we write these terms as functions of the demographic parameters to show how the estimation of survival rates is influenced by age and time specificity of recovery rates. The following steps are important to separate juvenile from adult survival parameters, within these four functions, and to compare analytically the two likelihoods to understand how the bias is induced. We consider, for simplicity, the case of square matrices so that  $J = I$ ,  $L^* = I - 1$  and  $m_{i.}^d = m_{iI}^d$ .

As shown in equation A.1 of Appendix A,  $\log K_{I-1}^1(\Phi)$  can be written as :

$$\log K_{I-1}^1(\Phi) = \sum_{i=1}^{I-1} m_i^1 \log(1 - \phi_{1i} g_{i+1}^1) \quad (4.1)$$

where  $g_{i+1}^1(\phi_{a1}^c)$ , defined for  $i = 1, \dots, I - 1$ , takes values in  $(0, 1)$  and can be interpreted as a discrete version of a Kaplan-Meier estimate for the probability that birds ringed in year  $i$  will survive at least until time  $I$ .

Similarly, for  $K_{I-1}^a(\Phi)$  (see equation A.2) we have:

$$\log K_{I-1}^a(\Phi) = \sum_{i=1}^{I-1} m_i^a \log(1 - g_i^a) \quad (4.2)$$

where  $g_i^a$  is now defined for  $i = 1, \dots, I$  and is equal to  $\phi_{a1} g_2^1$  when  $i = 1$ , while being identical to  $g_i^1$  for  $i = 2, \dots, I$ . When the number of birds ringed is available, from equation A.3, we have for juveniles:

$$\log \Delta_{I-1}^1(\Phi, \lambda) = \sum_{i=1}^{I-1} (R_i^1 - m_i^1) \log[1 - \lambda_{1i} + \phi_{1i}(\lambda_{1i} - h_{i+1}^1)] \quad (4.3)$$



where  $h_{i+1}^1(\phi_{a1}^c) = (1 - \phi_{ai+1})\lambda_{ai+1} + \phi_{ai+1}h_{i+2}^1$  takes values in  $(0,1)$ . Finally, when we consider birds ringed as adults we have from equation A.4:

$$\log \Delta_{I-1}^a(\Phi, \lambda) = \sum_{i=1}^{I-1} (R_i^a - m_i^a) \log(1 - h_i^a) \quad (4.4)$$

where  $h_i^a$  is now defined for  $i = 1, \dots, I$  and is equal to  $(1 - \phi_{a1})\lambda_{a1} + \phi_{a1}h_2^1$  when  $i = 1$ , while being identical to  $h_i^1$  for  $i = 2, \dots, I$ . The structure of these terms is important because recovery rates can only influence the estimation of survival rates through  $\Delta^1(\Phi, \lambda)$  and  $\Delta^a(\Phi, \lambda)$ . The assumptions we make on the structure of recovery rates play an important role when describing what mechanisms generated the data, but also when comparing conditional and unconditional models since they may imply differences in the estimation of survival rates when using (2.1) or (3.1).

#### 4.1 BIAS FROM IGNORING AGE SPECIFICITY OF RECOVERY RATES

We start by considering recovery rates that are constant over time:  $\lambda_{1i} = \lambda_1$  and  $\lambda_{ai} = \lambda_a$ . Freeman & Morgan (1992) reported that ignoring age specificity in recovery rates that are constant over time but different in the two age groups results in biased estimates for juvenile survival but not for adult survival. The authors fixed a distance between the two reporting rates and observed, with a simulation study, that if  $\lambda_1 < \lambda_a$  then  $\hat{\phi}_1 > \phi_1$  and viceversa, while the adult survival estimates are barely affected by the relationship between  $\lambda_1$  and  $\lambda_a$ . Here we show why this happens by directly comparing survival estimation in models (2.1) and (3.1), while in the next section we describe what happens when recovery rates vary also with time. We adopt a Bayesian approach and compare the estimated posterior distribution for parameters  $\phi_{1i}$  and  $\phi_{aj}$  with the two types of likelihood, so as to obtain simultaneously a point estimate and its associated variability. If we consider the juvenile survival then, given a prior distribution  $\pi(\phi_{1i})$ , by using the conditional model we can obtain a posterior  $\pi^{con}(\phi_{1i} | \mathbf{m}^1) \propto \pi(\phi_{1i})f_1(\phi_{1i}, \mathbf{m}^1)f_2(\phi_{a1}^c, \mathbf{m}^1)[K^1(\Phi)]^{-1}$ , where the form of  $f_1$  and  $f_2$  is given in (2.3); by using the unconditional model we obtain in-

stead  $\pi^{unc}(\phi_{1i} | \mathbf{m}^1, \mathbf{R}^1) \propto \pi(\phi_{1i})f_1(\phi_{1i}, \mathbf{m}^1)f_2(\phi_{a1}^c, \mathbf{m}^1)\Delta^1(\Phi, \lambda)$ . These posteriors are not of standard form but can be estimated, for example, with the Metropolis-Hastings algorithm (Chib & Greenberg 1995). If we denote by  $q(\phi_{1i}^C, \phi_{1i}^P)$  the proposal distribution, centred around  $\phi_{1i}^C$ , used to generate a new value  $\phi_{1i}^P$  and by  $A_i^{con}$  and  $A_i^{unc}$  the acceptance functions in the conditional and the unconditional models then, by taking logarithms, we obtain:

$$\log A_i^{con} = \log \left[ \frac{\pi(\phi_{1i}^P | \mathbf{m}^1)q(\phi_{1i}^P, \phi_{1i}^C)}{\pi(\phi_{1i}^C | \mathbf{m}^1)q(\phi_{1i}^C, \phi_{1i}^P)} \right] = \log \mathbf{K}_{\phi_{1i}=\phi_{1i}^C}^1 - \log \mathbf{K}_{\phi_{1i}=\phi_{1i}^P}^1 + a(\phi_{1i}^C, \phi_{1i}^P) \quad (4.5)$$

$$\log A_i^{unc} = \log \left[ \frac{\pi(\phi_{1i}^P | \mathbf{m}^1)q(\phi_{1i}^P, \phi_{1i}^C)}{\pi(\phi_{1i}^C | \mathbf{m}^1)q(\phi_{1i}^C, \phi_{1i}^P)} \right] = \log \Delta_{\phi_{1i}=\phi_{1i}^P}^1 - \log \Delta_{\phi_{1i}=\phi_{1i}^C}^1 + a(\phi_{1i}^C, \phi_{1i}^P). \quad (4.6)$$

where the function  $a(\cdot)$  depends on  $f_1, f_2$ , on the ratio of proposals and on the ratio of priors. The recovery rates appear only in the likelihood function so that the changing behaviour of the acceptance function does not depend on the particular choice of the prior or the proposal distribution and can be linked to what happens with alternative classical estimation methods. As shown in Appendix 6, the two expressions above can be simplified and compared to see how and a systematic difference in recovery rates produces a translation of the conditional acceptance function towards higher values of  $(\phi_{1i}^C, \phi_{1i}^P)$  when  $r_\lambda = \lambda_1/\lambda_a < 1$  and towards lower values when  $r_\lambda > 1$ . On the other hand when  $r_\lambda = 1$ , i.e.  $\lambda_1 = \lambda_a$ , the two functions are different but consider the same range of values for  $\phi_{1i}$  in the acceptance step. An example of the translation described above, which results in a systematic bias for survival rates, is plotted in Figure 1 where we compare the contour plots of  $\log A_i^{con}$  with those of  $\log A_i^{unc}$  for different values of  $r_\lambda$ . Since proposed values that lead to a negative log(acceptance function) are always accepted, the plots show how a similar range of values is likely to be accepted when  $\lambda = 1$  and how higher (lower) values are more likely to be accepted when  $r_\lambda < 1$  ( $r_\lambda > 1$ ). From Appendix 6 we also note that the form of  $\log \Delta^1(\Phi, \lambda)$  is such that the value of the ratio between  $\lambda_1$  and  $\lambda_a$  in-

fluences juvenile survival rates while adult survival rates only appear in the function  $g$ . On the other hand, the term  $\log \Delta^a(\Phi, \lambda)$  is not influenced by such a ratio because, although we estimate  $\phi_{a1}^c$  using both datasets, from the second year of age we only have adult recovery rates. Thus, the age-specific structure that we ignore with (3.1) is not shifting the conditional estimates for the adult survival rates towards higher or lower values, as noted by Freeman & Morgan (1992) in their simulation study. The bias induced on survival rates, that we described above, depends only on the structure of the recovery rates that is ignored. The acceptance functions, used in MCMC, are useful here because they allow a component wise comparison between the conditional and the unconditional estimates, while an explicit comparison by using the whole likelihood would not be feasible. From a practical perspective, when implementing rather than comparing these models, softwares such as WinBUGS or MARK may be used for a Bayesian or a Classical approach respectively, but with MARK it is not possible to estimate the variability of the conditional estimates for survival.

[Figure 1 about here.]

## 4.2 BIAS FROM IGNORING AGE AND TIME SPECIFICITY OF RECOVERY RATES

When recovery rates vary with both age and time we cannot use anymore the relationship  $h_{i+1}^1 = \lambda_a(1 - g_{i+1}^1)$  and it is harder to relate the behaviour of the acceptance function to estimation of survival. Since in the next section we have an example where recovery rates vary in this manner, we built a simulation study where we tried to decompose the two effects. In particular we considered the Godfinch species and, from the available information we had on  $\mathbf{R}^1$  and  $\mathbf{R}^a$ , we generated juvenile and adult survival rates from  $\phi_{1i} \sim \text{Unif}(0.3, 0.5)$ , for  $i = 1, \dots, I$ , and  $\phi_{aj} \sim \text{Unif}(0.4, 0.6)$ , for  $j = 1, \dots, J$ , to have an example where adult survival rates are slightly higher but the two series overlap in some years. We then considered four possible logistic linear trends over time for recovery rates:

(a) increasing but at the same level:  $\text{logit}(\lambda_{1i}) = \text{logit}(\lambda_{ai}) = -1.40 + 0.35i$

(b) decreasing but at the same level:  $\text{logit}(\lambda_{1i}) = \text{logit}(\lambda_{ai}) = -1.40 - 0.35i$

(c) increasing and parallel:  
 $\text{logit}(\lambda_{1i}) = -1.10 + 0.30i$   
 $\text{logit}(\lambda_{ai}) = -1.40 + 0.35i$

(d) decreasing and parallel:  
 $\text{logit}(\lambda_{1i}) = -1.10 - 0.30i$   
 $\text{logit}(\lambda_{ai}) = -1.40 - 0.35i$

These coefficients for the logistic regression produce increasing/decreasing trends between 10% and 30% or 15% and 35% as detailed in figure 2. We did not to explore the whole range  $[0, 1]$  for  $\boldsymbol{\lambda}$  because figures above 20-30% would be unrealistic in this context. Given these values we proceeded as follows:

1. Select a structure for recovery rates,  $(\boldsymbol{\lambda}^1, \boldsymbol{\lambda}^a)$ , from (a)-(d)
2. Calculate the matrices  $\mathbf{P}^1$  and  $\mathbf{P}^a$
3. generate the data  $(\mathbf{m}_1^1, \dots, \mathbf{m}_{N_{sim}}^1) \sim \text{Mult}(\mathbf{R}^1; \mathbf{P}^1)$  and  $(\mathbf{m}_1^a, \dots, \mathbf{m}_{N_{sim}}^a) \sim \text{Mult}(\mathbf{R}^a; \mathbf{P}^a)$ ,  
for  $N_{sim} = 100$
4. Fit model (3.1) to each pair of datasets  $(\mathbf{m}_n^1, \mathbf{m}_n^a)$ , for  $n = 1, \dots, N_{sim}$ , thus ignoring the information on the number of birds ringed.

Survival parameters were estimated with the Metropolis-Hasting algorithm, by assigning a uniform prior,  $\pi(\phi_{di}) = \text{Unif}(0, 1)$ , and proposing values from a truncated Normal distribution,  $q(\phi_{di}^C, \phi_{di}^P) = \text{Norm}(\phi_{di}^C, \sigma_{di}^2)I_{[0,1]}(\phi_{di})$ . For each  $n$ , we used an adaptive phase of 5000 iterations, preceding the actual generation of the chain, where the proposal variances  $\sigma_{di}^2$ , were tuned every  $m$  iterations, while the corresponding acceptance rates were monitored, as in Gelman et al. (1996), until they were between 25% and 45% for all parameters. We simulated a chain for 50000 iterations, of which we discarded the first half as a burnin, and then we calculated

an average of the posterior mean estimates for the survival rates and of the corresponding standard errors over the  $N_{sim}$  generated datasets. The first two rows of Figure 2 show the behaviour of the conditional estimates when recovery rates are increasing or decreasing but equal in the two age groups. The conditional model (3.1) ignores the presence of a time structure with a consequent small but systematic bias in both juvenile and adult survival estimates: a positive (negative) bias for both when recovery rates increase (decrease). The third and fourth row show the behaviour of the conditional estimates when there is both a linear change over time and a systematic difference between the two age groups: the two effects are combined and, given the values we have chosen, the distance between  $\lambda_{1i}$  and  $\lambda_{ai}$  appears to be more important than their trend over time in determining the direction of the bias induced on the survival estimates by model (3.1).

[Figure 2 about here.]

## 5 APPLICATION

We applied the results from the previous sections to the Starling data, relative to the period 1965-1999, that have been recently analysed by Freeman et al. (2007) and for which we do have  $R_i^1$  and  $R_i^a$ . We first estimated recovery rates (as well as survival) from model (9) and for the Starling these vary with both age and time, as in §4.2. Recovery rates were regressed against time, using a logistic link, and the estimated coefficients for juvenile and adults are reported in Table 2 together with their 95% Highest Posterior Density Intervals (HPDI). We then ignored  $R_i^1$  and  $R_i^a$  and applied model (1) to see how much information we lose by ignoring the structure of the recovery rates. From Figure 3 we can see that the conditional estimates for survival rates are systematically biased and the conclusions for juvenile survival are particularly affected. The case of the Starling is interesting because recovery

rates are very small, with estimates for adults declining from 0.05 to 0.02 and estimates for juveniles declining from 0.03 to 0.02. With such a small difference we may naively think that the specification of an age-time structure is not important and that a conditional model would not produce a relevant change in the survival estimates. This example shows how these speculations can be wrong and how a small difference in recovery rates may induce a large distortion in the estimation of survival. Such small recovery rates also imply that we are dealing with sparse data matrices and if we use a full age-time specification, instead of a logistic regression, then recovery rates are heavily influenced by the prior distribution. In fact, the posterior estimates would have little precision and, for example, with a vague prior the posterior means would be too large overestimating recovery rates in some years. Recovery is negatively correlated with survival, thus the sparseness of data would also produce an underestimation of the latter, in the corresponding years, which would appear to be more variable over time than it actually is. The same problem occurs with a Classical approach, as estimates for recovery rates tend to have an erratic behaviour if modelled with a full age-time specification, as also reported in Freeman et al. (2007) who analysed the same Starling dataset by using the software MARK. Here the logistic regression improves enormously the estimation of parameters and provides a practical example of the model used in the simulation study of §4.2.

[Table 2 about here.]

[Figure 3 about here.]

## 6 CONCLUSIONS

Conditioning on recovery may be appealing since minor analytic modifications of the standard models allow us to deal with incomplete data arising, for example,

from the missing numbers of individuals ringed, (on which we focus), or from missing covariate information, as in Catchpole et al. (2004). However we believe it is important to highlight the implications that this type of conditioning has on the estimation of survival. If the distance between the age-specific survival rates is less than or equal to the bias induced by the conditional analysis then we may fail to identify an age structure. On the other hand the bias itself may create a spurious age structure in the survival rates. These analyses are often used to identify the demographic structure that explains how the population changes over time and a recent approach which has become very popular combines ring recovery data with census data by using state space models, as in Besbeas et al. (2002), Brooks et al. (2004), Thomas et al. (2005) and Mazzetta et al. (2007). This way of combining different sources of information has several advantages when the number of marked animals is known and is interesting both from a biological perspective, by allowing for a more comprehensive description of the system, and from a statistical perspective, by accounting for different sources of parameter uncertainty. However, some extra caution is needed if requested to use age-specific conditional ring recovery models in state space models, for example, which is the request that prompted this work. In fact, for species with a higher/lower juvenile probability of recovery there is a risk of using the wrong number of age classes to model important demographic changes. In general, conditional models are used when it is not possible to gather the information that is missing,  $R_i^d$  in this case, and this type of comparison may not be very helpful from a practical perspective. Here instead this evaluation may be useful to decide when it is appropriate to recover the information about  $R_i^d$ , often a quite expensive but possible process, and when survival rates obtained with the conditional analysis provide instead reasonable estimates. With this work we hope to facilitate, in a practical setting, the choice between implementing these models and recovering the information needed to perform an unconditional analysis.

## ACKNOWLEDGEMENTS

I would like to thank the BTO for providing the data and the many volunteers who collected such data. I also gratefully acknowledge the financial support of the UK Biotechnology and Biological Sciences Research Council and the help I received from Steve Freeman and Stephen Baillie, at the BTO, to understand the data collection and the ringing schemes.

## REFERENCES

- BAILLIE, S. R., MARCHANT, J. H., CRICK, H. Q. P., NOBLE, D. G., BALMER, D. E., COOMBES, R. H., DOWNIE, I. S., FREEMAN, S. N., JOYS, A. C., LEECH, D. I., RAVEN, M. J., ROBINSON, R. A. & THEWLIS, R. M. (2006). Breeding Birds in the Wider Countryside: their conservation status 2005. *BTO Research Report No. 435. BTO, Thetford. (<http://www.bto.org/birdtrends>)*.
- BESBEAS, P., FREEMAN, S. N., MORGAN, B. J. T. & CATCHPOLE, E. A. (2002). Integrating mark-recapture-recovery and census data to estimate animal abundance and demographic parameters. *Biometrics*, **58**, 540–547.
- BROOKS, S. P., KING, R. & MORGAN, B. J. T. (2004). A Bayesian Approach to Combining Animal Abundance and Demographic Data. *Animal Biodiversity and Conservation* **27**, 515–529.
- BURNHAM, K. P. (1990). Survival analysis of recovery data from birds ringed as young: efficiency of analyses when numbers ringed are not known. *The Ring*, **13**, 115–132.
- CATCHPOLE, E. A. & MORGAN, B. J. T. (1997). Detecting parameter redundancy. *Biometrika*, **84**, 187–196.
- CATCHPOLE, E. A., MORGAN, B. J. T. & COULSON, T. (2004). Conditional methodology for individual case history data. *Applied Statistics*, **53**, 123–131.



- CATCHPOLE, E. A., MORGAN, B. J. T. & FREEMAN, S. N.(1998). Estimation in Parameter-Redundant Models. *Biometrika*, **85**, 462–468.
- CHIB, S. & GREENBERG, E.(1995) Understanding the Metropolis-Hastings Algorithm. *The American Statistician*, **49**, 327–335.
- FREEMAN, S. N. & MORGAN, B. J. T.(1992). A modelling strategy for Recovery Data from Birds Ringed as Nestlings. *Biometrics*, **48**, 217–235 .
- FREEMAN, S. N.,ROBINSON, R. A., CLARK, J. A., GRIFFIN, B. M. & ADAMS, S. Y.(2007). Changing demography and population decline in the common Starling *Sturnus Vulgaris*: a multisite approach to integrated population monitoring. *Ibis*, **149**, 587–596.
- GELMAN, A., ROBERTS, G. O. & GILKS, W. R.(1996). Efficient Metropolis Jumping Rules. *BAYESIAN STATISTICS 5*Oxford: OUP 599–607.
- MAZZETTA, C., BROOKS, S. P. & FREEMAN, S. N.(2007). On Smoothing Trends in Population Index Modelling. *Biometrics*. **67**, 1007–1014.
- THOMAS, L., BUCKLAND, S. T., NEWMAN, K. B. & HARWOOD, J.(2005). A unified framework for modelling wildlife population dynamics. *Australian and New Zealand Journal of Statistics*. **47**, 19–34.
- WILLIAMS, B. K., NICHOLS, J. D. & CONROY, M. J.(2002). *Analysis and Management of Animal Populations*. Academic Press.

## Appendix A

Writing the terms  $K^1(\Phi)$ ,  $K^a(\Phi)$ ,  $\Delta^1(\Phi, \lambda)$  and  $\Delta^a(\Phi, \lambda)$  as functions of  $g$  and  $h$  to facilitate the comparison of the likelihood in the two models

$$\begin{aligned}
\log K^1(\Phi) &= \log K_{I-1}^1(\Phi) + m_{II}^1 \log(p_{II}^1) \\
\log K_{I-1}^1(\Phi) &= \sum_{i=1}^{I-1} m_i^1 \log(p_{ii}^1 + \sum_{j=i+1}^I p_{ij}^1) \\
&= \sum_{i=1}^{I-1} m_i^1 \log \left[ (1 - \phi_{1i}) + \phi_{1i} \sum_{j=i+1}^I \prod_{k=i+1}^{j-1} \phi_{ak} (1 - \phi_{aj}) \right] \\
&= \sum_{i=1}^{I-1} m_i^1 \log \left[ 1 - \phi_{1i} \left( 1 - \sum_{j=i+1}^I \prod_{k=i+1}^{j-1} \phi_{ak} [1 - \phi_{aj}] \right) \right] \\
&= \sum_{i=1}^{I-1} m_i^1 \log(1 - \phi_{1i} g_{i+1}^1)
\end{aligned} \tag{A.1}$$

$$\begin{aligned}
\log K^a(\Phi) &= \log K_{I-1}^a(\Phi) + m_{II}^a \log(p_{II}^a) \\
\log K_{I-1}^a(\Phi) &= \sum_{i=1}^{I-1} m_i^a \log(p_{ii}^a) \\
&= \sum_{i=1}^{I-1} m_i^a \log \left[ (1 - \phi_{ai}) + \phi_{ai} \sum_{j=i+1}^I \prod_{k=i+1}^{j-1} \phi_{ak} (1 - \phi_{aj}) \right] \\
&= \sum_{i=1}^{I-1} m_i^a \log(1 - g_i^a)
\end{aligned} \tag{A.2}$$

$$\begin{aligned}
\log \Delta^1(\Phi, \lambda) &= \log \Delta_{I-1}^1(\Phi, \lambda) + (R_I^1 - m_{II}^1) \log(1 - \psi_{II}^1) \\
\log \Delta_{I-1}^1(\Phi, \lambda) &= \sum_{i=1}^{I-1} (R_i^1 - m_i^1) \log(1 - \psi_{ii}^1 - \sum_{j=i+1}^I \psi_{ij}^1) \\
&= \sum_{i=1}^{I-1} (R_i^1 - m_i^1) \log \left[ 1 - (1 - \phi_{1i}) \lambda_{1i} - \phi_{1i} \sum_{j=i+1}^I \prod_{k=i+1}^{j-1} \phi_{ak} (1 - \phi_{aj}) \lambda_{aj} \right] \\
&= \sum_{i=1}^{I-1} (R_i^1 - m_i^1) \log \left[ 1 - \lambda_{1i} + \phi_{1i} (\lambda_{1i} - h_{i+1}^1) \right]
\end{aligned} \tag{A.3}$$

$$\begin{aligned}
\log \Delta^a(\Phi, \lambda) &= \log \Delta_{I-1}^a(\Phi, \lambda) + (R_I^a - m_{II}^a) \log(1 - \psi_{II}^a) \\
\log \Delta_{I-1}^a(\Phi, \lambda) &= \sum_{i=1}^{I-1} (R_i^a - m_i^a) \log \left( 1 - \psi_{ii}^a - \sum_{j=i+1}^I \psi_{ij}^a \right) \\
&= \sum_{i=1}^{I-1} (R_i^a - m_i^a) \log \left[ 1 - (1 - \phi_{ai}) \lambda_{ai} - \sum_{j=i+1}^I \prod_{k=i}^{j-1} \phi_{ak} (1 - \phi_{aj}) \lambda_{aj} \right] \\
&= \sum_{i=1}^{I-1} (R_i^a - m_i^a) \log \left[ 1 - \sum_{j=i}^I \prod_{k=i}^{j-1} \phi_{ak} (1 - \phi_{aj}) \lambda_{aj} \right] \\
&= \sum_{i=1}^{I-1} (R_i^a - m_i^a) \log(1 - h_i^a)
\end{aligned} \tag{A.4}$$

## Appendix B

Age structure for recovery rates and translation of the acceptance function

We can use appendix 6 to simplify (4.5) and (4.6) in §4.1 and write:

$$\log A_i^{con} = -m_i^1 \left[ \log(1 - \phi_{1i}^p g_{i+1}^1) - \log(1 - \phi_{1i}^c g_{i+1}^1) \right] + a(\phi_{1i}^c, \phi_{1i}^p).$$

For  $\lambda_{1i} = \lambda_1$  and  $\lambda_{ai} = \lambda_a$ , in the unconditional model we have:

$$\log A_i^{unc} = (R_i^1 - m_i^1) \left[ \log(1 - \lambda_1 - \phi_{1i}^p [\lambda_1 - h_{i+1}^1]) - \log(1 - \lambda_1 - \phi_{1i}^c [\lambda_1 - h_{i+1}^1]) \right] + a(\phi_{1i}^c, \phi_{1i}^p)$$

If we define the ratio  $r_\lambda = \lambda_1 / \lambda_a$  then we can write:

$$\begin{aligned} \log A_i^{unc} = & (R_i^1 - m_i^1) \left[ \log \left( \lambda_a \left[ \frac{1}{\lambda_a} - r_\lambda - r_\lambda \phi_{1i}^p + \phi_{1i}^p \frac{h_{i+1}^1}{\lambda_a} \right] \right) - \right. \\ & \left. \log \left( \lambda_a \left[ \frac{1}{\lambda_a} - r_\lambda - r_\lambda \phi_{1i}^c + \phi_{1i}^c \frac{h_{i+1}^1}{\lambda_a} \right] \right) \right] + a(\phi_{1i}^c, \phi_{1i}^p) \end{aligned}$$

and by replacing  $h_{i+1}^1 / \lambda_a$  with  $1 - g_{i+1}^1$  we obtain:

$$\begin{aligned} \log A_i^{unc} = & (R_i^1 - m_i^1) \left[ \log \left( r_\lambda \left( \frac{1}{\lambda_1} - 1 \right) - (r_\lambda - 1) \phi_{1i}^p - \phi_{1i}^p g_{i+1}^1 \right) - \right. \\ & \left. \log \left( r_\lambda \left( \frac{1}{\lambda_1} - 1 \right) - (r_\lambda - 1) \phi_{1i}^c - \phi_{1i}^c g_{i+1}^1 \right) \right] + a(\phi_{1i}^c, \phi_{1i}^p). \end{aligned}$$

Thus, the acceptance function is translated towards higher values of  $(\phi_{1i}^c, \phi_{1i}^p)$  when  $r_\lambda < 1$  and towards lower values when  $r_\lambda > 1$ .

Figure 1. Contour plots of the acceptance function for one component  $\phi_{1i}$  of the vector of survival probabilities for the juveniles. Each contour plot shows ten different levels of the acceptance function, from its minimum to its maximum value (labelled on the graph). The sequence of plots shows how the function concentrates on different ranges of values for  $\phi_{1i}^p$  (proposed) and  $\phi_{1i}^c$  (current) depending on the ratio  $r_\lambda$  between juvenile and adult recovery. From top left, clockwise: (i) conditional model; (ii) unconditional model  $r_\lambda = 1$ ; (iii) unconditional model  $r_\lambda = 1.25$  and (iv) unconditional model  $r_\lambda = 0.8$ . In this plot the remaining parameters were set to:  $g_{i+1}^! = 0.005$ ;  $S_1(i) = 30$ ;  $m_{ii}^! = 15$  and  $R_i^! = 300$ .

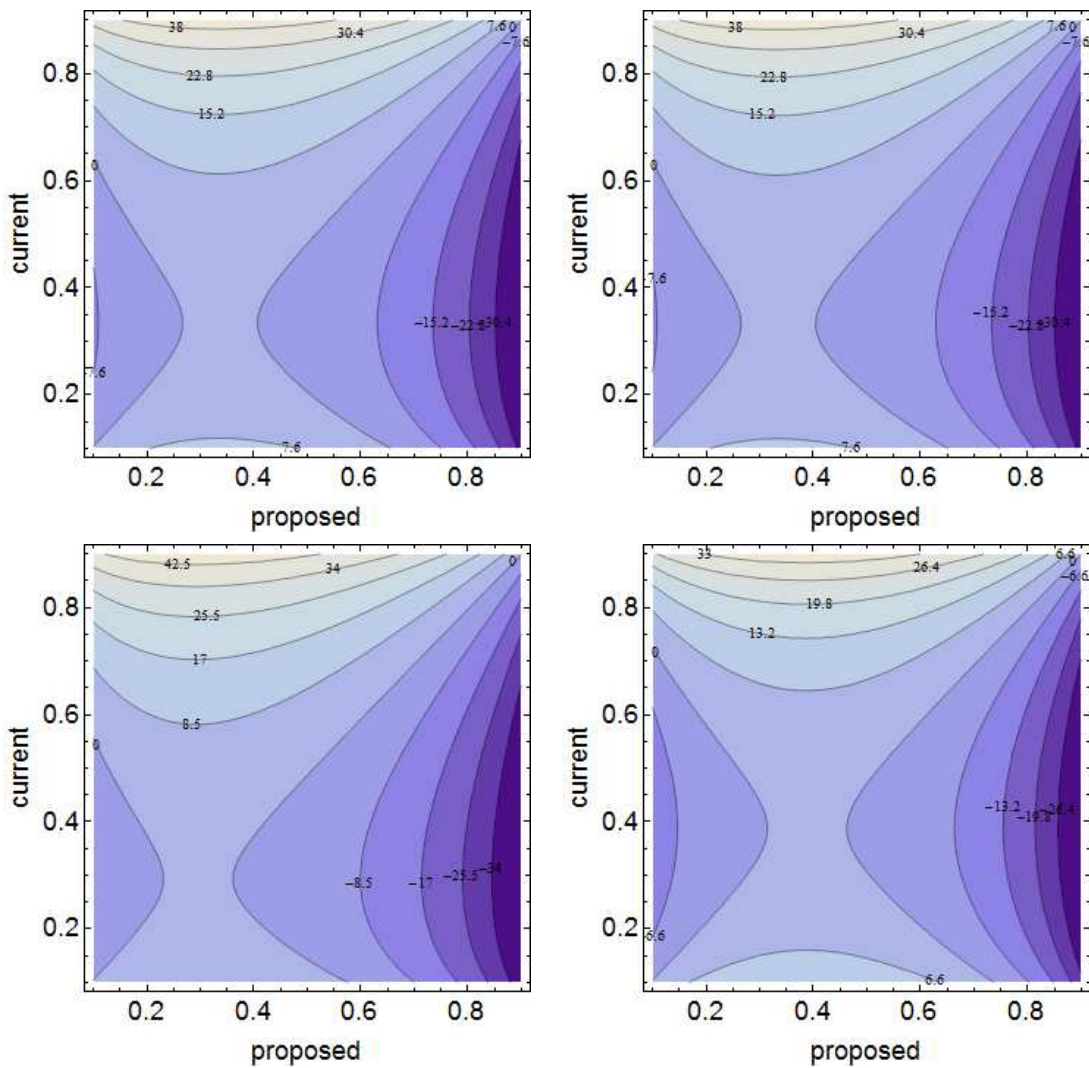


Figure 2. Conditional model fitted to simulated data: averaged survival estimates  $\pm$  two averaged standard errors over 100 simulated datasets. In case (a)  $\lambda_{1i} = \lambda_{ai}$  increase from 10% to 30%; (b)  $\lambda_{1i} = \lambda_{ai}$  decrease from 30% to 10%; (c)  $\lambda_{1i}$  increase from 15% to 35% and  $\lambda_{ai}$  increase from 10% to 30%; (d)  $\lambda_{1i}$  decrease from 35% to 15% and  $\lambda_{ai}$  decrease from 30% to 10%. True values for survival rates are in grey and plots (c) and (d) show the systematic underestimation of juvenile survival when juvenile recovery is higher than adult recovery.

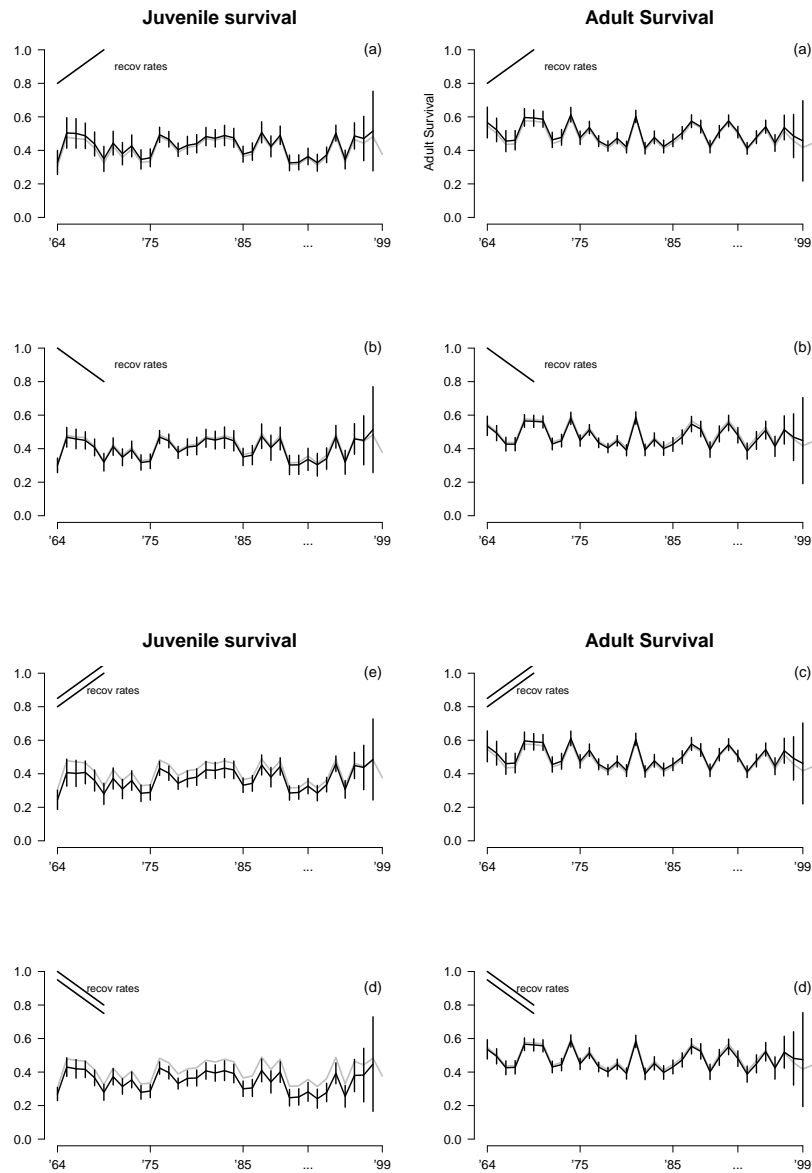


Figure 3. Starling data: bias in survival estimates when ignoring recovery rates that vary with age and time. Estimates and 95% HPDIs from the unconditional model are in grey while estimates and 95% HPDIs from the conditional model are in black. Here juvenile recovery is lower than adult recovery and when we ignore this information juvenile survival is overestimated.

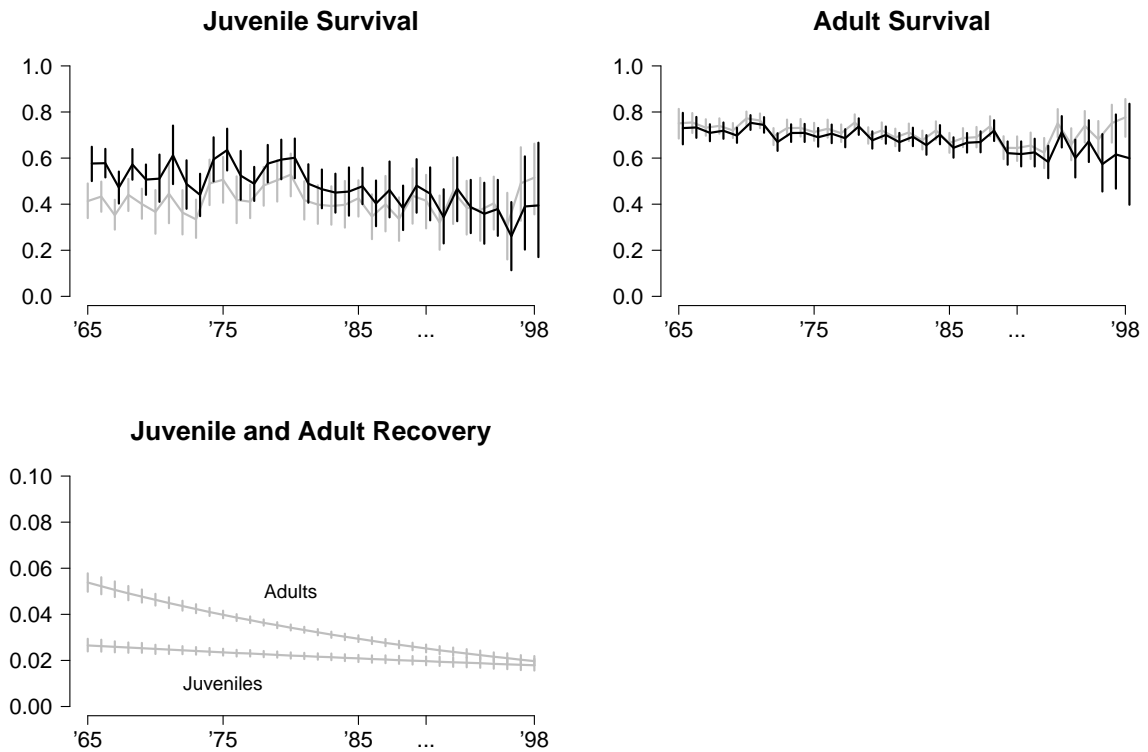


Table 1. Data structure for birds ringed at age  $d$  in year  $i$  and recovered in year  $j$ 

	1	2	3	...	I	...	J	J+1
1	$m_{11}^d$	$m_{12}^d$	$m_{13}^d$	...	$m_{1I}^d$	...	$m_{1J}^d$	$R_1^d - m_{1.}^d$
2		$m_{22}^d$	$m_{23}^d$	...	$m_{2I}^d$	...	$m_{2J}^d$	$R_2^d - m_{2.}^d$
3			$m_{33}^d$	...	$m_{3I}^d$	...	$m_{3J}^d$	$R_3^d - m_{3.}^d$
...			...		...		...	...
I					$m_{II}^d$	...	$m_{IJ}^d$	$R_I^d - m_{I.}^d$

Table 2. Starling data: posterior means (Standard Deviation) and 95% Highest Posterior Density Intervals for logistic regression coefficients for juvenile and adult recovery rates

	Mean (SD)	95% HPDI
Intercept Juveniles	-3.813 (0.035)	[-3.877, -3.742]
Slope Juveniles	-0.126 (0.035)	[-0.192, -0.056]
Intercept Adults	-3.404 (0.022)	[-3.448, -3.362]
Slope Adults	-0.323 (0.027)	[-0.374, -0.269]