



HAL
open science

Capture-recapture models with heterogeneity to study survival senescence in the wild

Guillaume G. Péron, Pierre-André Crochet, Rémi Choquet, Roger Pradel, Jean-Dominique Lebreton, Olivier Gimenez

► **To cite this version:**

Guillaume G. Péron, Pierre-André Crochet, Rémi Choquet, Roger Pradel, Jean-Dominique Lebreton, et al.. Capture-recapture models with heterogeneity to study survival senescence in the wild. *Oikos*, 2010. hal-03502442

HAL Id: hal-03502442

<https://hal.science/hal-03502442>

Submitted on 25 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Capture-recapture models with heterogeneity to study survival senescence in the wild

Authors:

Guillaume Péron ⁽¹⁾

Pierre-André Crochet ⁽¹⁾

Rémi Choquet ⁽¹⁾

Roger Pradel ⁽¹⁾

Jean-Dominique Lebreton ⁽¹⁾

Olivier Gimenez ⁽¹⁾

Affiliation :

⁽¹⁾ Centre d'Ecologie Evolutive et Fonctionnelle UMR 5175

CNRS, 1919 Route de Mende 34293 Montpellier, Cedex 5 – France

Author for correspondence:

Guillaume Péron

Phone : +33(0)4 67 61 32 98

e-mail: guillaume.peron@cefe.cnrs.fr

Abstract

Detecting senescence in wild populations and estimating its strength raises three challenges. First, in the presence of individual heterogeneity in survival probability, the proportion of high-survival individuals increases with age. This increase can mask a senescence-related decrease in survival probability when the probability is estimated at the population level. To accommodate individual heterogeneity we use a mixture model structure (discrete classes of individuals). Second, the study individuals can elude the observers in the field, and their detection rate can be heterogeneous. To account for detectability issues we use capture-mark-recapture (CMR) methodology, mixture models and data that provide information on individuals' detectability. Last, emigration to non-monitored sites can bias survival estimates, because it can occur at the end of the individuals' histories and mimic earlier death. To model emigration we use Markovian transitions to and from an unobservable state. These different model structures are merged together using hidden Markov chain CMR models, or multievent models. Simulation studies illustrate that reliable evidence for survival senescence can be obtained using highly heterogeneous data from non site-faithful individuals. We then design a tailored application for a dataset from a colony of black-headed gull *Chroicocephalus ridibundus*. Survival probabilities do not appear individually variable, but evidence for survival senescence becomes significant only when accounting for other sources of heterogeneity. This result suggests that not accounting for heterogeneity leads to flawed inference and/or that emigration heterogeneity mimics survival heterogeneity and bias senescence estimates.

Additional Key-Words

Ageing, Capture-recapture, Dispersal, E-SURGE, *Larus*.

1 **Introduction**

2 Senescence, the decline in fitness components with age due to internal physiological
3 deterioration (Medawar 1952), has been evidenced in several life-history traits in a variety of
4 wild vertebrates (e.g. Loison et al. 1999; Crespín et al. 2006; Nussey et al. 2006) and is
5 thought to be the rule for most (or all) vertebrate species (see Finch 1990; Jones et al. 2008).
6 Yet, several demographic studies of wild populations found no decline in individuals'
7 performance with age (Nichols et al. 1997; Miller 2001; Pistorius and Bester 2002; Congdon
8 et al. 2003), fuelling a debate over the validity of their results.

9 Among the potential flaws of demographic studies of senescence, the non-modelled
10 effect of individual heterogeneity is recurrent (Vaupel and Yashin 1985; Cam et al. 2002;
11 Zens and Peart 2003; van de Pol and Verhulst 2006). Individual heterogeneity can be defined
12 as the occurrence of systematic variation among individuals in demographic parameters. It
13 can originate from genetic differences or differences in the conditions experienced during
14 development (Fox et al. 2006), variation in individual strategies or quality (e.g. covariation
15 between reproductive effort and survival: Hamel et al. 2008), sex-bias in dispersal or
16 behaviour or the interplay of behavioural differences and study design (e.g. social status- or
17 body condition-dependent detectability: Whitehead and Wimmer 2005; Regehr et al. 2007;
18 Crespín et al. 2008). Individual heterogeneity can lead to population-level patterns that are not
19 always representative of the actual relationship at the individual level (Vaupel and Yashin
20 1985). In the case of senescence studies, since the proportion of individuals with high survival
21 probability will tend to increase with age, the age-specific population average of survival
22 probability, which is used in most studies of survival senescence, might not decrease or might
23 even increase with age (Vaupel and Yashin 1985; Fig. 1).

24 Heterogeneity can be modelled with known individual covariates: Regehr et al. (2007)
25 modelled a gender effect combined with an effect of the observation method; Fox et al. (2006)

26 modelled an effect of the year of birth and family structure. Yet, the precise cause of
27 heterogeneity is often not identified, or not measured. Two ways to accommodate individual
28 variation of unknown origin have been proposed: continuous random effects (e.g. Cam et al.
29 2002; Royle 2008) and mixture models that consider discrete classes of heterogeneity
30 (Pledger et al. 2003; Pradel 2009).

31 In addition, it is well known that analysing data on marked individuals with models that
32 do not formally estimate detection probabilities together with other demographic parameters
33 (Lebreton et al. 1992) can bias estimations of demographic parameters, including the rate of
34 senescence (Gimenez et al. 2008). Individual heterogeneity in detection rate also biases
35 survival estimates downwards if not accounted for (Pradel et al. 1997; Prévot-Julliard et al.
36 1998a) and it violates the fundamental assumption of parameter homogeneity in CMR models
37 (Lebreton et al. 1992), which can lead to flawed inference (Burnham and Anderson 2002).

38 Last, temporary emigration to non-monitored sites, such that individuals missing during
39 several capture occasions might still be alive elsewhere (Burnham 1993; Fujiwara and
40 Caswell 2002), evidently impacts on the estimation of “true” (as opposed to “local”) survival
41 probability. If emigration probability is subject to individual heterogeneity, patterns similar to
42 heterogeneity in survival might appear in the data. Consequently, heterogeneity in temporary
43 emigration can affect the detection of survival senescence as well. In short, we identified three
44 features of population studies (individual heterogeneity, imperfect detectability, temporary
45 emigration out of the study area) which, when not included in the population models, can bias
46 estimates of senescence and/or cause a lack of fit that leads to flawed inference.

47 Multievent models have been introduced as a unified framework by Pradel (2005). They
48 extend multistate models, in which individuals move between states or die and can at each
49 occasion be detected or not, by considering that the state of an individual is imperfectly
50 determined when it is observed. Their structure rests on the more general framework of

51 hidden Markov chain models (McDonald and Zucchini 1997). The introduction of this model
52 structure in the field of CMR data analysis was initially motivated by capture heterogeneity as
53 reviewed by Pledger (2000). The use of hidden Markov chains in cases not related to
54 individual heterogeneity was to our knowledge initiated by Nichols et al. (2004). These
55 authors modelled a situation where males and females had different survival probabilities, but
56 could not always be separated in the field because of reduced sexual dimorphism. Since then,
57 following the development of the software E-SURGE (Choquet et al. 2009a) numerous
58 applications have been proposed (see Pradel 2009). Among others, the implementation of
59 memory models (where demographic parameters depend on the states occupied during the
60 two preceding time steps) is made more straightforward through the use of hidden states
61 (Rouan et al. 2009), and epidemiology models can be fitted to data where health status
62 determination is uncertain or incomplete (Conn and Cooch 2009).

63 In this paper we show how this recently developed modelling framework can be used to
64 overcome the types of heterogeneity that typically plague senescence studies. Although these
65 developments are quite general and can be applied to other study situations and taxa, we
66 specifically tailor our example to a study of black-headed gulls *Chroicocephalus ridibundus*.
67 This is a species for which survival senescence is likely to occur (Pugesek et al. 1995; Cam et
68 al. 2002 in related species), but a population in which strong heterogeneities are expected (see
69 “Study site and population” in the method section). After presenting the dataset and the
70 features that suggested the need for this new development in CMR models, we present this
71 development and provide simulations that illustrate their performance.

72 **Methods**

73 **Study site and population**

74 Black-headed Gull (*Chroicocephalus ridibundus*) are long-lived Charadriiform birds
75 (maximum longevity recorded in our study area: 30 years) and breed colonially, often on
76 vegetated ponds. The data come from a long-term monitoring program of Black-headed Gulls
77 breeding in La Ronze (noted LR) pond, a large (more than 4000 pairs in recent years) colony
78 located in the Forez basin, at Craintilleux, central France (45°35'N 4°14'E). In this population,
79 detection is known to vary between individuals because nests are built within vegetation or at
80 its edge and because a large proportion of the re-sightings are made on the nests (Prévot-
81 Julliard et al. 1998a). Note that some re-sightings are made on other perches so that all birds
82 are potentially detectable even if their nests are not visible. Additionally, preliminary results
83 in the same population indicated that dispersal rates were individually variable, in particular
84 because of differences between males and females (Grosbois 2001), as is commonplace in
85 birds (Greenwood 1980). Yet, for both detection and emigration, we only had very partial
86 information on the characteristics of the birds in the field. Detectability cannot be assessed for
87 nests which are not visible of course and can't be evaluated for birds seen on other perches.
88 Furthermore, Black-headed Gulls are only weakly sexually dimorphic, which precludes
89 sexing of most birds in the field. It was thus clear that accounting for unknown or unmeasured
90 sources of variation between individuals would be very useful if we were to assess survival
91 senescence in this population.

92 Previous work in the same population indicated that time effects on survival
93 probabilities were reduced or absent (Prévot-Julliard et al. 1998a; Grosbois 2001). We were
94 thus confident that, despite most observations of old individuals occurred at the end of the
95 time series, unaccounted time effects could not confound age effects.

96 Data collection

97 Chicks were ringed before fledging with stainless steel rings. The use of stainless rings is
98 particularly appropriate for the study of senescence since they almost do not wear with age (in
99 all occasions when a known-age adult was physically recaptured at age >14, the code was
100 perfectly readable and the ring could not be removed even with the use of pliers).

101 Observations of ringed adults were conducted using a floating blind from which metal
102 ring codes could be read with a telescope (Lebreton, 1987). At each observation of a ringed
103 adult, we recorded whether it attended a nest or was feeding chicks, which would confirm its
104 status as breeder. Now, an individual nesting in an accessible location had a high probability
105 to be confirmed as breeder, whereas an individual breeding in a less accessible part of the
106 colony was more often observed on roosts or other perches and was more likely to end up
107 with an unconfirmed breeding status. The confirmation of breeding status thus potentially
108 yielded information on the detection probability.

109 We analyzed the survival of 1556 stainless-ringed adults, for a period of 28 years (from
110 1978 to 2006), starting from their first re-observation in LR, which occurred from age two to
111 age 23 (mean 5.5; SD 4.1). The large range of age at first resighting is partly due to the fact
112 that black-headed gulls are not present on the colony before starting to reproduce between two
113 and six year-old (Clobert et al. 1994), partly due to the non-exhaustiveness of detection and
114 partly due to the fact that some birds start breeding on other colonies and disperse toward LR
115 after several breeding attempts (Péron et al. submitted). The data were coded with one digit
116 per year: '0' (not observed), '1' (confirmed breeder) or '2' (non-confirmed breeder)

117 **Model description**

118 *1) Multievent models for the study of black-headed gull senescence*

119 As introduced above, our approach was based upon multievent CMR models (Pradel 2005).

120 The observer records “events” (here “not seen”, “confirmed breeder”, “non-confirmed
121 breeder”) that carry uncertain information on the state that the individual occupies at the

122 current sampling occasion (see below and Appendix S1 for state description). The

123 relationship between states and events is thus probabilistic (Pradel 2005).

124 All models were fully described by first considering the vector of probabilities of
125 initial presence in the various states (Π -vector), then linking states at successive sampling
126 occasions by the matrix of survival/transition probabilities (Φ -matrix), exactly like in
127 multistate models, while the events were linked to states by the matrix of event probabilities
128 (B -matrix). For convenience we separated Φ in two steps (S -matrix for survival probabilities
129 and Ψ -matrix for dispersal probabilities), and B (P -matrix for detection probabilities and R -
130 matrix for probabilities to confirm breeding status when detected). Full details on the model
131 structure and examples of these matrices are presented in Appendix S1.

132 *2) Individual heterogeneity*

133 Discrete classes of individuals were built to accommodate heterogeneity, each class being

134 associated with a distinct value of the parameter(s) (Pradel 2009); these classes were the

135 actual states of the multievent model. For a simple example, in a model with a two-class

136 heterogeneity structure for survival probability and no possibility to emigrate [model $\{\varphi(h_2),$

137 $p(\cdot)\}$ of Pledger et al. (2003)], there will be a state “low survival” and a state “high survival”.

138 In this model, the probability of the five first events in the first example history is:

$$\begin{aligned}
 &P(11010) = \\
 139 &\pi_1^H s_1^H p_2 s_2^H (1 - p_3) s_3^H p_4 [s_4^H (1 - p_5) + (1 - s_4^H)] \\
 &+ (1 - \pi_1^H) s_1^L p_2 s_2^L (1 - p_3) s_3^L p_4 [s_4^L (1 - p_5) + (1 - s_4^L)]
 \end{aligned}$$

140 Superscripts H and L refer to high and low survival classes respectively; subscripts refer
141 to time-dependence. π , s , p stand respectively for the probabilities of initial state, survival, and
142 detection. The situation is similar to having two possible paths at first observation: one low-
143 survival and one high-survival path (Fig. 2). The “low survival” path has the greatest
144 probability in “short” histories like ‘110100000000’ and the lowest in “long” histories like
145 ‘1000001000001’ (Fig. 2).

146 In a more complex model, there are a larger number of possibilities when individuals
147 are not encountered. In the presence of several kinds of heterogeneity (survival, emigration,
148 and detection), an individual can be in the low or the high-value class for each type of
149 heterogeneity. There is up to 8 “classes of heterogeneity” (see Appendix S1 part 4) in the
150 models, which greatly complicates the computation of history probabilities. The need for an
151 algorithm to calculate history probabilities should therefore be apparent. We used program E-
152 SURGE 1.1.1 (Choquet et al. 2009a) to obtain maximum likelihood estimates of the
153 parameters and perform model selection. A more rigorous and general development of the
154 likelihood using matrix notation is presented in Appendix S1.

155 In the following, survival is time-independent and detection is time-dependent, based on
156 results of Prévot-Julliard et al. (1998a) and Grosbois (2001).

157 3) *Temporary emigration*

158 We modelled temporary emigration as Markovian or state-dependent transitions to and from a
159 site where detection probability is zero (Fujiwara and Caswell 2002; Schaub et al. 2004; Fig
160 2; Appendix S1 part 2). Colonies that were not searched for marked individuals were grouped
161 in a single non-observable “site” denoted “alive elsewhere” (AE hereafter). This site was
162 included in the usual multisite CMR models formulation (Arnason, 1972, 1973; Schwarz et al.
163 1993) with the only difference that the detection rate was zero. Once in the state AE,
164 individuals lost their classification as low/high detection and low/high emigration. Therefore,

165 upon returning to the study site, they could become more/less detectable, or more/less site-
166 faithful, than what they were before emigrating (see Appendix S1 part 2 for justification).

167 *4) Modelling age-effects on survival when individuals enter the dataset at various ages*

168 The straightforward implementation of age effects in the sense of CMR models (Lebreton et
169 al. 1992) corresponds to the effect of time elapsed since first occurrence in the dataset
170 (hereafter TFC to match a previous acronym: Crespin et al. 2006). In our case gulls were
171 marked as chicks and were thus of known age, but they entered the dataset as adults in the
172 colony at a varying age (Clobert et al. 1994). TFC did thereby not correspond to true age. To
173 model the effect of true age, we had to constrain survival to vary with time across as many
174 groups as there were ages at first occurrence in the dataset (hereafter “group approach”;
175 described in details in Appendix S1 part 2). This procedure was computer-time-hungry
176 [around 24h were needed to fit such a model using an Intel Pentium 4HT, 2.6GHz (3.25 x
177 800) processor with 512 Mb of system memory, vs. less than 1h for a TFC model] and was
178 thus impractical for model selection which required running many models sequentially.

179 Yet using TFC as a proxy for true age in a similar study design does not prevent the
180 detection of survival senescence as shown by Crespin et al. (2006). A test of power (Crespin
181 et al. 2006) indicated that sample size rather than the use of TFC vs. true age is the most
182 critical factor preventing the detection of senescence. The main drawback of using TFC is that
183 individuals of various true ages are mixed in a same TFC-class, thus creating noise and
184 increasing the standard error on the estimation of the strength of senescence. Model selection
185 using TFC was therefore considered conservative for what concerns the detection of
186 senescence. We thus used TFC instead of age to select for the best model, and then confirmed
187 our results by running the preferred model with true age instead of TFC.

188 To represent a biologically sound relationship between age (or TFC) and survival we
189 used a constrained piecewise relationship; we modelled a separate survival probability at age

190 1, a plateau lasting until 14-years old and a log-linear decrease in survival with age afterwards
191 (see Appendix S1 part 2 for justification).

192 5) *Modelling observations that provide information on detectability*

193 As described in the “data collection” section, the confirmation of breeding status conveyed
194 information on the detectability of individuals. Models accommodating such data structure
195 were introduced by Nichols et al. (2004; the individual status was the gender, documented by
196 behaviour) and Conn and Cooch (2009; the individual status was the state of health,
197 documented by visible symptoms). In the present paper, the class of detection heterogeneity
198 was documented by the breeding status (Appendix S1 for practical implementation).

199 **Model selection**

200 There were 16 models in our candidate set representing every combination of
201 presence/absence of the four considered effects: age-effect on survival, heterogeneity in
202 survival, detection and emigration probabilities. The most general model we considered
203 included heterogeneity in survival, emigration and detection probabilities as well as age-effect
204 and had 11 states (model denoted $\{S_{H+a}; D_H; E_H\}$; Appendix S1 part 4). Subscripts H and 0
205 referred to models with and without heterogeneity in survival (denoted S), detection (denoted
206 D) and emigration (denoted E) probabilities, while subscripts $+a$ and $+\bar{a}$ referred to models
207 with and without age-effect on survival. The lowest AIC-model (with a two AIC-points
208 difference) was preferred (Akaike’s Information Criterion; Burnham and Anderson 2002).
209 However, all models in which a given effect was included could be considered as
210 “redundant”. Therefore, the importance value of each effect was computed as the sum of the
211 AIC-weights of the models including the considered effect, and it was interpreted as the
212 probability that the effect was present in the data (Burnham and Anderson 2002).

213 Goodness-of-fit

214 Goodness-of-fit (GOF) tests were performed on raw capture/non capture histories (formed of
215 '0' and '1' thus discarding the details about age at first detection and breeding status; note that
216 all the data concerned a single site). One of the main consequences of heterogeneity in
217 detection probability is an excess (when compared to homogeneous datasets) of encounter
218 histories with consecutive "captures" (i.e. runs of '1') and consecutive "non-captures" (i.e.
219 runs of '0'). Such histories indicate the presence of highly and poorly detectable individuals,
220 respectively. Some of the runs of '0' will occur at the end of the capture history. As a
221 consequence detection heterogeneity tends to induce both "transience" (i.e. lower chance of
222 recapture of first-encountered individuals than already encountered ones; Pradel et al. 1997)
223 and "trap-happiness" (i.e. higher probability to encounter at time $t+1$ the individuals
224 encountered at time t than the individuals not encountered at time t but known to be alive
225 because of previous and future recaptures; Pradel 1993).

226 One-sided directional test statistics are the signed square roots of the chi-squared
227 statistics for the corresponding tests: Test3.SR for transience and Test2.CT for trap-
228 dependence (Pradel 1993; Pradel et al. 1997; Pradel et al. 2005; practical implementation
229 detail in Choquet et al. 2005). They are the most relevant statistics for the detection of
230 transience and trap-happiness respectively (Pradel et al. 2005) and are inflated by individual
231 heterogeneity.

232 We used techniques for partitioning chi-squares variables (Rao, 1973 pp. 185 and
233 following) to approximate a GOF test for a time-dependent model with capture heterogeneity
234 as follow: (i) we computed the overall GOF chi-squared statistics for transience and trap-
235 dependence, from Test3 and Test2 respectively (Pradel 1993; Pradel et al. 1997; Choquet et
236 al. 2005) (ii) we computed the directional statistics, from Test3.SR and Test2.CT respectively
237 (iii) we removed from the overall statistics the corresponding squared directional statistics

238 (which are asymptotically distributed as χ^2_1), and we obtained non-directional components
239 with one degree of freedom less. These components corresponded to a GOF test for a model
240 where sources of transience and trap-happiness (here, individual heterogeneity) were
241 accounted for. If this corrected test still proved statistically significant, we used an
242 overdispersion coefficient \hat{c} , computed as the ratio between the chi-squared statistic and the
243 degree of freedom, in the model selection procedure (Burnham and Anderson 2002). All the
244 GOF test components were computed using U-CARE (Choquet et al. 2009b)

245 **Simulation study**

246 To judge whether the results provided by the proposed methodology were reliable in a
247 complex but known case (three kinds of heterogeneity and a decrease in survival probability
248 with age), we carried out Monte Carlo simulation studies (Appendix S2). Very briefly, these
249 exercises illustrated that the multievent framework allowed (i) detecting simultaneously and
250 using AIC all three kinds of heterogeneities when present, and (ii) obtaining reliable evidence
251 and precise estimates for survival senescence by accounting for these heterogeneities.

252 More precisely the simulation study indicated (i) that the magnitude of the age-effect on
253 survival was very precisely retrieved when the heterogeneity structure in the model exactly
254 matched the simulated structure, and (ii) that bringing additional information on detectability
255 of individuals was sufficient (and necessary) to separate detectability and emigration
256 heterogeneities. In our case, such information was provided by the confirmation of breeding
257 status. The main drawbacks of the method were the non-reliability of emigration probability
258 estimates (the presence of emigration heterogeneity was retrieved, but the actual values of the
259 parameters were not), and the fact that, most probably because the data on them were sparse,
260 estimates for parameters associated to low-survival individuals were often inaccurate.

261 **Results**

262 **Black-headed gull dataset: Goodness-of-fit**

263 As expected, the directional tests for transience and trap-happiness were statistically
264 significant, which is a cue for heterogeneity in detection (Table 1). The corrected Test3.SR
265 and Test2.CT were statistically non significant, and the overall corrected GOF test indicated
266 that there was no need to account for any overdispersion (Table 1). This test indicated that a
267 time-dependent model with heterogeneity in detection probability fitted the data. Thus,
268 accounting for other sources of heterogeneity could only improve the fit.

269 **Modelling age-dependence and testing for senescence**

270 *1) Model selection using TFC*

271 There was strong support for a model with individual heterogeneity in both detection
272 and emigration probabilities, along with TFC-effect on survival (Table 2: model
273 $\{S_{0+a};D_H;E_H\}$). This model was nearly four AIC-points lower and three times more likely than
274 the next model (as indicated by the ratio of AIC-weights). The importance value (computed as
275 the sum of the AIC-weights of the models in which the considered effects occurred) of the
276 TFC-effect on survival probability was 0.83, which we interpret as a high probability for a
277 decrease in survival with age. These results supported the existence of survival senescence in
278 the population, although the slope of the TFC-effect was statistically not different from zero: -
279 0.16 (95% CI: -0.49; 0.17); see next section.

280 Importance values of heterogeneity in survival, detection and emigration were 0.09,
281 0.94, and 0.78, respectively. We interpret these values as high probability that two-class
282 heterogeneity structure was present in detection and emigration probabilities, and low
283 probability that such heterogeneity was present in survival probability. As discussed later, we

284 do not exclude that the discrete-class heterogeneity models were unable to detect small,
285 continuous individual variation in survival probability.

286 The second best model (Table 2: model $\{S_{0+\bar{a}}; D_H; E_0\}$), as opposed to the best model,
287 did not account for heterogeneity in emigration probability and, most importantly, did not
288 include any variation in survival probability with TFC. This result means that, when not
289 accounting for heterogeneity in emigration probabilities, the selection procedure discarded
290 TFC-effect on survival, in other words the detection of survival senescence was prevented.

291 2) *Parameter estimates in the true age formulation*

292 Parameter estimates are from the preferred model $\{S_{0+a}; D_H; E_H\}$, ran using true age instead of
293 TFC. Prime age survival (between 2 and 14 years old) was 0.84 (95% CI: 0.79; 0.88) and the
294 slope of the decrease in survival after age 14 was -0.16 (-0.30; -0.02) on a logit scale (Fig. 3).
295 Thus, when comparing true age- (this model) and TFC-models (previous section), we
296 observed that the use of TFC increased the error on the estimated slope of the decrease in
297 survival with age, as expected, but did not modify the value of the estimate.

298 The estimated temporary emigration probabilities to state AE were 0.17 (0.03; 0.54) in
299 stayers and 0.67 (0.28; 0.91) in movers. The estimated detection probabilities (averaged over
300 time) were 0.08 (0.05; 0.10) and 0.48 (0.41; 0.55) in the low- and high-detectability classes.
301 The estimated proportion of sightings on the nest or with chicks was 0.53 (0.44; 0.63) for the
302 low-detectability class and 0.61 (0.56; 0.66) for the high-detectability class.

303 **Discussion**

304 **Detecting senescence in wild populations**

305 We used multievent framework to combine three pre-existing types of CMR models that were
306 potentially required to fit the gull data: heterogeneity models with discrete classes (Pledger et
307 al. 2003; Pradel 2009), temporary emigration models (Fujiwara and Caswell 2002; Schaub et

308 al. 2004), and models accommodating partial information on individual status (Nichols et al.
309 2004; Conn and Cooch 2009). The framework was flexible enough to combine these model
310 structures which had only been used separately until now.

311 Survival senescence in our population was only detected after accounting for
312 heterogeneity in temporary emigration in the models (Table 2), which illustrates the interest
313 of complex models mixing several possible causes of heterogeneity. This result was perhaps
314 not unexpected since temporary emigration events occurring at the end of an individual's life,
315 when not followed by a detection event, could mimic earlier death. Thereby, emigration
316 heterogeneity might have created patterns in the data similar to heterogeneity in survival, and
317 prevented the detection of survival senescence when not accounted for. However, our
318 simulation studies only weakly supported this purported effect of non modelled heterogeneity
319 in temporary emigration on the detection of survival senescence. The result that senescence
320 was discarded when not accounting for emigration heterogeneity might therefore originate
321 from a lack-of-fit impairing the model selection when models did not include the appropriate
322 heterogeneity structure.

323 Definitive emigration has evidently an even greater impact on survival estimates than
324 temporary emigration, but it can only be modelled if recoveries data (i.e. rings recovered on
325 birds shot or found dead throughout the year) are available (Burnham 1993).

326 In senescence studies, we thus recommend that (i) emigration is modelled when field
327 observations indicate its presence and (ii) either individual heterogeneity is explicitly
328 considered or evidence for its absence is provided by GOF tests or biological considerations.
329 In particular, empirical support for the absence of senescence in animals (e.g. Nichols et al.
330 1997; Miller 2001; Pistorius and Bester 2002; Congdon et al. 2003) should be considered with
331 caution until the results are verified with analyses accounting for heterogeneity.

332 Sources of heterogeneity in the gull dataset: emigration, detectability

333 Although there was a strong support for the existence of low- and high- emigration classes in
334 our population, biological explanations are not straightforward. Emigration heterogeneity
335 might also include the skipping of breeding attempts and the effect of early nest failures.
336 These two phenomena result in an absence of the individuals from the colonies during field-
337 work (thereby mimicking temporary emigration), and they are influenced by individual
338 quality (Calladine and Harris 1997), a well known source of heterogeneity in demographic
339 parameters (Hamel et al. 2008 and references therein). Moreover, true temporary emigration,
340 reproduction failures and reproduction skipping do not occur at the same frequency which
341 might create individual heterogeneity when they are modelled using a same transition
342 probability. Alternatively, individual heterogeneity in true temporary emigration rate could
343 stem from sex-biased dispersal (Greenwood, 1980) or from heterogeneity in individual quality
344 and/or the conditions experienced during early life or previous reproduction attempts, both of
345 which are known to affect habitat choice and dispersal behaviours (Switzer 1997; Clobert et
346 al. 2009). Last, and perhaps more speculatively, heterogeneity in dispersal behaviour can
347 result from genetically determined differences in “personalities” or other behavioural
348 syndromes (see e.g. Cote & Clobert, 2007; Clobert *et al.* 2009). Tradeoffs between investment
349 in exploration-dispersal and in other traits (Wolf et al. 2007) or frequency-dependent selection
350 acting through environmental stochasticity (see Dingemanse et al. 2004) can maintain stable
351 polymorphism in dispersal tendencies.

352 Heterogeneity in detection probability was probably related to the fact that high
353 vegetation density hindered the detection of birds breeding far inside the vegetation. These are
354 likely to have strongly contributed to the class with a low detection probability and a low
355 proportion of sightings on a nest. Such heterogeneity in detection induced by habitat

356 heterogeneity in the study site is supposedly common, and we recommend it be explicitly
357 incorporated.

358 **Individual heterogeneity of unknown origin: modelling approach**

359 As in most modelling exercises, our study relied on some untested assumptions. Modelling
360 individual heterogeneity using discrete classes was *a priori* appropriate in our study: birds
361 could breed inside vegetation vs. outside vegetation, their movements could occur within vs.
362 outside of the colony. However, we do not claim that the discrete-class model represented
363 individual variation in survival probabilities better than a continuous individual random effect
364 could (e.g. Royle 2008, but see Pledger 2005). Concerning the number of heterogeneity
365 classes, Pledger (2005) indicated that, as a theoretic and approximated representation of
366 individual heterogeneity, the two-class models were more parsimonious than models with
367 more classes, excepted in the presence of strong multimodality in the true distribution of the
368 parameters of interest. Standard quantitative tools such as AIC generally fail to separate
369 different forms of heterogeneity models (own results not shown; Pledger 2005). We suggest
370 that information on the study system from the field might be the most reliable cue when
371 deciding which form of individual heterogeneity to incorporate in CMR models.

372 **Performance and identifiability issues**

373 The simulation studies illustrated that reliable evidence for survival senescence could be
374 obtained in the simultaneous presence of three kinds of heterogeneity. When the data included
375 events whose frequency depended on detectability, the three sources of heterogeneity were
376 separately identifiable, although in most cases only the structure and not the actual values of
377 emigration probabilities were retrieved. The magnitude of the age-effect was very precisely
378 retrieved when the heterogeneity structure in the model exactly matched the way data was
379 simulated. In the gull application we did not detect any parameter redundancy issue using the

380 rank of the models' numeric derivative matrix (Rouan et al. 2009 Appendix A; Choquet et al.
381 2009a pp. 56-57). Overall our results are therefore encouraging for the application of
382 multievent models to the study of population dynamics of species with complex life-histories,
383 weak or variable site-fidelity, or inhabiting very heterogeneous habitats.

384

385 **References**

- 386 Arnason, A.N. 1972. Parameter estimates from mark-recapture experiments on two
387 populations subject to migration and death. – Res. Popul. Ecol. 13: 97-113.
- 388 Arnason, A.N. 1973. The estimation of population size, migration rates and survival in a
389 stratified population. – Res. Popul. Ecol. 15: 1-8.
- 390 Burnham, K. P. 1993. A theory for combined analysis of ring recovery and recapture data. –
391 In: Lebreton, J. D. and North, P. H. (eds.), Marked individuals in the study of bird
392 population. Birkhäuser Verlag, Basel, Switzerland.
- 393 Burnham, K.P. and Anderson, D.R. 2002. Model selection and multimodel inference: a
394 practical information-theoretic approach, 2nd edition – Springer-Verlag, New York.
- 395 Calladine, J. and Harris, M.P. 1997. Intermittent breeding in the Herring gull *Larus*
396 *argentatus* and the Lesser black-backed gull *Larus fuscus*. – Ibis 139: 259–63.
- 397 Cam, E. et al. 2002. Individual covariation in life–history traits: seeing the trees despite the
398 forest. – Am. Nat. 159: 96–105.
- 399 Choquet, R. 2008. Automatic generation of multistate capture–recapture models. – Can. J.
400 Stat. 36: 43–57.
- 401 Choquet, R. et al. 2005. U-CARE 2.2 User's Manual – CEFE, Montpellier, France.
- 402 Choquet, R. et al. 2009a. Program E-SURGE: a software application for fitting Multievent
403 models. – In: Thomson, D. L. Cooch, E. G. and Conroy, M. J. (eds.), Modelling
404 Demographic Processes in Marked Populations, pp. 845–865.

- 405 Choquet, R. et al. 2009b. U-CARE: Utilities for performing goodness of fit tests and
406 manipulating CAPture-REcapture data. *Ecography*. In press.
- 407 Clobert, J. et al. 1994. The estimation of age-specific breeding probabilities from recaptures
408 or resightings in vertebrate populations.2. Longitudinal models. – *Biometrics* 50: 375–87.
- 409 Clobert, J. et al. 2009. Informed dispersal, heterogeneity in animal dispersal syndromes and
410 the dynamics of spatially structured populations. – *Ecol. Lett.* 12: 197-209.
- 411 Congdon, J.D. et al. 2003. Testing hypotheses of aging in long-lived painted turtles
412 (*Chrysemys picta*). – *Exp. Gerontol.* 38: 765–72.
- 413 Conn, P.B. and Cooch, E.G. 2009. Multistate capture–recapture under imperfect state
414 observation: an application to disease models. – *J. Appl. Ecol.* 46: 486-492.
- 415 Cote, J. and Clobert, J. 2007. Social personalities influence natal dispersal in a lizard. - P.
416 *Roy. Soc. Lond. B Bio.* 274: 383-390.
- 417 Crespin, L. et al. 2008. Is heterogeneity of catchability in capture–recapture studies a mere
418 sampling artefact or a biologically relevant feature of the population? – *Popul. Ecol.* 50:
419 247-256.
- 420 Crespin, L. et al. 2006. Increased adult mortality and reduced breeding success with age in a
421 population of Common guillemot *Uria aalge* using marked birds of unknown age. – *J.*
422 *Avian Biol.* 37: 273-282.
- 423 Dingemanse, N. J. Both, C. Drent, P. J. et al. 2004. Fitness consequences of avian
424 personalities in a fluctuating environment. – *P. Roy. Soc. Lond. B Bio.* 271: 847-852.
- 425 Finch, C. E. 1990. Longevity, senescence, and the genome. – University of Chicago Press,
426 Chicago, USA.
- 427 Fox, G.A. et al. 2006. Consequences of heterogeneity in survival probability in a population
428 of Florida scrub-jays. – *J. Anim. Ecol.* 75: 921-927.

- 429 Fujiwara, M. and Caswell, H. 2002. A general approach to temporary emigration in mark–
430 recapture analysis. – Ecology 83: 3266-3275.
- 431 Gimenez, O. et al. 2008. The risk of flawed inference in evolutionary studies when
432 detectability is less than one. – Am. Nat. 172: 441-448.
- 433 Greenwood, P.J. 1980. Mating systems, philopatry and dispersal in birds and mammals. –
434 Anim. Behav. 28: 1140-1162.
- 435 Grosbois, V. 2001. La dispersion : trait d'histoire de vie et paramètre démographique : étude
436 empirique dans une population de mouette rieuse. PhD thesis – Univ. Montpellier II,
437 Montpellier, France. In French.
- 438 Hamel, S. et al. 2008. Individual variation in reproductive costs of reproduction: high-quality
439 females always do better. – J. Anim. Ecol. 78: 143-151.
- 440 Jones, O. R., Gaillard, J.-M., Tuljapurkar, S., et al. 2008. Senescence rates are determined by
441 ranking on the fast-slow life-history continuum. – Ecol. Lett. 11: 664-673.
- 442 Lebreton, J.D. 1987. Régulation par le recrutement chez la Mouette rieuse *Larus ridibundus*.
443 – Rev. Ecol.- Terre Vie Suppl. 4: 173-187. In French.
- 444 Lebreton, J.D. et al. 1992. Modelling survival and testing biological hypotheses using marked
445 animals - a unified approach with case-studies. – Ecol. Monogr. 62: 67-118.
- 446 Loison, A. et al. 1999. Age-specific survival in five populations of ungulates: evidence of
447 senescence. – Ecology 80: 2539-2554.
- 448 MacDonald, I. L. and Zucchini, W. 1997. Hidden Markov and other models for discrete-
449 valued time series. - Chapman and Hall, London.
- 450 Medawar, P.B. 1952. An unsolved problem of biology. – Lewis, London.
- 451 Miller, J.K. 2001. Escaping senescence: demographic data from the Three-toed box turtle
452 (*Terrapene carolina triunguis*). – Exp. Gerontol. 36: 829-832.

- 453 Nichols, J.D. et al. 1997. Test for senescent decline in annual survival probabilities of
454 Common pochard, *Aythya ferina*. – Ecology 78: 1009-1018.
- 455 Nichols, J.D. et al. 2004. Estimation of sex-specific survival from capture-recapture data
456 when sex is not always known. – Ecology 85: 3192-3201.
- 457 Nussey, D.H. et al. 2006. The rate of senescence in maternal performance increases with
458 early-life fecundity in red deer. – Ecology Letters 9: 1342-1350.
- 459 Pistorius, P.A. and Bester, M.N. 2002. A longitudinal study of senescence in a pinniped. –
460 Can. J. Zool. 80: 395-401.
- 461 Pledger, S. 2000. Unified maximum likelihood estimates for closed capture-recapture models
462 using mixtures. - Biometrics 56: 434-442.
- 463 Pledger, S. 2005. The performance of mixture models in heterogeneous closed population
464 capture-recapture. – Biometrics 61: 868-873.
- 465 Pledger, S. et al. 2003. Open capture–recapture models with heterogeneity: I. Cormack-Jolly-
466 Seber model. – Biometrics 59: 786-794.
- 467 Pradel, R. 1993. Flexibility in survival analysis from recapture data: handling trap-
468 dependence. – In: Lebreton, J. D. and North, P. H. (eds.), Marked individuals in the study of
469 bird population, pp. 29-37.
- 470 Pradel, R. 2005. Multievent: an extension of multistate capture-recapture models to uncertain
471 states. – Biometrics 61: 442-447.
- 472 Pradel, R. 2009. The stakes of capture-recapture models with state uncertainty. – In:
473 Thomson, D. L. Cooch, E. G. and Conroy, M. J. (eds.), Modeling Demographic Processes in
474 Marked Populations, pp. 781-795.
- 475 Pradel, R. et al. 2005. Principles and interests of GOF tests for multistate capture–recapture
476 models. – Anim. Biodivers. and Conserv. 28: 189-204.

- 477 Pradel, R. et al. 1997. Capture-recapture survival models taking account of transients. –
478 *Biometrics* 53: 60–72.
- 479 Pradel, R. et al. 2007. Estimation of sex-specific survival with uncertainty in sex assessment.
480 – *Can. J. Stat.* 36: 29-42.
- 481 Prévot-Julliard, A.C. et al. 1998a. Re-evaluation of adult survival of Black-headed gulls
482 (*Larus ridibundus*) in presence of recapture heterogeneity. – *Auk* 115: 85-95.
- 483 Prévot-Julliard, A.C. et al. 1998b. Evidence for birth-site tenacity in breeding Common black-
484 headed gulls *Larus ridibundus*. – *Can. J. Zool.* 76: 2295-2298.
- 485 Pugsek, B.H. et al. 1995. Mark-resighting analysis of a California gull population. – *J. Appl.*
486 *Stat.* 22: 625-39.
- 487 Rao, C.R. 1973. Linear statistical inference and its applications. – Wiley, New York
- 488 Regehr, E.V. et al. 2007. Effects of earlier sea ice breakup on survival and population size of
489 Polar Bear in Western Hudson Bay. – *J. Wildlife Manage.* 71: 2673-2683.
- 490 Rouan, L. et al. 2009. A general framework for modeling memory in Capture-Recapture data.
491 *J. Agric. Biol. Envir. S.* In press.
- 492 Royle, J.A. 2008. Modeling individual effects in the Cormack-Jolly-Seber model: a state-
493 space formulation. – *Biometrics* 64: 364-370.
- 494 Schaub, M. et al. 2004. Estimating survival and temporary emigration in the multistate
495 capture-recapture framework. – *Ecology* 85: 2107-2113.
- 496 Schwarz, C.J. et al. 1993. Estimating migration rates using tag-recovery data. – *Biometrics*
497 49: 177-193.
- 498 Switzer, P.V. 1997. Past reproductive success affects future habitat selection. – *Behav. Ecol.*
499 *Sociobiol.* 40: 307-312.
- 500 van de Pol, M. and Verhulst, S. 2006. Age-dependant traits: a new statistical model to
501 separate within- and between-individual traits. – *Am. Nat.* 167: 766-773.

- 502 Vaupel, J.W. and Yashin, A.I. 1985. Heterogeneity's ruses: some surprising effects of
503 selection on population dynamics. – Am. Stat. 39: 176-185.
- 504 Whitehead, H. and Wimmer, T. 2005. Heterogeneity and the mark-recapture assessment of
505 the Scotian Shelf population of northern bottlenose whales (*Hyperoodon ampullatus*). –
506 Can. J. Fish. Aquat. Sci. 62: 2573-2585.
- 507 Wolf, M. van Doorn, G. S. Leimar, O. et al. 2007. Life-history trade-offs favour the evolution
508 of animal personalities. - Nature 447: 581-584.
- 509 Zens, M.S. and Peart, D.R. 2003. Dealing with death data: individual hazards, mortality and
510 bias. – Trends Ecol. Evol. 18: 366-373.

511

512 **Supplementary material**

513 The following supplementary material is available for this article.

514 **Appendix S1.** Model description

515 **Appendix S2.** Simulation study

516

517

518 **Tables**

519 Table 1: Components of an approximate goodness-of-fit (GOF) test for a model with
 520 heterogeneity, obtained by removing from the components of the standard GOF test for the
 521 time-dependent model the squared directional test statistics (see methods). The overall GOF
 522 test shows no sign of lack-of-fit for a model correcting transience and trap-happiness. df is the
 523 degree of freedom. \hat{c} is the overdispersion coefficient computed as the ratio between the chi-
 524 squared statistic and the degree of freedom.

525

	Test3: transience		Test2: trap dependence		Total
	Chi- squared statistic	Squared directional statistic	Chi- squared statistic	Squared directional statistic	Chi- squared statistic
Time-dependent model	64.19	38.67	25.88	9.21	164.39
df	22	1	22	1	131
P-level	< 0.0001	< 0.0001	0.26	0.002	0.02
\hat{c}	2.92		1.18		1.25
Time-dependent model with heterogeneity of detection	25.52		16.67		116.52
df	21		21		129
P-level	0.23		0.73		0.78
\hat{c}	1.22		0.79		0.90

526

527

528 Table 2: Model selection. The 16 candidate models vary in the presence/absence (Y/N)
 529 of heterogeneity and of the age-effect on survival. For each model the number of parameters
 530 (np), deviance (Dev), AIC and AIC-weight are given. Subscripts H and 0 referred to models
 531 with and without heterogeneity in survival (denoted S), detection (denoted D) and emigration
 532 (denoted E) probabilities, while subscripts a and \bar{a} referred to models with and without age-
 533 effect on survival. The models are sorted by AIC.

model	Heterogeneity in:			Age-	np	Dev	AIC	AIC weight
	Survival	Detection	Emigration	effect				
$\{S_{0+a};D_H;E_H\}$	N	Y	Y	Y	43	7781.88	7867.88	0.761
$\{S_{0+\bar{a}};D_H;E_0\}$	N	Y	N	N	36	7799.44	7871.44	0.128
$\{S_{H+a};D_H;E_0\}$	Y	Y	N	Y	47	7779.50	7873.50	0.046
$\{S_{H+\bar{a}};D_0;E_0\}$	Y	N	N	N	35	7803.62	7873.62	0.043
$\{S_{0+a};D_0;E_H\}$	N	N	Y	Y	37	7802.12	7876.12	0.012
$\{S_{H+a};D_0;E_0\}$	Y	N	N	Y	38	7801.99	7877.99	0.005
$\{S_{0+\bar{a}};D_H;E_H\}$	N	Y	Y	N	40	7800.28	7880.28	0.002
$\{S_{0+a};D_H;E_0\}$	N	Y	N	Y	41	7797.91	7879.91	0.002
$\{S_{H+a};D_0;E_H\}$	Y	N	Y	Y	42	7797.25	7881.25	0.001
$\{S_{H+a};D_H;E_H\}$	Y	Y	Y	Y	53	7777.12	7883.12	0.000
$\{S_{H+\bar{a}};D_H;E_0\}$	Y	Y	N	N	43	7802.21	7888.21	0.000
$\{S_{0+\bar{a}};D_0;E_0\}$	N	N	N	N	31	7856.10	7918.10	0.000
$\{S_{0+a};D_0;E_0\}$	N	N	N	Y	34	7820.42	7888.42	0.000
$\{S_{H+\bar{a}};D_0;E_H\}$	Y	N	Y	N	37	7852.13	7926.13	0.000
$\{S_{H+\bar{a}};D_H;E_H\}$	Y	Y	Y	N	49	7834.15	7932.15	0.000
$\{S_{0+\bar{a}};D_0;E_H\}$	N	N	Y	N	34	7857.33	7925.33	0.000

535

536 **Figure legend**

537 Fig. 1: Example heterogeneous population where the average survival rate is not
538 representative of the true ageing processes. This example population consists of 10% of
539 “initially robust” individuals and 90% of “initially weak” individuals, each subpopulation
540 experiencing the same slow decrease in survival with age. (a) Age variation in the true
541 survival rate for the two subpopulations (grey lines) and in the average survival rate for the
542 population (black line). (b) Age variation in the proportion of robust individuals.

543

544 Fig. 2: A tree diagram describing the hidden Markov chain probability structure for a
545 model with a two-class individual heterogeneity structure on survival probabilities, the
546 possibility to temporarily emigrate, and a detailed observation structure for the breeding
547 status. Black boxes indicate the 5 possible states (S1: alive in the study site with high survival
548 probability; S2: alive in the study site with low survival probability; AE1: alive outside the
549 study site with high survival probability; AE2: alive outside the study site with low survival
550 probability; dead), while grey boxes represent the 3 possible observations following initial
551 release (0: not seen, 1: seen and breeding status confirmed; 2: seen but breeding status not
552 confirmed). The probability for observing a particular encounter history is obtained by
553 summing the probability of all possible paths leading to a given encounter history (for the
554 sake of clarity at time $t+1$ only the states “AE” are represented; paths from states “S” are the
555 same as at time t and the state “dead” is absorbing, i.e. there is no path out of it). The
556 probability of a given path can be obtained by multiplying the probabilities appearing
557 alongside its component arrows. These probabilities consist of functions of π , the initial state
558 probabilities; ϕ , apparent survival probabilities; ψ , state transition probabilities; p , detection

559 probabilities; and r , the probabilities to confirm breeding status. A more formal matrix
560 description of the same model is provided in Table 1.

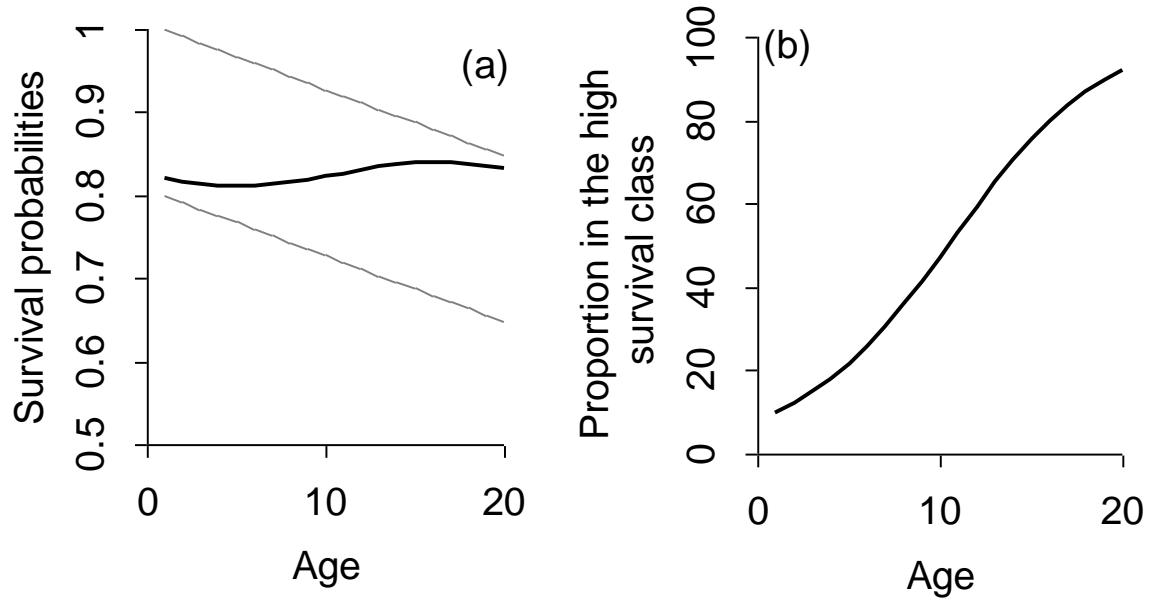
561

562 Fig. 3: Survival probabilities for the Black-headed gulls as a function of true age.
563 Parameters estimates are from model $\{S_{0+a}; D_H; E_H\}$ which includes heterogeneity in both
564 detection and emigration probabilities, and a piecewise constrained relationship between age
565 and survival probabilities. Dotted lines correspond to the 95% asymptotic confidence interval.
566 Black dots are boundary estimates, which come with no standard errors. The survival
567 probability for the first age class (age 2) was estimated separately and fell on the line for older
568 age classes.

569

570 **FIGURE**

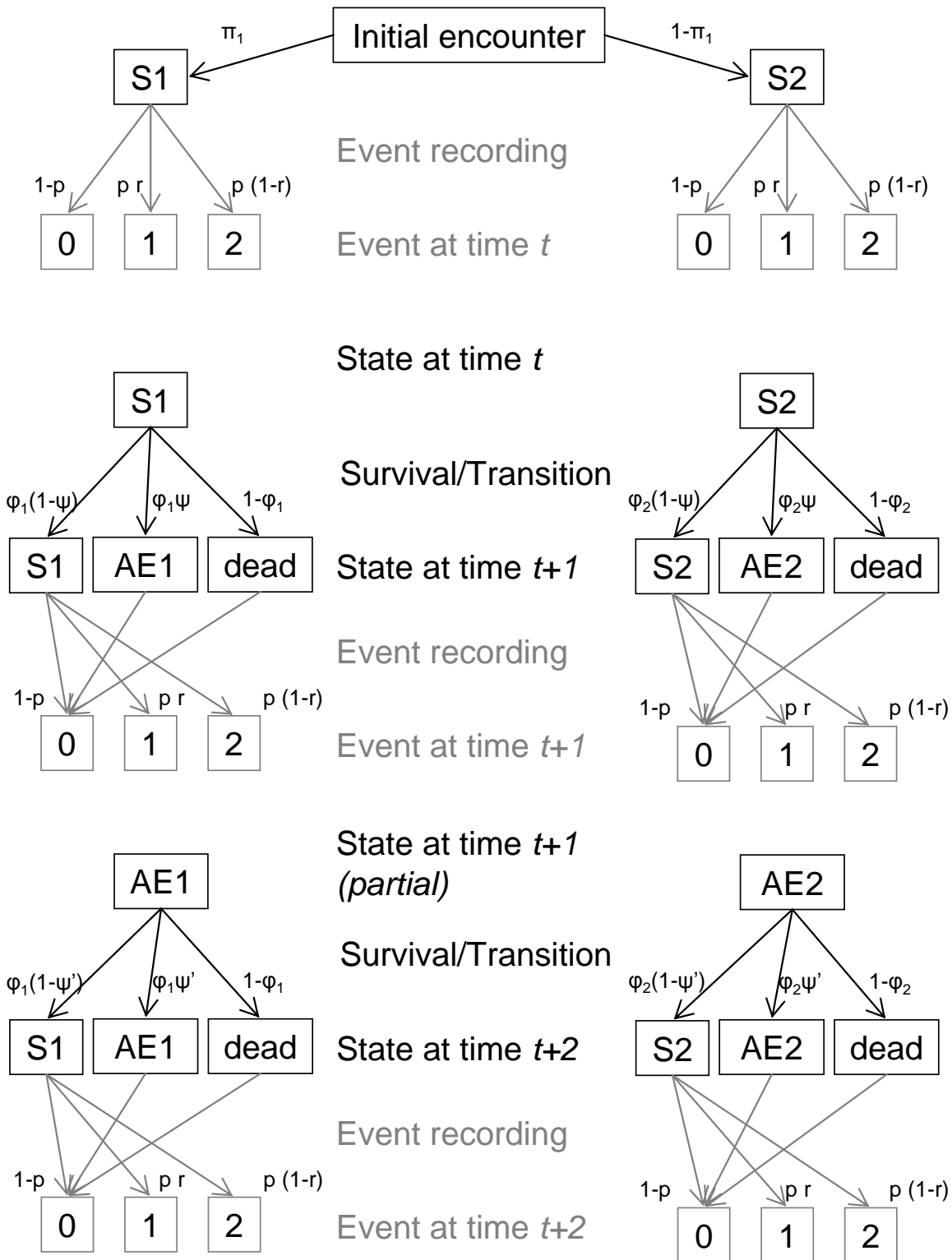
571 Figure 1



572
573

574
575
576

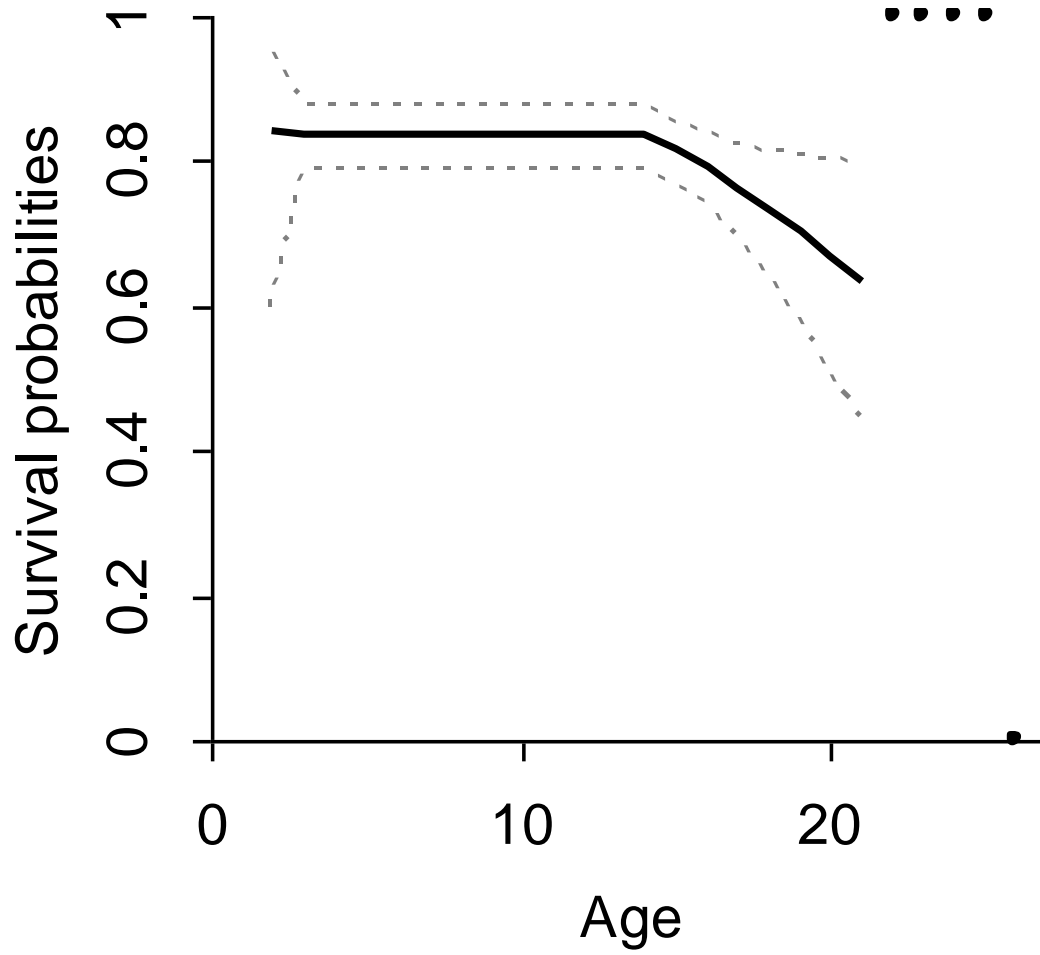
Figure 2



577
578
579
580

581
582
583
584
585

Figure 3



586
587
588

589 **Appendix S1: Model description**

590

591 **Part 1. Individual heterogeneity in multievent models**

592 There were 2^k observable states (k being the number of parameters types affected by
593 heterogeneity in the model), plus one or two AE-states depending on the presence of survival
594 heterogeneity in the model, and the state dead. We describe in this appendix two of the
595 considered models: (i) the model $\{S_H; E_0; D_0\}$ including heterogeneity in survival
596 probabilities only ($2^1+2+1=5$ states; part 3) and (ii) the model $\{S_H; E_H; D_H\}$ including the
597 three kinds of heterogeneity (survival, detection and emigration: $2^3+2+1=11$ states; part 4).

598 When considering detection heterogeneity, we did not allow individuals to move
599 between high- and low-detection classes, based on the finding that individuals tend to be site-
600 faithful on a small scale when breeding in LR (Prévot-Julliard et al. 1998b). However, when a
601 bird emigrated, the model did not keep memory of its previous classes of detection and
602 emigration, letting it free to change either of these classes upon return. The model permitted
603 returning individuals to settle in a different zone of the pond or to change of dispersal
604 behaviour.

605

606

607 **Part 2. Parameterisation**

608 The complex model structure rendered necessary the use of some constraints on the
609 parameters. These and the effects that were considered in the model selection procedure are
610 presented below:

611 *1) – Initial proportions*

612 The Π -vector represents the probability of being in the different classes of
613 heterogeneity at first detection. We assumed that Π was constant over time, hence assuming
614 that the proportions of the different classes of heterogeneity did not vary over time.

615 Preliminary results (not shown) indicated a poor fit of a model in which Π varied with
616 age, probably because the small number of individuals entering the dataset after age 14 (onset
617 of senescence) rendered some Π -parameters non-estimable. Moreover, when Π varied with
618 age, there might have been identifiability issues because high-survival individuals entering
619 the dataset at old age do have a low survival indeed.

620 *2) – Dispersal probabilities*

621 Dispersal probabilities were modelled as constant over time, for parsimony and
622 because parameter identifiability required that the probability of returning from the
623 unobservable state to the study colony was kept constant over time and that the survival
624 probabilities were the same in the states AE and in the states available for detection (Fujiwara
625 and Caswell 2002). Heterogeneity in emigration probability was modelled in a fashion similar
626 to “mover–stayer models” that consider a class with high site–tenacity and a “volatile” class
627 (Goodman 1961). Once in the state AE, individual lost their classification as low/high
628 detection and low/high emigration. Therefore, after emigration, they could become more
629 detectable, or more site-faithful, and reversely (see Table S3)

630 3) – *Detection probabilities*

631 The matrix of the probabilities of being detected when alive and present on the study
632 site (P) was modelled as time–dependent to fit with the known variation in resighting effort
633 over years (Prévot-Julliard et al. 1998a). Heterogeneity was modelled in a fashion additive to
634 time. Given the high small-scale site fidelity exhibited by breeders within the study colony
635 (Prévot-Julliard et al. 1998b), exchanges between detection classes were expected to be low
636 and we chose not to allow for direct transitions between heterogeneity classes. Indirect
637 transitions were possible after temporary emigration.

638 4) – *Breeding status confirmation probabilities*

639 They were modelled as constant over time in order to keep parsimonious. Indeed, if
640 detection probabilities varied with time because of between-year variation in detection effort,
641 breeding status confirmation probabilities were supposedly more related to the configuration
642 of the colony (distribution of the roost sites, density of the vegetation), and could be
643 considered time-invariant in a first approach. If there was heterogeneity in detection
644 probabilities in the model considered, breeding status confirmation probability was modelled
645 as depending on detectability.

646 5) – *Survival probabilities: modelling age–effect*

647 We considered the age–effect on survival as being additive (on a logit scale) to the
648 heterogeneity level, if any.

649 Models with a complete age–effect (age or TFC treated as categorical variables: 29
650 levels in our case) are the straightforward way to model age–dependency on survival because
651 they make no hypothesis on the shape of that dependency. However, studying the strength of
652 senescence requires the estimation of a slope describing the decrease in survival probabilities
653 with age (Jones et al. 2008), something that can be computed using age or TFC as a

654 continuous variable in a constrained log–linear relationship that moreover reduces the number
655 of parameters in the model (e.g. Loison et al. 1999; Nichols et al. 1997).

656 However, the sparseness of data at old ages (only 14 individuals were observed at an
657 age ≥ 22 , of which none was seen after age 25) renders maximum likelihood estimators
658 unstable, and produces boundary survival estimates (boundary estimates are age–specific
659 survival probabilities that are estimated at 0 or 1 because of low sample size). For example,
660 the occasional individual who will survive 1 year after the others will induce a survival
661 probability of 1 at that age. These boundary estimates are likely to hamper the accurate
662 detection and estimation of the strength of survival senescence, particularly when using
663 constrained regressions (Jones et al. 2008). We first ran a model with complete age–effect,
664 and it appeared that the age–threshold after which boundary estimates were an issue was 22.
665 Boundary estimates were excluded from the constrained relationship presented below.

666 From age 1 to 22, age–dependency was modelled using a three–parameter piecewise
667 relationship. Age one had a separate survival estimate (because we expected it to be lower;
668 e.g. Pugsek et al. 1995 in a related species). Then, we modelled a plateau that lasted up to an
669 age–threshold (which is species–specific: Jones et al. 2008). It was estimated to be 14–year
670 old in black–headed gull by mean of life–history comparisons (G.P. et al. unpublished
671 results); in other Larid species, the threshold was found around 14 years in *Larus californicus*
672 (Pugsek et al. 1995) and between 13 and 16 in *Rissa tridactyla* (Aebischer and Coulson
673 1990; Frederiksen et al. 2004). Afterwards we modelled a logit–linear decline in survival
674 probabilities (Loison et al. 1999).

675 6) – *Implementation of true age in CMR models when first capture occurs at a varying age*

676 True age (as opposed to TFC, see main text) was implemented with a computer
677 intensive procedure. E–SURGE uses the GEMACO syntax to constrain the parameters. The
678 phrase required for modelling the dependency of CMR parameters on true age, when

679 individuals enter the dataset at various ages, was based upon the use of groups. Individuals
 680 entering the dataset at age A were in the A^{th} group. In the A^{th} group at the t^{th} session after their
 681 first observation, individuals had the age $A+t$. The syntax for this relationship, to be entered in
 682 GEMACO, is started below:

```
683      {g(1).a(1)   +   g(1).a(2)&g(2).a(1)   +   g(1).a(3)&g(2).a(2)&g(3).a(1)+...
684      age1           age2           age3           ...
```

685 Further details can be found in (Choquet 2008). We provide below a short and flexible
 686 R-program (R Project Core Team 2008) that generates this syntax for any number of age
 687 classes:

```
688 RES<-" "
689 #maximum age at first entry in the dataset:
690 Amax=23
691 for (A in 2:Amax) {
692   for (i in 2:A) {
693     RES=paste(RES, "g(", as.character(i-
694     1), ").a(", as.character(A+1-i), ")&", sep = " ")
695   }
696   RES=paste(RES, "+", sep=" ")
697 }
698 #maximum potential age:
699 Apot=29
700 for (A in (Amax+1):Apot) {
701   for (i in 2:Amax) {
702     RES=paste(RES, "g(", as.character(i-
703     1), ").a(", as.character(A+1-i), ")&", sep = " ")
704   }
705   RES=paste(RES, "+", sep=" ")
706 }
707 RES<- gsub("(\\+) (\\})", " }",
708   paste("{", gsub("(\\) &(\\+) ", ")+", RES), "}")
709 )
710 RES
```

711 With the values 23 and 29 for Amax and Apot (these values correspond to the black-
 712 headed gull dataset), we obtained the following sentence to be entered as a shortcut for True
 713 Age in GEMACO:

```
714
715 {g(1).a(1)+g(1).a(2)&g(2).a(1)+g(1).a(3)&g(2).a(2)&g(3).a(1)+g(1).a(4)&g(2).a(3)&g(3).a(2)&g(4).a(1)+g(1).a
716 (5)&g(2).a(4)&g(3).a(3)&g(4).a(2)&g(5).a(1)+g(1).a(6)&g(2).a(5)&g(3).a(4)&g(4).a(3)&g(5).a(2)&g(6).a(1)+g
717 (1).a(7)&g(2).a(6)&g(3).a(5)&g(4).a(4)&g(5).a(3)&g(6).a(2)&g(7).a(1)+g(1).a(8)&g(2).a(7)&g(3).a(6)&g(4).a(
```

718 5)&g(5).a(4)&g(6).a(3)&g(7).a(2)&g(8).a(1)+g(1).a(9)&g(2).a(8)&g(3).a(7)&g(4).a(6)&g(5).a(5)&g(6).a(4)&g(7).a(3)&g(8).a(2)&g(9).a(1)+g(1).a(10)&g(2).a(9)&g(3).a(8)&g(4).a(7)&g(5).a(6)&g(6).a(5)&g(7).a(4)&g(8).a(3)&g(9).a(2)&g(10).a(1)+g(1).a(11)&g(2).a(10)&g(3).a(9)&g(4).a(8)&g(5).a(7)&g(6).a(6)&g(7).a(5)&g(8).a(4)
 720 &g(9).a(3)&g(10).a(2)&g(11).a(1)+g(1).a(12)&g(2).a(11)&g(3).a(10)&g(4).a(9)&g(5).a(8)&g(6).a(7)&g(7).a(6)
 721 &g(8).a(5)&g(9).a(4)&g(10).a(3)&g(11).a(2)&g(12).a(1)+g(1).a(13)&g(2).a(12)&g(3).a(11)&g(4).a(10)&g(5).a
 722 (9)&g(6).a(8)&g(7).a(7)&g(8).a(6)&g(9).a(5)&g(10).a(4)&g(11).a(3)&g(12).a(2)&g(13).a(1)+g(1).a(14)&g(2).a
 723 (13)&g(3).a(12)&g(4).a(11)&g(5).a(10)&g(6).a(9)&g(7).a(8)&g(8).a(7)&g(9).a(6)&g(10).a(5)&g(11).a(4)&g(1
 724 2).a(3)&g(13).a(2)&g(14).a(1)+g(1).a(15)&g(2).a(14)&g(3).a(13)&g(4).a(12)&g(5).a(11)&g(6).a(10)&g(7).a(9)
 725 &g(8).a(8)&g(9).a(7)&g(10).a(6)&g(11).a(5)&g(12).a(4)&g(13).a(3)&g(14).a(2)&g(15).a(1)+g(1).a(16)&g(2).a
 726 (15)&g(3).a(14)&g(4).a(13)&g(5).a(12)&g(6).a(11)&g(7).a(10)&g(8).a(9)&g(9).a(8)&g(10).a(7)&g(11).a(6)&g
 727 (12).a(5)&g(13).a(4)&g(14).a(3)&g(15).a(2)&g(16).a(1)+g(1).a(17)&g(2).a(16)&g(3).a(15)&g(4).a(14)&g(5).a
 728 (13)&g(6).a(12)&g(7).a(11)&g(8).a(10)&g(9).a(9)&g(10).a(8)&g(11).a(7)&g(12).a(6)&g(13).a(5)&g(14).a(4)&g
 729 (15).a(3)&g(16).a(2)&g(17).a(1)+g(1).a(18)&g(2).a(17)&g(3).a(16)&g(4).a(15)&g(5).a(14)&g(6).a(13)&g(7).a
 730 (12)&g(8).a(11)&g(9).a(10)&g(10).a(9)&g(11).a(8)&g(12).a(7)&g(13).a(6)&g(14).a(5)&g(15).a(4)&g(16).a(3)&g
 731 (17).a(2)&g(18).a(1)+g(1).a(19)&g(2).a(18)&g(3).a(17)&g(4).a(16)&g(5).a(15)&g(6).a(14)&g(7).a(13)&g(8).
 732 a(12)&g(9).a(11)&g(10).a(10)&g(11).a(9)&g(12).a(8)&g(13).a(7)&g(14).a(6)&g(15).a(5)&g(16).a(4)&g(17).a(3)
 733 &g(18).a(2)&g(19).a(1)+g(1).a(20)&g(2).a(19)&g(3).a(18)&g(4).a(17)&g(5).a(16)&g(6).a(15)&g(7).a(14)&g
 734 (8).a(13)&g(9).a(12)&g(10).a(11)&g(11).a(10)&g(12).a(9)&g(13).a(8)&g(14).a(7)&g(15).a(6)&g(16).a(5)&g(1
 735 7).a(4)&g(18).a(3)&g(19).a(2)&g(20).a(1)+g(1).a(21)&g(2).a(20)&g(3).a(19)&g(4).a(18)&g(5).a(17)&g(6).a(1
 736 6)&g(7).a(15)&g(8).a(14)&g(9).a(13)&g(10).a(12)&g(11).a(11)&g(12).a(10)&g(13).a(9)&g(14).a(8)&g(15).a(7
 737)&g(16).a(6)&g(17).a(5)&g(18).a(4)&g(19).a(3)&g(20).a(2)&g(21).a(1)+g(1).a(22)&g(2).a(21)&g(3).a(20)&g
 738 (4).a(19)&g(5).a(18)&g(6).a(17)&g(7).a(16)&g(8).a(15)&g(9).a(14)&g(10).a(13)&g(11).a(12)&g(12).a(11)&g(1
 739 3).a(10)&g(14).a(9)&g(15).a(8)&g(16).a(7)&g(17).a(6)&g(18).a(5)&g(19).a(4)&g(20).a(3)&g(21).a(2)&g(22).a
 740 (1)+g(1).a(23)&g(2).a(22)&g(3).a(21)&g(4).a(20)&g(5).a(19)&g(6).a(18)&g(7).a(17)&g(8).a(16)&g(9).a(15)&
 741 g(10).a(14)&g(11).a(13)&g(12).a(12)&g(13).a(11)&g(14).a(10)&g(15).a(9)&g(16).a(8)&g(17).a(7)&g(18).a(6)
 742 &g(19).a(5)&g(20).a(4)&g(21).a(3)&g(22).a(2)+g(1).a(24)&g(2).a(23)&g(3).a(22)&g(4).a(21)&g(5).a(20)&g(6
 743).a(19)&g(7).a(18)&g(8).a(17)&g(9).a(16)&g(10).a(15)&g(11).a(14)&g(12).a(13)&g(13).a(12)&g(14).a(11)&g
 744 (15).a(10)&g(16).a(9)&g(17).a(8)&g(18).a(7)&g(19).a(6)&g(20).a(5)&g(21).a(4)&g(22).a(3)+g(1).a(25)&g(2).a
 745 (24)&g(3).a(23)&g(4).a(22)&g(5).a(21)&g(6).a(20)&g(7).a(19)&g(8).a(18)&g(9).a(17)&g(10).a(16)&g(11).a(1
 746 5)&g(12).a(14)&g(13).a(13)&g(14).a(12)&g(15).a(11)&g(16).a(10)&g(17).a(9)&g(18).a(8)&g(19).a(7)&g(20).
 747 a(6)&g(21).a(5)&g(22).a(4)+g(1).a(26)&g(2).a(25)&g(3).a(24)&g(4).a(23)&g(5).a(22)&g(6).a(21)&g(7).a(20)
 748 &g(8).a(19)&g(9).a(18)&g(10).a(17)&g(11).a(16)&g(12).a(15)&g(13).a(14)&g(14).a(13)&g(15).a(12)&g(16).a
 749 (11)&g(17).a(10)&g(18).a(9)&g(19).a(8)&g(20).a(7)&g(21).a(6)&g(22).a(5)+g(1).a(27)&g(2).a(26)&g(3).a(25)
 750 &g(4).a(24)&g(5).a(23)&g(6).a(22)&g(7).a(21)&g(8).a(20)&g(9).a(19)&g(10).a(18)&g(11).a(17)&g(12).a(16)
 751 &g(13).a(15)&g(14).a(14)&g(15).a(13)&g(16).a(12)&g(17).a(11)&g(18).a(10)&g(19).a(9)&g(20).a(8)&g(21).a
 752 (7)&g(22).a(6)+g(1).a(28)&g(2).a(27)&g(3).a(26)&g(4).a(25)&g(5).a(24)&g(6).a(23)&g(7).a(22)&g(8).a(21)&
 753 g(9).a(20)&g(10).a(19)&g(11).a(18)&g(12).a(17)&g(13).a(16)&g(14).a(15)&g(15).a(14)&g(16).a(13)&g(17).a(1
 754 2)&g(18).a(11)&g(19).a(10)&g(20).a(9)&g(21).a(8)&g(22).a(7)}
 755
 756

757

758 **Part 3.EXAMPLE 1: Matrix description of one of the simplest heterogeneity model**

759 The 5–state model with heterogeneity in survival only (Fig. 2 in the main text) is
760 described in Table S1. As all models considered in this study, it includes the possibility to
761 emigrate temporarily, and also to record information on the breeding status; otherwise it
762 reduces to Pledger et al.’s (2003) two–class finite mixture model $\{\Phi(h_2), p(t)\}$ (see also
763 Pradel 2009). If the two survival-classes and the two “status confirmation events” are merged
764 together this model becomes the temporary emigration model of Fujiwara and Caswell
765 (2002).

766

767 TABLE S1

768 Matrix description of a multievent model accounting for survival heterogeneity using two

769 classes, while allowing for temporary emigration and detailed observation structure. (a)

770 Survival probabilities (S–matrix): ϕ_1 and ϕ_2 are the two survival probabilities (high and low).771 (b) Transition probabilities (Ψ –matrix): ψ and ψ' are the emigration and return transition772 probabilities respectively. The product $\Phi=S \times \Psi$ constitutes the survival/transition matrix that773 is common to multistate and multievent models. (c) Event probabilities (B–matrix): p and r

774 are respectively detection and breeding status assignment probabilities. There is no

775 heterogeneity in detection in this example model.

776 States S1 and S2 code for the two survival classes in the study site, while AE1 and AE2 code

777 for them in the “Alive Elsewhere” site. There were three possible observations following

778 initial release (0: not seen, 1: seen and breeding status confirmed; 2: seen but breeding status

779 not confirmed), which correspond to the three columns of the B–matrix.

780

(a) S	S1	S2	AE1	AE2	Dead
S1	ϕ_1	0	0	0	$1-\phi_1$
S2	0	ϕ_2	0	0	$1-\phi_2$
AE1	0	0	ϕ_1	0	$1-\phi_1$
AE2	0	0	0	ϕ_2	$1-\phi_2$
dead	0	0	0	0	1

(b) Ψ	S1	S2	AE1	AE2	Dead
S1	$1-\psi$	0	ψ	0	0
S2	0	$1-\psi$	0	ψ	0
AE1	ψ'	0	$1-\psi'$	0	0
AE2	0	ψ'	0	$1-\psi'$	0
dead	0	0	0	0	1

(c) B=P×R	0	1	2
S1	$1-p$	$p \times r$	$p \times (1-r)$
S2	$1-p$	$p \times r$	$p \times (1-r)$
AE1	1	0	0
AE2	1	0	0
dead	1	0	0

781

782

783 **Part 4. EXAMPLE 2: Matrix description of the most complex model**

784 The multievent model $S_H;D_H;E_H$ has three kinds of heterogeneity. As all models
785 considered in this study, it includes the possibility to emigrate temporarily, and also to record
786 information on the breeding status; otherwise it reduces to Pledger et al.'s (2003) two-class
787 finite mixture model $\{\Phi(h_2), p(h_2+t)\}$ (see also Pradel, 2008). If the eight heterogeneity
788 classes and the two “status confirmation events” are merged together this model becomes the
789 temporary emigration model of Fujiwara and Caswell (2002).

790 There are 11 states in this model which is the most detailed we considered. A bird
791 breeding in LR can be in one of the 8 states denoted $S_iD_jE_k$, where i, j, k can be “low” or
792 “high”. An individual in state $S_{low}D_{high}E_{high}$ for example has low survival, high capture and
793 high emigration probabilities. Alternatively, the bird can be “Alive elsewhere” in states
794 AES_{low} and AES_{high} , or it can be dead, a state that is explicitly included in the model. The
795 recorded events are ‘0’ for “not seen”, ‘1’ for “seen as confirmed breeder”, and ‘2’ for “seen
796 as unconfirmed breeder”.

797 In addition to the matrix description we provide for this model more detail of the
798 practical implementation, in the form of the GEPAT/GEMACO instructions. These are the
799 interfaces of E-SURGE that are used to constrain multievent models (Choquet et al. 2009a).

800

801 TABLE S2

802 Survival probabilities (S–matrix) of multievent model accounting for survival,
 803 detection and emigration heterogeneity using two classes, while allowing for temporary
 804 emigration and detailed observation structure. φ_1 and φ_2 are the two survival probabilities.
 805 States denominations are described above.

		State at time t										
To\from		$S_{low,D_{low},E_{low}}$	$S_{low,D_{high},E_{low}}$	$S_{low,D_{low},E_{high}}$	$S_{low,D_{high},E_{high}}$	$S_{high,D_{low},E_{low}}$	$S_{high,D_{high},E_{low}}$	$S_{high,D_{low},E_{high}}$	$S_{high,D_{high},E_{high}}$	AES _{low}	AES _{high}	dead
State before dispersal	$S_{low,D_{low},E_{low}}$	φ_1	0	0	0	0	0	0	0	0	0	$1-\varphi_1$
	$S_{low,D_{high},E_{low}}$	0	φ_1	0	0	0	0	0	0	0	0	$1-\varphi_1$
	$S_{low,D_{low},E_{high}}$	0	0	φ_1	0	0	0	0	0	0	0	$1-\varphi_1$
	$S_{low,D_{high},E_{high}}$	0	0	0	φ_1	0	0	0	0	0	0	$1-\varphi_1$
	$S_{high,D_{low},E_{low}}$	0	0	0	0	φ_2	0	0	0	0	0	$1-\varphi_2$
	$S_{high,D_{high},E_{low}}$	0	0	0	0	0	φ_2	0	0	0	0	$1-\varphi_2$
	$S_{high,D_{low},E_{high}}$	0	0	0	0	0	0	φ_2	0	0	0	$1-\varphi_2$
	$S_{high,D_{high},E_{high}}$	0	0	0	0	0	0	0	φ_2	0	0	$1-\varphi_2$
	AES _{low}	0	0	0	0	0	0	0	0	φ_1	0	$1-\varphi_1$
	AES _{high}	0	0	0	0	0	0	0	0	0	φ_2	$1-\varphi_2$
	dead	0	0	0	0	0	0	0	0	0	0	1

806

807 The modelling of survival as in Table S2 is done through GEPAT and GEMACO as follow:

808 GEPAT instruction:

```

    φ - - - - - - - - - *
    - φ - - - - - - - - - *
    - - φ - - - - - - - - - *
    - - - φ - - - - - - - - - *
    - - - - φ - - - - - - - - - *
    - - - - - φ - - - - - - - - - *
    - - - - - - φ - - - - - - - - - *
    - - - - - - - φ - - - - - - - - - *
    - - - - - - - - φ - - - - - - - - - *
    - - - - - - - - - φ - - - - - - - - - *
    - - - - - - - - - - φ - - - - - - - - - *
    - - - - - - - - - - - φ - - - - - - - - - *
    - - - - - - - - - - - - φ - - - - - - - - - *
    - - - - - - - - - - - - - φ - - - - - - - - - *
    - - - - - - - - - - - - - - φ - - - - - - - - - *
    - - - - - - - - - - - - - - - φ - - - - - - - - - *
    
```

809

810 GEMACO instruction (for the TFC-effect):

811 **f(1:4;9,5;8;10).a(1,2:24)+a(14_24)*x(1)+a(20:24)+others**

812

813 A *.txt* file is to be entered as an external covariable and codes for the linear effect of age
 814 varying between 14 and 24 (11 age classes):

815 1
 816 11
 817 0.090909091 0.181818182 0.272727273 0.363636364 0.454545455 0.545454545
 818 0.636363636 0.727272727 0.818181818 0.909090909 1
 819

820 TABLE S3

821 Dispersal probabilities (Ψ -matrix) of multievent model accounting for survival,
 822 detection and emigration heterogeneity using two classes, while allowing for temporary
 823 emigration and detailed observation structure. The product $\Phi=S \times \Psi$ constitutes the
 824 survival/transition matrix that is common to multistate and multievent models. ψ_1 and ψ_2 are
 825 the two probabilities of transition from the study site to states “Alive Elsewhere”,
 826 corresponding to high and low emigration rates. After a temporary emigration, gulls could
 827 completely change the location of their nest, thus change of detectability class. Similarly, after
 828 an emigration event, gulls could be more site-faithful (if emigration aims at sparing energy for
 829 subsequent breeding) or less site faithful (if the first emigration event is the symptom of a
 830 decrease in health status). Therefore our model did not constraint the way individual behaved
 831 after a temporary emigration: four different probabilities to return to the study site were
 832 implemented. The probabilities of coming back to the study colony are four: $\psi_3, \psi_4, \psi_5, \psi_6$
 833 respectively the transitions, from the corresponding “Alive Elsewhere” states, to states
 834 $S_{low}D_{low}E_{low}$ and $S_{high}D_{low}E_{low}$, $S_{low}D_{high}E_{low}$ and $S_{high}D_{high}E_{low}$, $S_{low}D_{low}E_{high}$ and $S_{high}D_{low}E_{high}$,
 835 $S_{low}D_{high}E_{high}$ and $S_{high}D_{high}E_{high}$. $\sum \psi$ stands for $\psi_3 + \psi_4 + \psi_5 + \psi_6$. States denominations are
 836 described above.

837

		State before dispersal											
		To\from	$S_{low}D_{low}E_{low}$	$S_{low}D_{high}E_{low}$	$S_{low}D_{low}E_{high}$	$S_{low}D_{high}E_{high}$	$S_{high}D_{low}E_{low}$	$S_{high}D_{high}E_{low}$	$S_{high}D_{low}E_{high}$	$S_{high}D_{high}E_{high}$	AES _{low}	AES _{high}	dead
Stat	↑	$S_{low}D_{low}E_{low}$	$1-\psi_1$	0	0	0	0	0	0	0	ψ_1	0	0
	↓	$S_{low}D_{high}E_{low}$	0	$1-\psi_1$	0	0	0	0	0	0	ψ_1	0	0

$S_{low}D_{low}E_{high}$	0	0	$1-\psi_2$	0	0	0	0	0	ψ_2	0	0
$S_{low}D_{high}E_{high}$	0	0	0	$1-\psi_2$	0	0	0	0	ψ_2	0	0
$S_{high}D_{low}E_{low}$	0	0	0	0	$1-\psi_1$	0	0	0	0	ψ_1	0
$S_{high}D_{high}E_{low}$	0	0	0	0	0	$1-\psi_1$	0	0	0	ψ_1	0
$S_{high}D_{low}E_{high}$	0	0	0	0	0	0	$1-\psi_2$	0	0	ψ_2	0
$S_{high}D_{high}E_{high}$	0	0	0	0	0	0	0	$1-\psi_2$	0	ψ_2	0
AES_{low}	ψ_3	ψ_4	ψ_5	ψ_6	0	0	0	0	$1-\sum\psi$	0	0
AES_{high}	0	0	0	0	ψ_3	ψ_4	ψ_5	ψ_6	0	$1-\sum\psi$	0
dead	0	0	0	0	0	0	0	0	0	0	1

838

839 The modelling of dispersal as in Table S3 is done through GEPAT and GEMACO as follow:

840 GEPAT instruction:

```

* 0 0 0 0 0 0 0 0 0 0 0
0 * 0 0 0 0 0 0 0 0 0 0
0 0 * 0 0 0 0 0 0 0 0 0
0 0 0 * 0 0 0 0 0 0 0 0
0 0 0 0 * 0 0 0 0 0 0 0
0 0 0 0 0 * 0 0 0 0 0 0
0 0 0 0 0 0 * 0 0 0 0 0
0 0 0 0 0 0 0 * 0 0 0 0 0
ψ ψ ψ ψ 0 0 0 0 * 0 0 0
0 0 0 0 ψ ψ ψ ψ 0 * 0 0
0 0 0 0 0 0 0 0 0 0 *

```

841

842 GEMACO instruction:

843 **f(1 3 5 7,2 4 6 8)+f(9 10).to**

844

845 TABLE S4

846 Detection probabilities (P-matrix) and breeding status confirmation probabilities (R-
847 matrix) of multievent model accounting for survival, detection and emigration heterogeneity
848 using two classes, while allowing for temporary emigration and detailed observation
849 structure. Product $P \times R$ constitutes the matrix of event probabilities (B-matrix), which
850 represents the probabilities of recording the events conditional on the occupied state. The
851 “observation status” [“not seen”, “seen 1” (seen with a high detection probability), “seen 2”

852 (seen with a low detection probability)] corresponds to transitory states. Then there were three
 853 possible events (0: not seen, 1: seen and breeding status confirmed; 2: seen but breeding status
 854 not confirmed), which correspond to the three columns of R. p_{low} and p_{high} are the two
 855 detection probabilities; r_{low} and r_{high} are the corresponding probabilities of confirming
 856 breeding status. States denominations are described above.

(a) P	Observation status			(b) R	Event recorded				
	not seen	Seen1	Seen2		0	1	2		
State at time $t+1$	$S_{low}D_{low}E_{low}$	$1-p_{low}$	p_{low}	0	Observation status	not seen	1	0	0
	$S_{low}D_{high}E_{low}$	$1-p_{high}$	0	p_{high}		Seen1	0	r_{low}	$1-r_{low}$
	$S_{low}D_{low}E_{high}$	$1-p_{low}$	p_{low}	0		Seen2	0	r_{high}	$1-r_{high}$
	$S_{low}D_{high}E_{high}$	$1-p_{high}$	0	p_{high}					
	$S_{high}D_{low}E_{low}$	$1-p_{low}$	p_{low}	0					
	$S_{high}D_{high}E_{low}$	$1-p_{high}$	0	p_{high}					
	$S_{high}D_{low}E_{high}$	$1-p_{low}$	p_{low}	0					
	$S_{high}D_{high}E_{high}$	$1-p_{high}$	0	p_{high}					
	AES _{low}	1	0	0					
	AES _{high}	1	0	0					
	dead	1	0	0					

857

858 The modelling of event recording as in Table S4 is done through GEPAT and GEMACO as
 859 follow:

860 GEPAT instruction for detection:

```

*  β  -
*  -  β
*  β  -
*  -  β
*  β  -
*  -  β
*  β  -
*  -  β
*  -  -
*  -  -
*  -  -
    
```

861 GEPAT instruction for breeding status confirmation:

```

*  -  -
-  β  *
-  β  *
    
```

862

863 GEMACO instruction for detection (modelling time-dependence):

864 **a(1)+a(2:29).[to+t]**

865 GEMACO instruction for breeding status confirmation:

866 **from**

867

868

869 **References quoted in this Appendix**

870 Aebischer, N.J. and Coulson, J.C. 1990. Survival of the Kittiwake in relation to sex, year,
871 breeding experience and position in the colony. – J. Anim. Ecol 59: 1063–71.

872 Choquet, R. 2008. Automatic generation of multistate capture–recapture models. – Can. J.
873 Stat. 36: 43–57.

874 Choquet, R. et al. 2009a. Program E–SURGE: a software application for fitting Multievent
875 models. – In: Thomson, D. L. Cooch, E. G. and Conroy, M. J. (eds.), Modelling

876 Frederiksen, M. et al. 2004. Estimating true age–dependance in survival when only adults can
877 be observed: an example with Black–legged Kittiwakes. – Anim. Biodivers. Conserv. 27:
878 541–48.

879 Fujiwara, M. and Caswell, H. 2002. A general approach to temporary emigration in mark–
880 recapture analysis. – Ecology 83: 3266–75.

881 Goodman, L.A. 1961. Statistical methods for the mover–stayer model. – J. Am. Stat. Assoc.
882 56: 841–68.

883 Jones, O.R. et al. 2008. Senescence rates are determined by ranking on the fast–slow life–
884 history continuum. – Ecol. Lett. 11: 664–73.

885 Loison, A. et al. 1999. Age–specific survival in five populations of ungulates: evidence of
886 senescence. – Ecology 80: 2539–54.

887 Nichols, J.D. et al. 1997. Test for senescent decline in annual survival probabilities of
888 common pochard, *Aythya ferina*. – Ecology 78: 1009–18.

- 889 Pledger, S. et al. 2003. Open capture–recapture models with heterogeneity: I. Cormack–Jolly–
890 Seber model. – *Biometrics* 59: 786–94.
- 891 Pradel, R. 2009. The stakes of capture–recapture models with state uncertainty. – In:
892 Thomson, D. L. Cooch, E. G. and Conroy, M. J. eds. *Modelling Demographic Processes in*
893 *Marked Populations*, pp. 781–795.
- 894 Prévot-Julliard, A.C. et al. 1998a. Re-evaluation of adult survival of Black-headed Gulls
895 *Larus ridibundus* in presence of recapture heterogeneity. – *Auk* 115: 85–95.
- 896 Prévot-Julliard, A.C. et al. 1998b. Evidence for birth–site tenacity in breeding Common
897 Black-headed Gulls, *Larus ridibundus*. – *Can. J. Zool.* 76: 2295–98.
- 898 Pugsek, et al. 1995. Mark–resighting analysis of a California gull population. – *J. Appl. Stat.*
899 22: 625–39.
- 900 R Development Core Team 2008. *R: A language and environment for statistical computing.* –
901 R Foundation for Statistical Computing, Vienna, Austria.
- 902

903 **Appendix S2: Simulation study**

904

905 Because the performances of multievent CMR models have never been formally tested in the
906 presence of several sources of heterogeneity, we carried out three simulation studies.

- 907 - The first simulated situation was the absence of senescence and the presence of three
908 kinds of heterogeneity (survival, detection, emigration). It aimed at evaluating the
909 performance of our information-theoretic approach concerning type II errors (i.e. the
910 probability to falsely detect the presence of senescence).
- 911 - The second simulated case was the presence of senescence and of the three kinds of
912 heterogeneity, but without the detailed information on detectability, i.e. there was only
913 two events possible (“seen” / “not seen”). It aimed at illustrating potential issues related to
914 the mixing of detection and emigration heterogeneity.
- 915 - The third simulated case was the presence of senescence, three sources of heterogeneity,
916 and detailed data structure (i.e. the frequency of the observed events depended on the
917 detection probability; see “data simulation” below). There were three events possible
918 (“seen 1”, “seen 2”, “not seen”).

919 **Data simulation**

920 The simulation studies were based on 30 simulated datasets (except for case #2 where
921 we used 60 datasets). The time needed to perform each of these exercises exceeded 10 days
922 each with an Intel Pentium 4HT, 2.6GHz (3.25 x 800) processor with 512 Mb of system
923 memory.

924 Datasets were simulated as follows. At each of 20 capture sessions, 50 newly marked
925 individuals were released in each of the eight classes of heterogeneity, forming an overall
926 sample size of 8000 individuals. Between each session, individuals were allowed to move to

927 or from a site where detection was impossible. The eight classes had different values of
 928 demographic parameters: probability of survival (at age 1 before senescence) was 0.8 or 0.5;
 929 of detection 0.5 or 0.1; and of changing site 0.4 or 0.1.

930 When present in the simulated case, the detailed observation structure was simulated as
 931 follows. In the high-detectability class, at each session 80% of the detected individuals
 932 produced the event “seen 1” and 20% the event “seen 2”. In the low-detectability class the
 933 frequency of “seen1” was 20% and the frequency of “seen2” was 80%.

934 When present in the simulated case, senescence was simulated as a linear decrease in
 935 survival with age starting at age 1 and with a slope of -0.033 and -0.021 year^{-1} in the high and
 936 low survival classes respectively (Fig. S1), i.e. survival probability was zero at age 25. We
 937 deliberately chose to simulate a linear effect (and not a logit-linear effect), so that, as is
 938 supposedly the case in real datasets, the logit-linear relationship between age and survival that
 939 was used in the CMR model was only approximating the underlying process.

940 **Data analysis**

941 Ten models were fitted on each simulated datasets: $\{S_{0+\bar{a}};D_0;E_0\}$, $\{S_{0+a};D_0;E_0\}$,
 942 $\{S_{0+\bar{a}};D_H;E_H\}$, $\{S_{0+a};D_H;E_H\}$, $\{S_{H+\bar{a}};D_0;E_H\}$, $\{S_{H+a};D_0;E_H\}$, $\{S_{H+\bar{a}};D_H;E_0\}$, $\{S_{H+a};D_H;E_0\}$,
 943 $\{S_{H+\bar{a}};D_H;E_H\}$ and $\{S_{H+a};D_H;E_H\}$. See main text for the meaning of the symbols. For each
 944 dataset-model, we computed the AIC, AIC-weights, and parameter estimates, and for each
 945 dataset we computed the importance value of the four considered effects (age-effect on
 946 survival, heterogeneity in survival, detection and emigration probabilities; see method in main
 947 text). To spare computer time we did not estimate the rank of the models and therefore did not
 948 formally detect identifiability problems if they occurred (Rouan et al. 2009; Choquet et al.
 949 2009a), nor did we compute standard errors on the estimates.

950 **RESULTS**951 **CASE #1: ABSENCE OF SENESCENCE, DETAILED DATA**

952 - Model selection

953 The three kinds of heterogeneity were unambiguously selected. Their importance
954 values were 1. Among the 30 datasets, and using the 2-point AIC threshold, model
955 $\{S_{H+a};D_H;E_H\}$ was selected 18 times and model $\{S_{H+\bar{a}};D_H;E_H\}$ (i.e. the model used to simulate
956 the data) 16 times. Thus, despite its absence, the age-effect on survival was selected in more
957 than half the cases.

958 - Slope of the age-effect where it should have been absent

959 We computed the average estimated slope of the age-effect (on a logit scale) over the
960 30 models $\{S_{H+a};D_H;E_H\}$. In the low-survival class, the slope estimate was always very
961 positive (average \pm SD over the 30 datasets: $+0.26 \pm 0.05$), i.e. not compatible with
962 senescence. Why the models included this strong increase from a low survival at age 1 to a
963 high survival at old age in the low-survival class is not known, but supposedly stems from the
964 quick decrease with age in sample size of low-survival individuals (see discussion).

965 In the high-survival class, the estimated slope was on average positive and closer to
966 zero (average \pm SD over the 30 datasets: $+0.042 \pm 0.059$). Type II errors for senescence (i.e.
967 negative slopes) occurred in 9 out of 30 cases. Over these 9 cases, the slope was still on
968 average close to zero (-0.027 ± 0.022 on a logit scale, compared to the value obtained in the
969 gull application: -0.16). In short, even if the age effect was selected in more than half of the
970 datasets, the estimated slope in the high-survival class was small and most of the time positive
971 and thus did not provide strong evidence for senescence.

972

973 CASE #2: PRESENCE OF SENESCENCE, NON-DETAILED DATA**974 - Model selection**

975 Among the 60 datasets, and using the 2-point AIC threshold, model $\{S_{H+a};D_0;E_H\}$ was
976 selected 50 times, model $\{S_{H+a};D_H;E_H\}$ (i.e. the model used to simulate the data) 20 times
977 only, model $\{S_{H+a};D_H;E_0\}$ 11 times, and model $\{S_{0+a};D_H;E_H\}$ once. Thus, the age-effect and
978 the survival heterogeneity were generally selected. The two other kinds of heterogeneity
979 proved harder to detect. Importance values of the considered effects were high (Table S5),
980 with the only exception of heterogeneity in detection probability. This issue is discussed
981 under “potential identifiability issues”.

982 When neglecting heterogeneity in survival, in 45 out of 60 cases the age-effect was
983 discarded (by AIC); this confirmed the strong bias due to neglecting survival heterogeneity.
984 When neglecting heterogeneity in detection or emigration the age-effect was always selected
985 except for 3 cases. Therefore, under the simulated scenario, these sources of heterogeneity
986 had a reduced impact on the detection of survival senescence.

987 - Slope of senescence

988 In the high-survival class the estimated slope of -0.11 ± 0.02 on a logit scale (average \pm
989 SD over the 60 datasets) translated into a -0.023 year^{-1} (expected value: -0.033 year^{-1}).

990 The slope was of slightly lower magnitude in the model without emigration
991 heterogeneity than in the model with it (it was the case in 49 cases out of 60; average slope -
992 0.09 ± 0.02 instead of -0.11 ± 0.02). Therefore, the impact of emigration heterogeneity on the
993 estimation of senescence received some support.

994

995 CASE #3: PRESENCE OF SENESCENCE, DETAILED DATA (3 EVENTS)**996 - Model selection**

997 Among the 30 datasets, and using the 2-point AIC threshold, model $\{S_{H+a};D_H;E_H\}$ (i.e.
998 the model used to simulate the data) was selected 26 times, model $\{S_{H+a};D_H;E_0\}$ 11 times,
999 model $\{S_{0+a};D_H;E_H\}$ twice, and model $\{S_{H+a};D_H;E_H\}$ three times. The age-effect, the survival
1000 heterogeneity, and the detection heterogeneity (contrary to Case #2) were unambiguously
1001 selected when relying on importance values (Table S5). Emigration heterogeneity was
1002 slightly less frequently detected. The three preferred models often ended up very close to one
1003 another.

1004 - Slope of senescence

1005 In the high-survival class the estimated slope of -0.16 ± 0.02 (average \pm SD over the
1006 30 datasets) on a logit scale translated into $-0.0329 \text{ year}^{-1}$ on a linear scale (expected value: -
1007 0.033 year^{-1}).

1008 The slope was not of lower magnitude in the model without emigration heterogeneity
1009 than in the model with it (all cases), and discarding emigration heterogeneity did not prevent
1010 to detect senescence (all cases). Therefore, the impact of emigration heterogeneity on
1011 senescence was not supported in Case #3.

1012 **DISCUSSION**

1013 - Type-II errors

1014 In Case #1 the selection of an age-effect when there was none in the data was quite
1015 frequent. However, in most cases the estimated slope of the age-effect in the high-survival
1016 class was small and positive; if we had computed confidence intervals (not performed due to
1017 computer time constraints), we believe these slopes would moreover have appeared non-
1018 significantly different from zero. The risk of type-II errors for senescence *per se* therefore
1019 appeared low. Further investigation concerning the surprising find that the low-survival class
1020 was impacted by a strong increase in estimated survival probability is required; for low-

1021 survival individuals the interval between the sessions might have been too short to study age
1022 effects on survival.

1023 - Potential identifiability issues

1024 In Case #2, emigration and detection heterogeneity proved difficult to detect.

1025 Moreover when the importance value of emigration heterogeneity was low, it tended to be
1026 negatively correlated to the importance value of detection heterogeneity (Fig. S2).

1027 Identifiability problems might thus have occurred in some of the simulated cases, between
1028 heterogeneity in detection and in emigration. In other words, the model might have

1029 difficulties separating individuals not seen because emigrated from those not seen because not
1030 detected, hence the large SD on the corresponding importance values in Table S5.

1031 Case #3 however illustrated how the use of detailed observation data (three events
1032 instead of two, the frequency of each event depending on detectability) made the separation
1033 between detection and emigration processes possible. We thereby suggest that in the study we
1034 present in the main text, our use of the information about breeding status confirmation did
1035 overcome the issue of parameter identifiability raised by Case #2. Furthermore the tool
1036 implemented in E-SURGE did not detect any problem of parameter redundancy in the gull
1037 study (see main text).

1038

1039

1040

1041

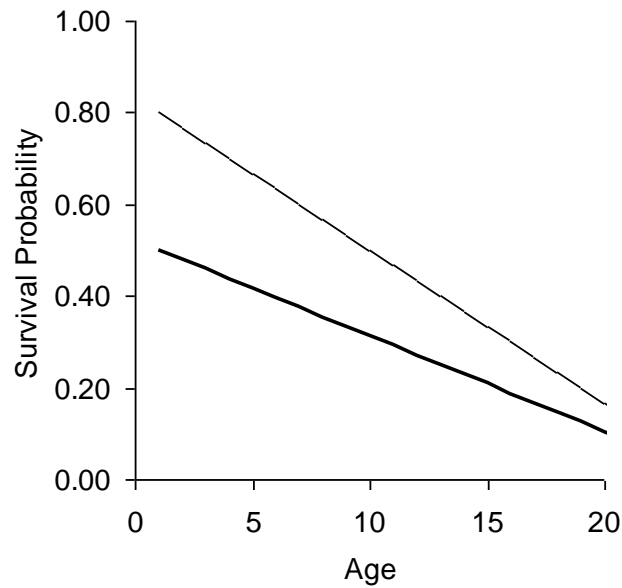
1042 Table S5: Importance value (IV; see method in main text) of the four considered effects in
 1043 three simulation studies. See “data simulation” for the description of the considered cases.
 1044 Mean and SD were computed on 30 (Case #1 and 3) or 60 (Case #2) simulated datasets.

Effect considered	CASE #1		CASE #2		CASE #3	
	Mean	SD	Mean	SD	Mean	SD
age-effect on survival	0.536	0.410	0.987	0.081	0.910	0.137
heterogeneity in survival	1.000	0.000	0.987	0.083	0.932	0.088
heterogeneity in emigration	1.000	0.000	0.874	0.224	0.759	0.228
heterogeneity in detection	1.000	0.000	0.335	0.307	1.000	0.000

1045

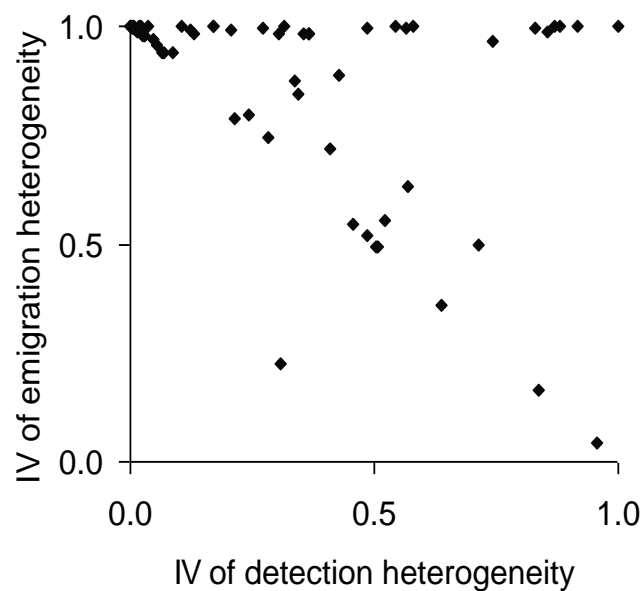
1046

1047 Figure S1: Simulated variation in survival rate with age in the low- (starting at 0.5) and high-
1048 (starting at 0.8) survival classes.



1049

1050 Figure S2: Importance values (IV; see methods in main text) of detection and emigration
1051 heterogeneity computed from the 60 datasets in Case #2 where the detailed information on
1052 detectability was not used.



1053

1054