The Auk
Ornithological Advances

AmericanOrnithology.org

Volume 137, 2020, pp. 1–17
DOI: 10.1093/auk/ukaa022

RESEARCH ARTICLE

Breeding performance of Common Terns (*Sterna hirundo*) does not decline among older age classes

Ian C. T. Nisbet,[1,#,*] David Iles,[2,a,#] Andrew Kaneb,[2,b] Carolyn S. Mostello,[3] and Stéphanie Jenouvrier[2,4]

[1] I. C. T. Nisbet & Company, North Falmouth, Massachusetts, USA
[2] Biology Department, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts, USA
[3] Massachusetts Division of Fisheries and Wildlife, Westborough, Massachusetts, USA
[4] Centre d'Etudes Biologiques de Chizé, Centre National de la Recherche Scientifique, Villiers en Bois, France
[a] Current address: Canadian Wildlife Service, Environment and Climate Change Canada, Ottawa, Ontario, Canada
[b] Current address: Princeton University, Princeton, New Jersey, USA
[#] These authors contributed equally to the paper.
*Corresponding author: icnisbet@verizon.net

Submission Date: October 8, 2019; Editorial Acceptance Date: April 7, 2020; Published May 16, 2020

ABSTRACT

Declines in reproductive performance among older age classes have been reported in many bird and mammal species, and are commonly presented as demonstrating reproductive senescence. However, no declines in performance could be demonstrated in studies of several bird species. We measured reproductive performance in Common Terns (*Sterna hirundo*) of known age (2–28 yr) during a 19-yr period at a site in Buzzards Bay, Massachusetts, USA. We measured 6 components of reproductive performance and used generalized additive mixed models (GAMMs) in a Bayesian framework to analyze dependence of each measure on parental age, while controlling for variations among years and indices of individual quality. Four measures of performance improved (earlier laying date, higher values of clutch size, fledging success, and productivity) with age, most rapidly between ages 2 and 10 yr; egg mass and hatching success varied only slightly with age. No measure of performance showed reversals among the older age classes; fledging success and productivity continued to increase through at least age 22 yr. These findings are consistent with results from an earlier study of the same species. Continued increase in reproductive performance through the oldest age classes is not incompatible with "reproductive senescence" (decline in physiological or other functions required for successful reproduction) if either reproductive effort or efficiency continue to increase. Studies within our population have yielded no evidence for age-related increase in reproductive effort, but 3 studies have suggested that older Common Terns can raise chicks more successfully than younger birds without increasing reproductive effort, probably by more efficient foraging and chick provisioning. Our findings suggest that Common Terns offset reproductive senescence by continuing to improve efficiency through at least age 22 yr. Age-related changes in efficiency should be investigated in other species with similar life-history traits.

Keywords: age, breeding, Common Tern, laying date, productivity, reproductive effort, reproductive senescence, *Sterna hirundo*

LAY SUMMARY

- We studied Common Terns (*Sterna hirundo*) at a breeding colony in Massachusetts from 1970 until 2003: we banded chicks when they hatched so that we could determine their age when they returned to breed at the same site in subsequent years.
- The terns' breeding success continued to improve with age throughout their lives and was still increasing among birds that were 22 years old.
- Common Terns continually improve their skills at finding fish and provisioning their chicks, even after 20 years of practice.

Absence de sénescence reproductive chez les Sternes Pierregarin (*Sterna hirundo*)

RÉSUMÉ

Chez de nombreuses espèces d'oiseaux et de mammifères, une baisse des performances reproductives a été mise en évidence chez les individus plus âgées, généralement présentée comme démontrant une sénescence de la reproduction. Cependant, aucun vieillissement des capacités de reproduction n'a pu être démontré chez plusieurs espèces d'oiseaux. Cette étude mesure la performance de reproduction chez des Sternes Pierregarin (*Sterna hirundo*) d'âge connu (2–28 ans) pendant une période de 19 ans à Buzzards Bay, Massachusetts, USA. Six composantes de la performance de reproduction ont été mesurées et des modèles statistiques mixtes additifs généralisés (GAMM) dans un cadre Bayésien ont été développés pour analyser la dépendance de chaque mesure à l'âge des parents, tout en contrôlant les variations aléatoires entre les années et individus (indices de qualité individuelle). Quatre mesures des performances se sont améliorées (date de ponte avancée, valeurs plus élevées: du nombre de poussins, du succès à l'envol et de la productivité) avec l'âge, et ceci plus rapidement entre 2 et 10 ans. Cependant, la masse des œufs et le succès d'éclosion varient peu avec l'âge. Aucune mesure de la performance n'a montré vieillissement des capacités de reproduction; le succès à l'envol et la productivité ont continué d'augmenter jusqu'à l'âge de 22 ans. Ces résultats sont cohérents avec ceux d'une étude antérieure sur la même espèce. L'augmentation des performances de reproduction dans les classes d'âge les plus vielles n'est pas incompatible avec la «sénescence de la reproduction» (déclin des fonctions physiologiques ou autres nécessaires à une reproduction réussie) si l'effort de reproduction ou l'efficacité à élever des poussins continue à augmenter. Des études au sein de cette population de Sternes Pierregarins à Buzzards Bay n'ont fourni aucune évidence robuste d'une augmentation de l'effort de reproduction liée à l'âge, mais trois études ont suggéré que les individus plus âgées peuvent élever les poussins plus efficacement que les individus plus jeunes sans augmenter l'effort de reproduction, probablement par une recherche de nourriture et un approvisionnement des poussins plus efficaces. Nos résultats suggèrent que les Sternes Pierregarins compensent la sénescence reproductrice en continuant à améliorer l'efficacité de reproduction jusqu'à l'âge de au moins 22 ans. Les changements d'efficacité liés à l'âge devraient être étudiés chez d'autres espèces ayant des caractéristiques biologiques similaires.

Mots-clés: âge, date de ponte, effort de reproduction, productivité, reproduction, sénescence reproductive, *Sterna hirundo*, Sterne Pierregarin

INTRODUCTION

Late-life declines in reproductive performance have been reported in many animal species (Nussey et al. 2013). However, in several studies of birds no declines could be detected despite intensive study to advanced ages, potentially indicating a lack of reproductive senescence (Altwegg et al. 2007, Berman et al. 2009, Blas et al. 2009). If negligible reproductive senescence could be demonstrated in wild animals, this would be a phenomenon of great biological interest (Ottinger 2007, Monaghan et al. 2008, Nussey et al. 2008), so cases in which a species is found to show no evidence of late-life reproductive declines should be checked carefully for possible confounding factors.

Several processes can obscure late-life declines in reproductive performance when they do in fact exist. These include individual heterogeneity (Vedder and Bouwhuis 2018), selective disappearance of low-quality individuals from study populations (Vaupel and Yashin 1985), and temporal variation in environmental conditions, such as resource availability (Nisbet 2001). Individual heterogeneity and selective disappearance can be controlled for by using mixed models to estimate changes within individuals (van de Pol and Verhulst 2006, Nussey et al. 2008). Changes in the environment over the course of a study can similarly be controlled for by including fixed or random effects that account for these changes, or by standardizing measures of performance relative to the population mean for the same year (Reid et al. 2003).

Environmental conditions that influence late-life reproductive performance are more difficult to control for and could make it difficult to detect late-life declines. Indeed, theory and empirical evidence suggest that patterns of senescence can vary in response to environmental regimes that determine intrinsic mortality risks (Austad 1993, McNamara et al. 2009). For example, Austad (1993) reported delayed actuarial senescence in populations of Virginia opossums (*Didelphis virginiana*) that experienced reduced exposure to predators. Similarly, Campbell et al. (2017) found that the onset of reproductive senescence in wild Eurasian beavers (*Castor fiber*) occurred much earlier in low-quality environments. Thus, long-term studies of age-related changes in reproductive performance under different environmental regimes may be necessary to understand whether the presence or lack of reproductive senescence is a general property of any particular species.

In a detailed longitudinal study of Common Terns (*Sterna hirundo*) at a breeding colony in Germany, Zhang et al. (2015b) found no significant change in several aspects of reproductive performance (laying date, egg volume, and clutch size) between ages 10 and 19 yr, whereas productivity (number of chicks raised to fledging per breeding attempt) appeared to increase continuously until at least age 19 yr. A second analysis of the same dataset (Zhang et al. 2015a) showed no increase in productivity after age 14 yr; the 2 sets of results were mutually consistent and together showed no evidence of late-life declines in productivity. Here, we evaluate evidence for late-life reproductive declines in a second population of Common Terns breeding in Buzzards Bay, Massachusetts, USA. The Buzzards Bay population appears to have lower adult survival than the German population (mean 0.85–0.89 yr^{-1} across ages 4–14 yr: Breton et al. 2014, vs. 0.90–0.92 yr^{-1} across all ages: Szostek and Becker 2012), potentially leading to stronger selection for high reproductive performance in midlife and more rapid senescence (McNamara et al. 2009). Both populations nest at high densities (0.3–1.8 nests m^{-2}; Szostek et al. 2014, Nisbet et al. 2020a); both have exhibited density-dependent reductions in productivity as total numbers have increased, but the Buzzards Bay colony appears to have been resource-limited throughout the period of our study whereas the German colony was resource-limited only during the second half of the study period (Tims et al. 2004, Szostek et al. 2014, Nisbet et al. 2020a). Accordingly, we investigated whether our population may have exhibited stronger signs of senescence than the German population.

Nisbet et al. (2002) previously reported that the oldest individuals (ages 18–23 yr) in the Buzzards Bay colony had higher performance in several components of reproduction than any matched group of younger birds. However, that study was cross-sectional in design and was conducted in only 1 yr. Here, we report a comprehensive analysis of more than 1,400 individual birds studied over a 19-yr period. We use a Bayesian mixed-modeling approach to examine the relationship between age and overall productivity (number of young fledged per nest) as well as lower-level functional traits (clutch size, hatching success, breeding phenology, and egg mass) while controlling for known confounding factors. Our investigation provides a complement to the German studies (Zhang et al. 2015a, 2015b), generating broader insights into the presence (or lack) of age-related reproductive declines in this long-lived seabird species.

METHODS

Study Site and Data Collection

Fieldwork was conducted at Bird Island, Marion, Massachusetts, USA (41.669°N, 70.717°W); data collected from 1986 to 2004 are used here. At Bird Island during the study period, numbers of nesting Common Terns (hereafter, terns) increased from ~1,100 pairs in 1986 to ~1,900 pairs in 1989 and thereafter remained more or less constant at ~1,800 pairs until 2004 (see Appendix A, Appendix Table 1). Mean breeding productivity at Bird Island was ~1.0 fledgling pair^{-1}, declining slightly during the study period (Appendix Table 2), and was thought to have been limited by density-dependent factors, probably interference competition at foraging sites (Tims et al. 2004). Hence, our study was conducted in a period when breeding success was limited by extrinsic factors (food availability) and intraspecific competition, so that variations among individuals would have reflected differences in foraging skill and competitive ability.

Our analysis was limited to nesting terns that had been banded as chicks, and were thus of known age (Breton et al. 2014). Known-aged birds were located by trapping adults on nests. Sampling was rigorously randomized only in 1991–1992 (Nisbet 1996), 1995–1997 (Nisbet and Cam 2002), and 1999 (Nisbet et al. 2002), but in all other years birds were trapped throughout the nesting area and samples are thought to have been representative of all birds that nested on the island, except that younger birds (ages 2–4 yr) were probably under-sampled because few birds were trapped after the main peak of nesting in each year. We weighed and measured all trapped birds, allowing us to estimate and control for body condition in subsequent analyses of breeding performance (see Statistical Analyses). Approximately 86% of individuals could be sexed with ≥95% confidence using multiple characters (Appendix B). Head length (from tip of bill to back of skull) was used as a measure of body size and hence as part of an index of individual quality (Nisbet et al. 2007; Appendix C).

We searched the nesting area daily during the peak breeding period in each year and most nests were marked in sequence as soon as they were found. For each nest with a known-aged parent, we recorded sequential measures of breeding performance: (1) laying date, (2) egg mass, (3) clutch size, (4) hatching success, (5) fledging success, and (6) productivity (each of these variables is defined in Appendix A). All nests included in the study are assumed to be the first nestings for each pair in each year: failures and renestings were infrequent throughout the study period, and most or all renestings were excluded by excluding late outliers (Appendix D).

Data on breeding performance were available for ages 2–28 yr, but we limited our analysis to ages 2–23 yr because sample sizes were small for ages ≥24 yr (Appendix Table 5). After exclusion of cases for reasons specified in Appendices A–C, our dataset included 2,132 records of 1,401 unique individuals for laying date, 1,423 records of 1,031 individuals for egg mass, 1,980 records of 1,325 individuals for clutch size, 714 records of 615 individuals for hatching success, 302 records of 274 individuals for fledging success, and 345 records of 310 individuals for productivity. Appendix A contains additional details describing temporal dynamics of the Bird Island colony, field methods and definitions of variables.

Statistical Analyses

Our primary objective was to examine the relationship between age and each of the 6 measures of breeding performance (Appendix A), and to evaluate evidence for declines in 1 or more of these measures at the oldest ages. Following Zhang et al. (2015a, 2015b), we used generalized additive mixed models (GAMMs) to investigate these relationships, allowing for flexible and asymmetric effects of age on each measure of breeding performance. GAMMs fit curvilinear functions between prespecified knots within the data and use smoothing parameters to ensure that functional relationships are continuous across knot locations.

In each model, we specified 6 knots equally spaced across the range of observed ages, thereby balancing model flexibility and over-fitting, although preliminary analyses indicated that our conclusions were robust to other choices of knot number (e.g., 5–11 knots). We used cubic regression splines as a basis for all GAMMs.

We assumed a Gaussian error structure for models describing laying date and egg mass, and used an identity link function to relate covariates to the response variables. Clutch sizes were strongly under-dispersed relative to a Poisson distribution: most nests had either 2 or 3 eggs. We therefore categorized each clutch as either "large" (≥3 eggs) or "small" (≤2 eggs), and used a Bernoulli error structure to model the probability of having a "large" clutch, with a logit link function. We assumed binomial errors (and logit link function) for hatching success and fledging success, and Poisson error (and log link function) for productivity. In all models, we included a Gaussian random effect for "study year" to account for interannual differences in reproductive performance that could potentially obscure age-related effects (e.g., due to temporal variation in resource availability). We did not include study year as a linear covariate because there was little evidence for trends in outcome variables (Appendix Table 2): laying dates advanced during the study period, but most of this advance was attributable to the increase in mean ages (Appendix A).

In analyses examining the effect of age on each measure of breeding performance, we controlled for potentially confounding effects of individual quality in 3 ways. First, we included a random effect for "bird identity" that accounts for consistent differences in individual quality not explained by other covariates. Second, we used the methods of Peig and Green (2009) to calculate a body condition index (BCI) for each bird in each year: BCI indexes body mass corrected for structural size and is included as an explanatory covariate in all models (see Appendix C). BCI was calculated separately for males and females. We tested for dependence of BCI on age and laying date as described in Appendix C. Third, we extracted residuals from the selected model for laying date and included these as an explanatory covariate (resid-laydate) in each of the models for the subsequent measures of breeding performance. Resid-laydate describes whether a bird initiated its nest early or late relative to other birds of the same age and body condition in the same year, and is expected to reflect differences in individual quality in addition to effects of age and BCI.

We used Bayesian methods to fit each model for breeding performance using JAGS 4.3.0 interfaced with the R programming language version 3.4.4 (R Project Core Team 2018). We used the *jagam* function from the *mgcv* package (Wood 2019) to fit GAMMs in JAGS. We stored every fifth sample from each of 3 MCMC chains until 5,000 posterior samples were accumulated, after discarding a burn-in period of 10,000 iterations. We assessed model convergence visually and using the Gelman-Rubin convergence statistic; all models unambiguously converged. Although Bayesian methods are more computationally expensive than maximum likelihood approaches, we used them in our analyses because they enabled us to calculate probabilities that alternative functions would fit the data, specifically whether trends observed among young and middle-aged birds might be reversed among the oldest age classes (see Results).

Within the reduced sample of nests for which the ages of both parents were known (n = 448), we tested for effects of mate age on laying date and productivity as described in Appendix E. Data on mate retention from one year to the next are also presented in Appendix E. We tested for selective survival of high-performing birds as described in Appendix F.

RESULTS

Covariate Effects on Body Condition and Breeding Performance

Body Condition Index (BCI) was negatively related to laying date, but this relationship was manifested only among birds with laying dates later than 35 (June 4) and was not statistically significant (Appendix Figure 3A). BCI was not related to age (Appendix Figure 3B).

Laying date had a strongly nonlinear and negative relationship with age (Figure 1A). On average, 2-yr-old birds initiated their clutches 18.0 days later than 10-yr-old birds (95% credible interval [CI]: 17.0 to 19.1 days), and 21.5 days later than 23-yr-old birds (CI: 18.6 to 24.3 days). The index of body condition, BCI, had a weakly negative effect on laying date (mean effect = –0.14 days g^{-1}, 95% CI: –0.36 to 0.07 days g^{-1}); thus, a 1-g increase in BCI resulted in a 0.14-day advance in laying date. Random effect variances associated with "study year" and "bird identity" were 18.0 and 8.4 $days^2$, respectively, indicating large interannual variation in laying dates (e.g., due to temporal variation in environmental conditions) and moderate variation among individuals (e.g., due to differences in individual quality in addition to those explained by age and body condition); residual variance was 14.6 $days^2$.

Egg mass was not correlated with age (Figure 1B). BCI had a weakly positive effect on egg mass (mean effect = 0.08, 95% CI: 0.00 to 0.16), while resid-laydate had a negative effect (mean effect = –0.026 g day^{-1}, 95% CI: –0.044 to –0.009 g day^{-1}). Thus, for each 1-g increase in BCI, mean egg mass increased by 0.08 g, and for each 1-day delay in clutch initiation, mean egg mass decreased by 0.026 g. Egg mass varied among individuals (variance of bird identity random effect = 1.02 g^2), but little among years (variance of study year random effect = 0.06 g^2): residual variance was 1.37 g^2.

The probability that a bird laid a "large" clutch (≥3 eggs) increased strongly with age, from 0.21 at age 2 yr (CI: 0.12 to 0.32) to 0.64 at age 23 yr (CI: 0.42 to 0.82). The relationship was nonlinear, with a rapid increase from age 2 to 12 yr, leveling off after age ~15 yr (Figure 1C). BCI had a weakly negative effect on clutch size (mean effect on logit scale = −0.13 g^{-1}, 95% CI: −0.25 to −0.01 g^{-1}), after controlling for age. Resid-laydate also had a negative effect (mean effect on logit scale = −0.07 day^{-1}, 95% CRI: −0.10 to −0.05 day^{-1}). Thus, for each 1-day increase in resid-laydate, the probability of laying a large clutch decreased by 0.07 on the logit scale. Variances (on the logit scale) associated with random effects for study year and bird identity were 0.80 and 0.68, respectively.

Hatching success was high and increased slightly with age (Figure 1D), from a mean of 0.92 for 2-yr-old birds (CI: 0.84 to 0.97) to a mean of 0.96 for 23-yr-old birds (CI: 0.91 to 0.99). The effect of BCI on hatching success was negligible and credible intervals broadly overlapped zero (mean effect = 0.09 g^{-1}, CI: −0.11 to 0.30 g^{-1}). The effect of resid-laydate was also negligible (mean effect = −0.026 day^{-1}, 95% CI: −0.070 to 0.015 day^{-1}). Variances (on the logit scale) associated with random effects for study year and bird identity were 1.20 and 1.63, respectively.

In contrast to hatching success, fledging success increased strongly with age (Figure 1E), from a mean of 0.30 for 2-yr-old birds (CI: 0.21 to 0.40) to a mean of 0.60 for 22-yr-old birds (CI: 0.42 to 0.78). The effect of BCI on fledging success was negligible and credible intervals broadly overlapped zero (mean effect = −0.05 day^{-1}, 95% CI: −0.21 to 0.11 day^{-1}). Fledging success declined with residual laying date (mean effect = −0.07 g^{-1}, 95% CI: −0.23 to 0.10 g^{-1}). Thus, each 1-day delay in clutch initiation decreased fledging success by 0.07 on the logit scale; a 10-day delay in clutch initiation for 23-yr-old birds would reduce mean fledging success from 0.42 to 0.26. Variances (on the logit scale) associated with random effects for study year and bird identity were 0.080 and 0.032, respectively.

Productivity (number of young fledged per nest), which integrates all preceding measures of breeding performance, also increased with age (Figure 1F). Productivity increased almost linearly from 0.57 at age 2 yr (CI: 0.39 to 0.80) to 1.63 at age 22 yr (CI: 0.91 to 2.70). The effect of BCI on productivity was negligible (mean effect = −0.05 g^{-1}, CI: −0.16 to 0.06 g^{-1}). Productivity was negatively related to resid-laydate (mean effect = −0.045 day^{-1}, CI: −0.067 to −0.023 day^{-1}; Figure 2). Thus, each 1-day delay in clutch initiation decreased total productivity by 0.045 on the log scale; a 10-day delay in clutch initiation for 10-yr-old birds would reduce mean productivity from 0.92 to 0.59. Variances (on the logit scale) associated with random effects for study year and bird identity were 0.033 and 0.006, respectively.

Additional Effects of Mate Age

In 448 cases, the ages of both mates were known. Ages of mates were highly correlated, and mate retention was very high (Appendix E). Within the reduced sample of nests for which the ages of both mates were known, addition of residual mate age to the GAMM for laying date improved the fit (explained deviance increasing from 69.8% to 71.9%; mean effect = −0.41 days yr^{-1}, CI: −0.33 to −0.50 days yr^{-1}; n = 321). Addition of residual mate age to the GAMM for productivity did not improve the fit (explained deviance increasing from 15.8% to 16.1%; mean effect = −0.014 yr^{-1}, CI: −0.044 to 0.027 yr^{-1}; n = 95; Appendix E).

Evidence for Late-life Declines in Reproductive Performance

All of the 6 measures of reproductive performance we investigated changed monotonically with age up to 22 yr (23 yr for laying date, clutch size, and hatching success), and none displayed evidence of reversals among the oldest age classes. Bayesian analysis allowed us to directly determine the probability that age-related improvements in breeding performance were reversed for the oldest age classes. For laying date, the probability that 22-yr-old birds initiated clutches ≥1 day later on average than 21-yr-old birds was 0.002, providing strong evidence against late-life regression to later breeding. Similarly, the probabilities of a 1% (or larger) decrease in egg mass, clutch size, hatching success, fledging success, and total productivity between 21 and 22 yr of age were <0.001, 0.31, 0.005, 0.06, and 0.11, respectively. Thus, evidence of reversed trends among the oldest age classes was absent or extremely weak for all measures of reproductive performance except clutch size.

DISCUSSION

We found no evidence for late-life declines in reproductive performance in Common Terns at our study site: laying date, egg mass, clutch size, and hatching success were essentially constant (after controlling for other intrinsic variables) from age 12 yr onward (Figure 1A–D), and fledging success and productivity increased throughout the lifespan to at least age 22 yr (Figure 1E, F). In these respects, our results are closely parallel to those of Zhang et al. (2015b) in the same species: in particular, Figure 1F in our paper and Figure 3(g1) in Zhang et al. (2015b) show the same pattern of monotonic increase in productivity throughout the lifespan. Both studies used the same statistical methodology (GAMMs) and controlled for several confounding factors, but there were important differences

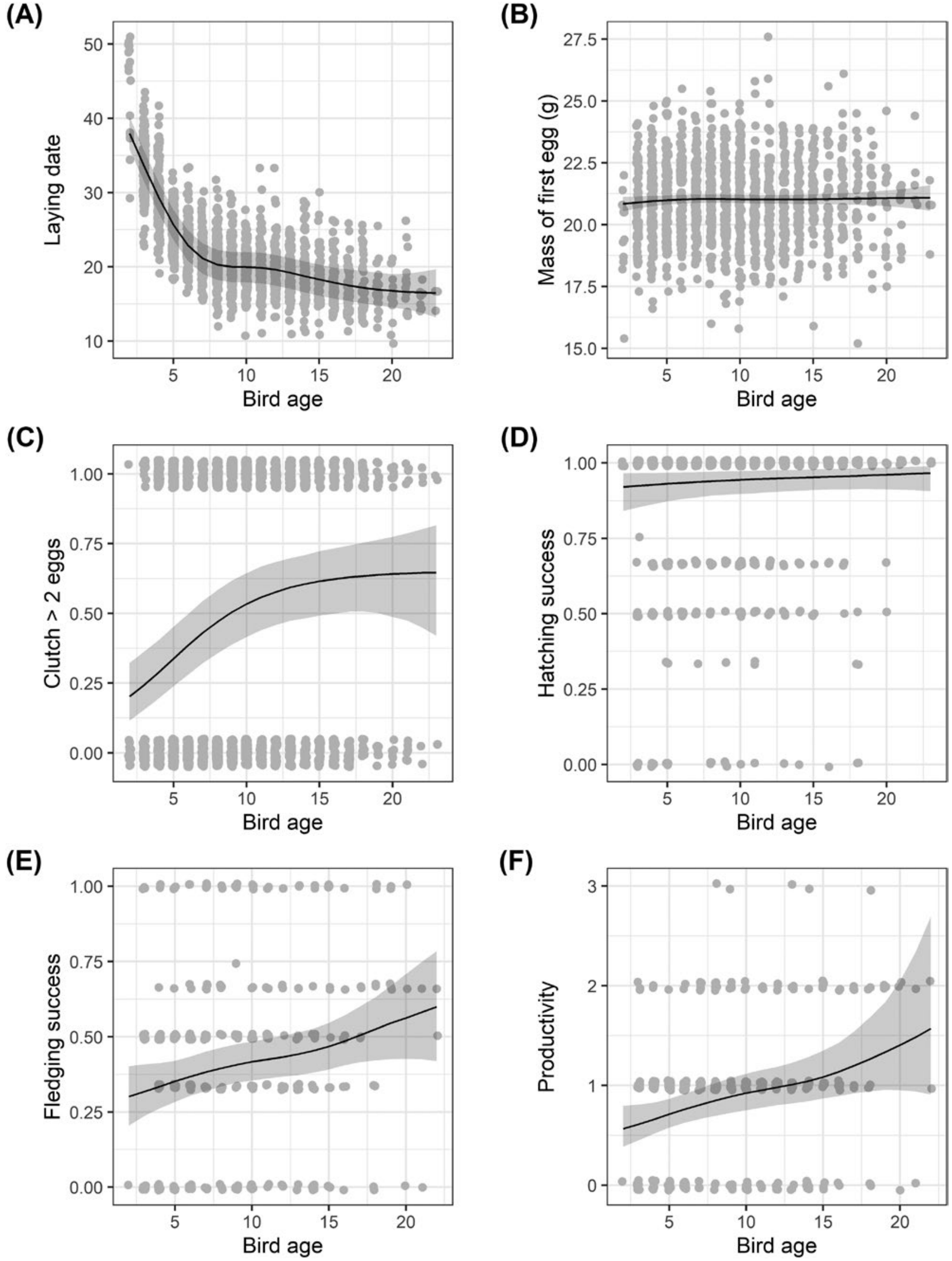


FIGURE 1. Dependence on age of (**A**) laying date, (**B**) egg mass, (**C**) clutch size, (**D**) hatching success, (**E**) fledging success, and (**F**) productivity in a Common Tern population in Massachusetts, USA. Relationships are derived using GAMMs controlling for bird identity, year, BCI, and resid-laydate (for details, see Methods). Shaded areas mark 95% CI. Gray dots plot individual cases; for discrete variables (**C–F**), these are "jittered" to show densities.

between the datasets and the variables included in the analyses (see below). Other differences were that the mortality rate appears to have been higher in our population and that our population was more resource-limited (see Introduction), which might have led to earlier or more rapid senescence. However, both populations exhibited the same pattern of increasing reproductive performance into advanced ages.

Methodological Comparison of the Two Studies

In the German study, Zhang et al. (2015b) used data for fewer individual birds than our study (473 vs. 1,401, respectively), but they had better longitudinal data than ours (mean 4.57 cases per bird vs. 1.52) and, consequently, they were able to consider several variables that we could not measure, including breeding probability, lifespan, and recruitment age (Zhang et al. 2015b). Conversely, we included in our analysis several other variables (BCI, residlaydate, and mate age) that Zhang et al. did not consider.

Both studies controlled for several important factors that can alter age-related patterns of reproductive performance within individuals and/or obscure these patterns at the population level. Below, we discuss the potentially confounding effects of each, how we controlled for them, and the resulting insights they provide into the life history of Common Terns.

Confounding Factors

Individual heterogeneity and selective disappearance. Differences in age-related reproduction among individuals can obscure patterns at the population level and lead to erroneous conclusions about the presence (or absence) of senescence. This may occur when "low-quality" individuals have lower survival rates than "high-quality" individuals, and consequently are less frequent in the older age classes. In these cases, population means of breeding performance and other demographic parameters may increase with age without any changes within individuals (Vaupel and Yashin 1985, Bouwhuis and Vedder 2017). In principle, accounting for the random effects of individual identity should control for these effects. Within our study, the residual variance among individuals was large in laying date, egg mass, and hatching success, but relatively low in fledging success and productivity. We found no evidence that birds with high productivity had higher survival than those with low productivity, although the sample size for this comparison was small (Appendix E). This is consistent with findings from the German population, for which the effects of individual heterogeneity and selective disappearance were reported to be minor (Zhang et al. 2015b, Vedder and Bouwhuis 2018).

Changes in the environment. Changes in the environment during the course of a study may also influence patterns of age-related changes in reproductive performance. In most studies of known-aged birds (including both studies of Common Terns), the proportion of older birds in the sampled population is low in the early years of the study and increases in later years (e.g., Appendix Table 2). In these circumstances, if environmental factors (e.g., food availability) improve (or deteriorate) during the course of a study, population means of breeding performance and other parameters may increase (or decrease) with age without any changes in capability within individuals (Nisbet 2001). In both studies of Common Terns, there was considerable interannual variation in mean laying date, clutch size, and productivity (Appendix Table 2; González-Solís et al. 2004), consistent with strong year-to-year differences in food availability. In Germany, productivity and other demographic parameters declined markedly during the period of study, attributed to a decrease in food availability (Szostek and Becker 2012). However, the use of random effects to account for interannual variation in both studies should have controlled for these temporal changes in the environment.

Mate age. In our study, as in earlier studies of Common Terns (Bridge and Nisbet 2004, Rebke et al. 2017), ages of mates were highly correlated (Appendix E). Hence, older birds had older mates and this would have enhanced the relationship between breeding performance and age: it could not have reduced the strength of that relationship or reversed it. We found that residual mate age was a minor but significant predictor of laying date after controlling for bird age; there was no significant effect of residual mate age on productivity, although the sample size was small for that comparison (Appendix E). Zhang et al. (2015b) did not consider mate age, but in a separate study within the German population, Rebke et al. (2017) reported large effects of mate age on productivity in addition to effects of age of the focal bird.

Pair-bond duration. In some studies of other species, breeding performance parameters were correlated with the duration of pair bonds, accounting for part or all of the apparent relationships with age (Nisbet and Dann 2009, and references therein). In our study, mate retention was very high (Appendix E), so it is likely that older birds had longer pair bonds than younger birds. However, we had no information on pair-bond duration for our birds. In the German population, Rebke et al. (2017) found no effect of pair-bond duration on productivity.

Breeding probability. If the probability of breeding decreases (or increases) among older age classes, this could offset increases (or decreases) in mean productivity. However, Zhang et al. (2015b) found that the mean probability of breeding in Common Terns varied only between 0.90 and 1.00 among age classes, with a minimum around age 10 yr, leveling off at about 0.95 after age 10 yr. We could not measure breeding frequencies of individual birds, but an earlier study within our dataset indicated that the mean probability of breeding was similar to that in the German study (0.93–1.00; Nisbet and Cam 2002).

Life span and recruitment age. We could not measure life span or recruitment age for individual birds. Zhang et al. (2015b: Table 1) found that both variables strongly influenced laying date: long-lived birds laid earlier, and birds that recruited to the breeding population at ages ≥4 yr laid later. However, they found no effect of life span on productivity after controlling for other variables, and only small effects of recruitment age on productivity (limited to the few birds that recruited to the breeding population at ages ≥5 yr).

Terminal investment. We could not identify ages at last reproduction for individual birds and could not investigate terminal effects (increased or decreased reproductive success in the last year of life). Zhang et al. (2015b) found that terns laid earlier in the last year of life, but there was no effect on productivity.

Individual quality. We measured 2 indices of individual "quality" in addition to age: BCI (body mass controlled for structural size) and resid-laydate (laying date controlled for effects of age and BCI). Both indices were significantly autocorrelated (Appendix F), so they apparently reflected persistent characteristics of individual birds as well as variations in these characteristics among years. BCI had a small effect on laying date (birds with high BCI laid earlier, after controlling for other factors), but no significant effect on productivity (see Results). Resid-laydate had a large effect on fledging success and productivity (birds that laid earlier than predicted for their age and BCI raised more young: Figure 2). Zhang et al. (2015b) did not explicitly consider indices of individual quality other than recruitment age (see above), but in both studies the use of mixed models with individual identity as a random factor would have controlled for most components of individual quality.

Summary. Many of the confounding factors listed above appear to have influenced laying dates of individual birds. Because several factors that we found to be important were not considered in the German studies, and vice versa, it is difficult to draw definitive conclusions about the dependence of laying date on age without full control of all identified confounding factors. However, it is unlikely that any of the confounding factors could have obscured a tendency for laying dates to become later in late life, if it had existed.

There is little evidence that any of the confounding factors listed above could have caused the observed late-life increase in productivity if it did not exist. We found a strong effect of resid-laydate (an index of individual quality, not considered by Zhang et al. 2015b) on fledging success and productivity (Figure 2). However, the late-life increase in productivity that we found controlled for the effect of this factor (Figure 1F).

Juvenile Survival and Transgenerational Effects

Our study was limited to following the breeding performance of terns through fledging of their offspring. Parental fitness can also decline among older age classes through effects that occur after fledging, including reduced juvenile survival or reduced breeding performance of the offspring after recruitment to the breeding population.

In a small sample within our population (prior to the period of our study), Nisbet (1996) found no effect of parental age on survival from fledging to recruitment. In the German population, Bouwhuis et al. (2015: figure S4) reported that survival from fledging to recruitment declined with the age of the male parent, but not with the age of the female parent. The study by Bouwhuis et al. (2015) was limited to "local recruitment" (recruitment to the breeding population at the natal site) and could have been confounded by differential emigration: emigration is known to be substantial in our population (Tims et al. 2004) and was taken into account in the study by Nisbet (1996).

Bouwhuis et al. (2015: Figure 2) also reported that the lifetime reproductive success of the recruits varied strongly with the age of their male and female parents, peaking at parental ages of 10–11 yr and declining by about 50% by parental age 16 yr. They concluded that the fitness of the parents declined among the oldest age classes because of this transgenerational effect, even though there had been no decline in reproductive performance through fledging of the offspring. Further study of transgenerational effects is needed.

Reproductive Senescence

Late-life declines in reproductive performance have frequently been reported as demonstrating "reproductive senescence" (e.g., Nussey et al. 2013), but the scientific basis for this claim is tenuous. We have not found a rigorous definition of reproductive senescence anywhere in the

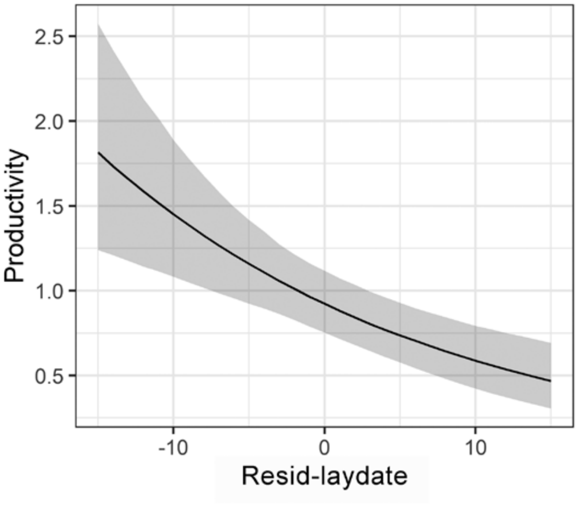


FIGURE 2. Dependence of productivity of Common Terns on resid-laydate, controlling for age and for other factors listed in the caption to Figure 1. Data are shown for birds aged 10 yr. The shaded area marks 95% CI.

scientific literature. Senescence is usually defined as "a progressive decline in vitality and functions" (Finch 1994) or "a decline in fitness with age caused by physiological deterioration" (Jones et al. 2008). Accordingly, reproductive senescence has been described in laboratory species in terms of decline in physiological parameters such as sperm function in males, ovarian function in females, and hormonal changes in both sexes (Ottinger 2007). Incommensurably, field studies of wild birds (including ours and those of Zhang et al. 2015b and Bouwhuis and Vedder 2017 in Common Terns) have measured late-life declines (or lack of declines) in various aspects of reproductive performance. Although studies in both populations of Common Terns have reported physiological parameters in known-aged birds (Nisbet et al. 1999, Apanius and Nisbet 2003, Heidinger et al. 2006, 2010; Riechert et al. 2012) none has related longitudinal measurements of breeding performance to physiological parameters in the same birds. Declines in survival among older age classes (actuarial senescence) have been reported in both populations (Breton et al. 2014, Zhang et al. 2015a).

Causes of Late-Life Increases in Breeding Performance

Given that age-related increases in breeding performance have been observed in 2 independent studies of Common Terns, after controlling for several confounding factors, we address the potential biological causes and implications of this pattern. Specifically, we examine 2 mechanisms that could enable reproductive performance to continue to improve with age even if the birds were experiencing physiological or functional senescence.

Increased reproductive effort. Life-history theory predicts that declining reproductive value (RV) should lead to increasing reproductive effort (Stearns 1992), which might account for some or all of the continuing increase in productivity with age at ages ≥7 yr reported here and by Zhang et al. (2015b). Indeed, Zhang et al. (2015a) reported that RV progressively declined among the older age classes in the German population of Common Terns. However, reproductive effort is difficult to measure independently of measures of performance such as chick provisioning rate and productivity. Clutch size and egg mass are sometimes used as indices of reproductive effort, but neither of these indices showed late-life changes in our study (Figure 1B, C) or in that of Zhang et al. (2015b). Other studies attempting to detect and measure age-related changes in reproductive effort within our population have shown no consistent patterns. Field metabolic rates (proxy for energy expended in parental care) and circulating levels of IgG (proxy for resources expended on self-maintenance) did not vary with age (Galbraith et al. 1999, Apanius and Nisbet 2003, respectively). In a handicapping study, older males lost mass at lower rates during chick-raising than younger males, opposite to the prediction from life-history theory (Nisbet et al. 2004). Endocrine responses (increase in corticosterone, decrease in prolactin) to a standardized stressor declined with age through age 29 yr, but there were no age-related changes in baseline levels of corticosterone or prolactin (Heidinger et al. 2006, 2010). In the German population, Riechert et al. (2012) reported age-related changes in baseline levels of prolactin and corticosterone, but the changes were very small and were only significant for prolactin in males after the 11th year of breeding. Sample sizes were small in several of these studies, but collectively they have yielded little evidence for age-related increase in reproductive effort that might account for the observed increase in reproductive performance.

Increased efficiency. In a study within our population, rates of feeding young declined with parental age from ages 12 to 20 yr without changes in mass provisioning rates to the young or field metabolic rates of the parents, implying that older parents brought larger prey items than younger parents and spent less time foraging (Galbraith et al. 1999). This suggests an age-related increase in efficiency without increase in effort. In 2 other studies, chicks of older parents grew more rapidly than those of younger parents, after controlling for effects of laying date (Nisbet et al. 2002, Tims et al. 2004), implying that older parents were better at provisioning their chicks. The effects were most pronounced in the asymptotic masses of *B*-chicks (second hatched in each brood): most *B*-chicks of parents aged 5–16 yr fledged in poor condition, whereas most *B*-chicks of parents aged 18–23 yr fledged in good condition (Nisbet et al. 2002).

Together, these results indicate that older Common Terns are able to raise chicks more successfully than younger birds without increasing reproductive effort, probably by more efficient foraging and chick provisioning. Common Terns are ecological generalists, with large variability in foraging locations, foraging methods, and prey selection, adjusted to local conditions over wide spatial and temporal scales (Becker and Ludwigs 2004, Goyert 2015, Nisbet et al. 2020a). Raising chicks to fledging in good condition requires skill and adaptability: achieving proficiency is known to require many years of experience (Galbraith et al. 1999) and it is plausible that efficiency might continue to increase even after 15–20 yr of experience. Adult Common Terns move frequently among the 3 breeding sites within Buzzards Bay, and rates of dispersal increase with age (Breton et al. 2014): we suggest that selecting the best site at which to nest at the start of each breeding season may be one of the skills that improves with age but requires many years of experience of several alternative sites.

Factors Associated with Maintenance of High Reproductive Success Among the Oldest Age Classes

Most bird and mammal species that have been studied sufficiently have shown late-life declines in reproductive performance: Nussey et al. (2013) listed 51 bird species in which these declines have been reported, although the evidence for declines was weak in some cases and the degree of control for confounding factors was variable. However, in several species of birds no declines could be detected despite intensive study to advanced ages (Jones et al. 2008). These species include Barn Owl (*Tyto alba*: females only; Altwegg et al. 2007), Snow Petrel (*Pagodroma nivea*; Berman et al. 2009), and Black Kite (*Milvus migrans*; Blas et al. 2009). However, none of the studies of these species controlled for as many confounding factors as in the 2 studies of Common Terns.

Although these 4 species are dissimilar in many ways, they share several common characteristics: they are all long-lived, with biparental care and high rates of mate retention; they are all predatory, and they all live in complex environments. We suggest that this combination of characteristics is important in enabling birds of these (and similar) species to gain experience through many years of breeding, and hence to maintain or improve breeding performance late in life, even while they may be experiencing senescent declines in survival and other functions.

Conclusions

Two independent sets of studies of Common Terns have yielded no evidence of late-life declines in reproductive performance: both sets of studies showed increases in productivity throughout the lifespan. Both sets of studies were controlled for several potential confounding factors. Studies within our population showed no evidence for age-related increases in reproductive effort. Thus, the most likely explanation for the lack of late-life declines in breeding performance in Common Terns is increased efficiency in foraging and raising young (attributable to progressive improvement in skill and experience), apparently increasing through at least age 23 yr. Future studies should focus on measuring reproductive effort and reproductive efficiency in this and other species with similar characteristics (i.e. high longevity, biparental care, high mate retention, and complex or highly variable environments).

ACKNOWLEDGMENTS

We thank many colleagues and field assistants who helped with data collection, the Town of Marion for permission to work at Bird Island, P. Szczys for sexing, A. Breton for help with the database, M. A. Ottinger, O. Vedder, and 2 anonymous reviewers for helpful discussion.

Funding statement: Field work in 1999–2004 was supported in part by the New Bedford Harbor Trustee Council through an Interagency Service Agreement with Massachusetts Division of Fisheries and Wildlife.

Ethics statement: All field work was conducted under appropriate state and federal permits, and prospectively conformed to the *Guidelines for the Use of Wild Birds in Research*, subsequently developed by the Ornithological Council. IACUC approval was not required at the time of field work.

Author contributions: I.C.T.N. conceived and designed the study; I.C.T.N. and C.S.M. performed field work; D.I., A.K., and S.J. conducted statistical analyses; I.C.T.N., D.I., and S.J. drafted the manuscript; all authors contributed to revision and approved the paper as submitted. I.C.T.N. and D.I. contributed equally to this work.

Data availability: Analyses reported in this article can be reproduced using the data provided by Nisbet et al. (2020b).

LITERATURE CITED

Altwegg, R., M. Schaub, and A. Roulin (2007). Age-specific fitness components and their temporal variation in the Barn Owl. The American Naturalist 169:47–61.

Apanius, V., and I. C. Nisbet (2003). Serum immunoglobulin G levels in very old Common Terns *Sterna hirundo*. Experimental Gerontology 38:761–764.

Arnold, J. M., J. J. Hatch, and I. C. T. Nisbet (2004). Seasonal declines in reproductive success of the Common Tern: Timing or parental quality? Journal of Avian Biology 35:33–45.

Austad, S. N. (1993). Retarded senescence in an insular population of Virginia opossums (*Didelphis virginiana*). Journal of Zoology 229:695–708.

Becker, P. H., and J.-D. Ludwigs (2004). Common Tern *Sterna hirundo*. BWP Update 6:91–137.

Berman, M., J. M. Gaillard, and H. Weimerskirch (2009). Contrasted patterns of age-specific reproduction in long-lived seabirds. Proceedings of the Royal Society of London B: Biological Sciences 276:375–382.

Blas, J., F. Sergio, and F. Hiraldo (2009). Age-related improvement in reproductive performance in a long-lived raptor: A cross-sectional and longitudinal study. Ecography 32:647–657.

Bouwhuis, S., and O. Vedder (2017). Avian escape artists? Patterns, processes and costs of senescence in wild birds. In The Evolution of Senescence in the Tree of Life (R. P. Shefferson, O. R. Jones, and R. Salguero-Gómez, Editors). Cambridge University Press, Cambridge, UK. pp. 156–174.

Bouwhuis, S., O. Vedder, and P. H. Becker (2015). Sex-specific pathways of parental age effects on offspring lifetime reproductive success in a long-lived seabird. Evolution 69:1760–1771.

Breton, A. R., I. C. T. Nisbet, C. S. Mostello, and J. J. Hatch (2014). Age-dependent survival and breeding dispersal within a metapopulation of Common Terns *Sterna hirundo*. Ibis 156:534–547.

Bridge, E. S., and I. C. T. Nisbet (2004). Wing-molt and assortative mating in Common Terns: A test of the molt-signaling hypothesis. The Condor 106:336–343.

Campbell, R. D., F. Rosell, C. Newman, and D. W. Macdonald (2017). Age-related changes in somatic condition and reproduction in the Eurasian beaver: Resource history influences onset of reproductive senescence. PLOS One 12:e0187484.

Finch, C. E. (1994). Longevity, Senescence, and the Genome. University of Chicago Press, Chicago, IL, USA.

Friar, M. S. (2004). Adoption behavior in Common Terns *Sterna hirundo*. Ph.D. thesis, The University at Albany, Albany, NY, USA.

Galbraith, H., J. J. Hatch, I. C. T. Nisbet, and T. W. Kunz (1999). Age-specific reproductive efficiency among Common Terns:

Measurement of energy expenditure using doubly-labelled water. Journal of Avian Biology 30:85–96.

González-Solís, J., P. H. Becker, L. Jover, and X. Ruiz (2004). Individual changes underlie age-specific pattern of laying date and egg-size in female Common Terns *Sterna hirundo*. Journal of Ornithology 145:129–136.

Goyert, H. F. (2015). Foraging specificity and prey utilization: Evaluating social and memory-based strategies in seabirds. Behaviour 152:861–895.

Heidinger, B. J., O. Chastel, I. C. T. Nisbet, and E. D. Ketterson (2010). Mellowing with age: Older parents are less responsive to a stressor in a long-lived seabird. Functional Ecology 24:1037–1044.

Heidinger, B., I. C. T. Nisbet, and E. D. Ketterson (2006). Older parents are less responsive to stress: A mechanism for increased reproductive performance with age? Proceedings of the Royal Society B, Biological Sciences 273:2227–2231.

Jones, O. R., J. M. Gaillard, S. Tuljapurkar, J. S. Alho, K. B. Armitage, P. H. Becker, P. Bize, J. Brommer, A. Charmantier, M. Charpentier, et al. (2008). Senescence rates are determined by ranking on the fast–slow life-history continuum. Ecology Letters 11:664–673.

McNamara, J. M., A. I. Houston, Z. Barta, A. Scheuerlein, and L. Fromhage (2009). Deterioration, death and the evolution of reproductive restraint in late life. Proceedings of the Royal Society B, Biological Sciences 276:4061–4066.

Monaghan, P., A. Charmantier, D. H. Nussey, and R. E. Ricklefs (2008). The evolutionary ecology of senescence. Functional Ecology 22:371–378.

Nisbet, I. C. T. (1996). Post-fledging survival of Common Terns in relation to brood order, hatching date and parental age. Colonial Waterbirds 19:253–255.

Nisbet, I. C. T. (2001). Detecting and measuring senescence in wild birds: Experience with long-lived seabirds. Experimental Gerontology 36:833–843.

Nisbet, I. C. T., and E. Cam (2002). Test for age-specificity in survival of the Common Tern. Journal of Applied Statistics 29:65–83.

Nisbet, I. C. T., and P. Dann (2009). Reproductive performance of Little Penguins *Eudyptula minor* in relation to year, age, breeding date, pair-bond duration and individual quality. Journal of Avian Biology 40:296–308.

Nisbet, I. C. T., and W. H. Drury (1972). Measuring breeding success in Common and Roseate terns. Bird-Banding 43:97–106.

Nisbet, I. C. T., V. Apanius, and M. S. Friar (2002). Breeding performance of very old Common Terns. Journal of Field Ornithology 73:117–124.

Nisbet, I. C. T., J. M. Arnold, H. Galbraith, and J. J. Hatch (2004). Responses of known-aged Common Terns to experimental shortening of the wings. Waterbirds 27:13–20.

Nisbet, I. C. T., J. M. Arnold, S. A. Oswald, P. Pyle, and M. A. Patten (2020a). Common Tern (*Sterna hirundo*), version 1.0. *In* Birds of the World (S.M. Billerman, *Editor*). Cornell Lab of Ornithology, Ithaca, NY, USA. https://doi.org/10.2173/bow.comter.01

Nisbet, I. C. T., E. S. Bridge, P. Szczys, and B. J. Heidinger (2007). Sexual dimorphism, female–female pairs and test for assortative mating in Common Terns. Waterbirds 30:169–179.

Nisbet, I. C. T., C. E. Finch, N. Thompson, E. Russek-Cohen, J. A. Proudman, and M. A. Ottinger (1999). Endocrine patterns during aging in the Common Tern (*Sterna hirundo*). General and Comparative Endocrinology 114:279–286.

Nisbet, I. C. T., D. Iles, A. Kaneb, C. S. Mostello, and S. Jenouvrier (2020b). Data from: Breeding performance of Common Terns (*Sterna hirundo*) does not decline among older age-classes. The Auk: Ornithological Advances 137:1–17. https://doi.org/10.5061/dryad.qrfj6q5c6

Nisbet, I. C. T., J. M. Winchell, and A. E. Heise (1984). Influence of age on the breeding biology of Common Terns. Colonial Waterbirds 7:117–126.

Nussey, D. H., T. Coulson, M. Festa-Bianchet, and J. M. Gaillard (2008). Measuring senescence in wild animal populations: Towards a longitudinal approach. Functional Ecology 22:393–406.

Nussey, D. H., H. Froy, J. F. Lemaitre, J. M. Gaillard, and S. N. Austad (2013). Senescence in natural populations of animals: Widespread evidence and its implications for bio-gerontology. Ageing Research Reviews 12:214–225.

Ottinger, M. A. (2007). Neuroendocrine aging in birds: Comparing lifespan differences and conserved mechanisms. Ageing Research Reviews 6:46–53.

Peig, J., and A. J. Green (2009). New perspectives for estimating body condition from mass/length data: The scaled mass index as an alternative method. Oikos 118:1883–1891.

Rahn, H., I. C. T. Nisbet, C. V. Paganelli, and G. C. Whittow (1976). Regulation of incubation water-loss in eggs of seven species of tern. Physiological Zoology 49:245–259.

Rebke, M., P. H. Becker, and F. Colchero (2017). Better the devil you know: Common Terns stay with a previous partner although pair bond duration does not affect breeding output. Proceedings of the Royal Society B, Biological Sciences 284:20161424.

Reid, J. M., E. M. Bignal, S. Bignal, D. I. McCracken, and P. Monaghan (2003). Age-specific reproductive performance in Red-billed Choughs *Pyrrhocorax pyrrhocorax*: Patterns and processes in a natural population. Journal of Animal Ecology 72:765–776.

Riechert, J., O. Chastel, and P. H. Becker (2012). Why do experienced birds reproduce better? Possible endocrine mechanisms in a long-lived seabird, the Common Tern. General and Comparative Endocrinology 178:391–399.

R Project Core Team (2018). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/

Saino, N., M. Fasola, and E. Crocicchia (1994). Adoption behaviour in Little and Common terns (Aves; Sternidae): Chick benefits and parents' fitness costs. Ethology 97:294–309.

Stearns, S. C. (1992). The Evolution of Life Histories. Oxford University Press, Oxford, UK.

Szostek, K. L., and P. H. Becker (2012). Terns in trouble: Demographic consequences of low breeding success and recruitment on a Common Tern population in the German Wadden Sea. Journal of Ornithology 153:313–326.

Szostek, K. L., P. H. Becker, B. C. Meyer, S. R. Sudmann, and H. Zintl (2014). Colony size and not nest density drives reproductive output in the Common Tern *Sterna hirundo*. Ibis 156:48–59.

Tims, J., I. C. T. Nisbet, M. S. Friar, C. Mostello, and J. J. Hatch (2004). Characteristics and performance of Common Terns in old and newly-established colonies. Waterbirds 27:321–332.

van de Pol, M., and S. Verhulst (2006). Age-dependent traits: A new statistical model to separate within- and between-individual effects. The American Naturalist 167:766–773.

Vaupel, J. W., and A. I. Yashin (1985). Heterogeneity's ruses: Some surprising effects of selection on population dynamics. American Statistician 39:176–185.

Vedder, O., and S. Bouwhuis (2018). Heterogeneity in individual quality in birds: Overall patterns and insights from a study on Common Terns. Oikos 127:719–727.

Wood, S. (2019). Package 'nmgcv'. R-package. https://cran.r-project.org/web/packages/mgcv/mgcv.pdf

Zhang, H., M. Rebke, P. H. Becker, and S. Bouwhuis (2015a). Fitness prospects: Effects of age, sex and recruitment age on reproductive value in a long-lived seabird. The Journal of Animal Ecology 84:199–207.

Zhang, H., O. Vedder, P. H. Becker, and S. Bouwhuis (2015b). Age-dependent trait variation: The relative contribution of within-individual change, selective appearance and disappearance in a long-lived seabird. The Journal of Animal Ecology 84:797–807.

APPENDIX A

Study Site, Field Methods, Definitions of Variables, and Summary Statistics

Bird Island (41.669°N, 70.717°W) is located in Buzzards Bay, Massachusetts, USA, a shallow bay separating Cape Cod from the mainland of North America. Common Terns (hereafter, terns) nesting at Bird Island form part of a metapopulation with those nesting at 2 other colony sites within the bay, Ram and Penikese islands, 10–26 km apart. Adult terns move fairly freely among these 3 sites (Breton et al. 2014): some terns included in this study are known through re-trapping to have nested one or more times at one of the other sites. However, total numbers were too large to keep track of individual birds from year to year. Appendix Table 1 summarizes numbers of tern nests at the 3 colony sites in Buzzards Bay, 1986–2004; the counts refer to the peak period of nesting in each year (for definition, see Appendix D).

At Bird Island, ~1,800 pairs of terns nested throughout the study period (Appendix Table 1). Terns nested at densities of 0.5–1.0 nests m^{-2} at low elevations (mostly <2 m above high-water mark) among scattered vegetation on substrates of sand, gravel, turf, and cobbles. In each year, the nesting area was checked daily or almost daily during the peak period and most nests were marked in sequence as soon as they were found. The laying date for each nest is defined as the date on which the first egg in the clutch was laid and was usually accurate to ±1 day. The distribution of laying dates was skewed, with most nests initiated within the first 25–30 days of the season and a scattering of nests initiated thereafter, some of which were known to be re-layings by pairs that had failed earlier in the season. To exclude these re-layings and other late outliers, we defined a cutoff date as 25 days after the 10th percentile laying date for each year. Cutoff dates were then adjusted for variations among age classes (see Appendix D for details). A total of 108 nests (2.9%) with laying dates after the cutoff date for their years and ages were excluded from all subsequent analyses.

Eggs were marked in sequence of laying and the first egg in each clutch was weighed within 4 days of laying: the fresh mass ("egg mass") was estimated to ±0.1 g by adding 0.12 g for each day of incubation prior to weighing (Rahn et al. 1976). Clutch size is the maximum number of eggs recorded in each nest within 7 days after the laying date (range: 1–4 eggs, usually 2–3): egg loss was rare during the study period and errors in recording clutch size were probably very infrequent.

In each year (except 1995), 3–10 intensive study plots of 10–100 m^2 were selected during the egg-laying or incubation periods. Study plots were selected to be representative of nests established in different parts of the nesting area, except that in some years, plots were established specifically to monitor the performance of older birds. Most adult terns that nested in study plots were trapped. Within study plots, nests were visited daily or near-daily during the hatching period and chicks were banded at hatching. Hatching success (a discrete variable with values 0, 0.33, 0.5, 0.67, 0.75, or 1.0) is defined as the proportion of eggs in each clutch that hatched.

Study plots were fenced to retain chicks for detailed study (Nisbet and Drury 1972). Chicks were banded at hatching and were weighed daily or on alternate days until they fledged, disappeared, or were found dead. Most surviving chicks in most years were followed until they could fly (usually at ages 22–27 days), but some chicks (mean ~6%) that hatched late in the season could not be followed to a certain outcome. In such cases, we deemed a chick to have survived to fledging if it reached the age of at least 17 days, attained a body weight of at least 85 g, was gaining weight at last encounter, and was not found dead subsequently. Chicks that did not attain a body weight of 85 g by age 17 days and were losing weight at last encounter were deemed to have died; other chicks were deemed indeterminate. Fledging success (a discrete variable with values 0, 0.33, 0.5, 0.67, or 1.0) is defined as the proportion of chicks in each brood that survived to fledging. Productivity (a discrete variable with values 0, 1, 2, or 3) is defined as the number of chicks raised to fledging from each nest, even in cases where chicks were known or suspected to have been adopted and raised by pairs other than their genetic parents (see Appendix G for details). Fledging success and productivity were assigned only if outcomes were determined for all chicks that hatched from the clutch.

Appendix Table 2 lists summary statistics for tern breeding performance at Bird Island, 1986–2004. The last 3 lines in Appendix Table 2 show regressions on year for each variable. Although laying date advanced significantly during the study period, most of this advance was attributable to the increase in mean age (27.6% of variance); the linear trend with year explained only 1.7% of variance after controlling for age.

The ages of both mates were known for 448 nests, so that each nest was included twice in the database. We selected one case at random for each of these nests and excluded the other 448 cases from all analyses.

Birds that were subjected to experimental treatments that might have affected their parental performance (Arnold et al. 2004, Nisbet et al. 2004) were excluded from analyses of hatching success, fledging success, and productivity, but data on their laying dates, egg masses, and clutch sizes were used; all data for these birds in nonexperimental years were used. All data for birds from which blood samples were taken were used, as pilot studies had shown that this did not affect any aspect of their breeding performance.

APPENDIX B

Sexing Criteria

Terns were sexed genetically (n = 92), by being seen to lay an egg (n = 3), by the presence or absence of a palpable egg in the oviduct or distended cloacae when trapped during egg-laying (n = 264), by body mass during egg-laying (n = 357), by observation of copulation or repeated courtship-feeding (n = 150), by head length (n = 1,201), by the same characters recorded in previous or subsequent years (n = 1,019), and/or from the sex of their mates in one or more years (n = 694); most terns were sexed using 2 or more of these characters. Although formal estimates of percent confidence are available only for sexing based on head length (Nisbet et al. 2007), we judged that 1,757/2,046 (86%) of the terns in the study were sexed with ≥95% confidence; the remainder were treated as of unknown sex and were excluded from most analyses because BCI could not be calculated (see Appendix C).

APPENDIX C

Calculation of Body Condition Index (BCI)

A body condition index (BCI) for each bird in each year was calculated using the methods of Peig and Green (2009). Head length (from back of skull to tip of bill) was used as an index of body size. A scaling exponent b was calculated by running a standardized major axis regression of ln(body mass) vs. ln(head length) and estimating b as the slope of the best-fit line. BCI for the ith bird was then calculated from the formula $BCI_i = M_i(H_m/H_i)^b$ where M_i and H_i are the body mass and head length of the ith bird and H_m is the mean head length of all birds of the same sex in the same year. H_m and b were calculated separately for males and females so that BCI was only estimated for birds of known sex. The units of BCI are grams.

We ran a GAMM to test for dependence of BCI on age and laying date, with year and bird identity as random factors. BCI declined with increasing laying date (Appendix Figure 3B), but the effect was only manifested among birds with laying dates later than 35 (June 4) and was not statistically significant (P = 0.12, n = 2,132). There was no relationship of BCI to age (Appendix Figure 3A).

APPENDIX TABLE 1. "Peak of season" counts of Common Tern nests at the 3 colony sites in Buzzards Bay, Massachusetts, 1986–2004. The "peak of season" is defined in Appendix D.

Year	Bird Island	Ram Island	Penikese Island	Total
1986	1,129	0	0	1,129
1987	1,337	0	1	1,338
1988	1,613	0	0	1,613
1989	1,879	0	0	1,879
1990	1,803	0	0	1,803
1991	1,780	0	0	1,780
1992	1,575	1	18	1,594
1993	1,829	36	0	1,865
1994	1,803	160	140	2,103
1995	1,590	431	9	2,030
1996	1,780	1,100	7	2,887
1997	2,033	930	8	2,971
1998	1,903	1,307	137	3,347
1999	1,836	1,887	101	3,824
2000	1,880	2,030	126	4,036
2001	2,136	1,890	278	4,304
2002	1,702	2,307	279	4,288
2003	2,054	2,000	661	4,715
2004	1,761	2,938	631	5,330

APPENDIX TABLE 2. Summary statistics (mean ± SD) for 3,787 nests of known-aged Common Terns at Bird Island, 1986–2004. Laying date is the date of laying of the first egg in the clutch, scaled so that May 1 = 1. For definition of BCI (body condition index), see Methods. A dash (–) indicates too few values (<10) for meaningful averages. Sample sizes for No. hatched and No. fledged were usually much smaller than the total number of nests studied (*n*). The last 3 rows in the table give regression coefficients (*b* ± SE) on year, sample sizes (*n*), and probabilities (*P*).

Year	*n*	Age (yr)	Clutch	No. hatched	No. fledged	Laying date	Egg mass (g)	BCI (g)
1986	71	8.82 ± 3.56	2.80 ± 0.43	–	–	20.0 ± 5.8	21.56 ± 1.48	130.1 ± 14.2
1988	67	7.06 ± 3.87	2.34 ± 0.51	2.50 ± 0.50	–	24.4 ± 7.4	21.40 ± 1.29	123.0 ± 13.5
1989	66	7.65 ± 4.01	2.42 ± 0.55	2.02 ± 0.79	0.90 ± 0.72	24.2 ± 8.5	21.26 ± 1.50	121.0 ± 10.9
1990	36	8.67 ± 4.58	2.42 ± 0.55	2.09 ± 0.79	0.97 ± 0.81	24.6 ± 7.7	21.31 ± 1.57	125.1 ± 11.5
1991	416	7.10 ± 3.68	2.46 ± 0.57	2.06 ± 0.70	0.86 ± 0.69	30.0 ± 6.6	21.00 ± 1.45	116.5 ± 16.5
1992	282	8.07 ± 3.82	2.21 ± 0.60	2.71 ± 0.45	0.92 ± 0.81	29.7 ± 7.6	20.65 ± 1.61	130.4 ± 15.9
1993	32	8.87 ± 3.28	2.73 ± 0.44	2.58 ± 0.57	1.83 ± 0.90	20.9 ± 4.5	21.12 ± 1.61	124.5 ± 13.8
1994	51	9.96 ± 3.66	2.35 ± 0.59	1.96 ± 0.92	1.04 ± 0.72	21.8 ± 5.6	20.91 ± 1.07	130.5 ± 16.4
1995	606	9.45 ± 3.67	2.67 ± 0.53	2.34 ± 0.57	–	19.9 ± 6.6	20.86 ± 1.63	123.6 ± 12.4
1996	476	9.71 ± 4.13	2.38 ± 0.59	1.92 ± 0.61	0.87 ± 0.53	23.7 ± 5.6	20.58 ± 1.63	128.5 ± 14.9
1997	522	9.76 ± 4.30	2.53 ± 0.61	2.10 ± 0.92	0.85 ± 0.68	21.7 ± 6.8	20.80 ± 1.78	126.0 ± 12.3
1998	166	10.00 ± 4.77	2.71 ± 0.47	2.70 ± 0.45	0.80 ± 0.60	17.2 ± 5.8	21.64 ± 1.28	130.8 ± 13.6
1999	229	11.59 ± 4.47	2.53 ± 0.54	2.18 ± 0.91	0.89 ± 0.67	13.6 ± 5.6	20.81 ± 1.79	130.9 ± 13.1
2000	143	10.06 ± 4.93	2.48 ± 0.53	2.38 ± 0.70	1.38 ± 0.62	20.3 ± 7.0	21.07 ± 1.32	131.6 ± 11.9
2001	161	10.76 ± 4.88	2.50 ± 0.55	2.28 ± 0.61	–	16.9 ± 5.0	21.20 ± 1.57	126.5 ± 11.2
2002	207	10.40 ± 5.28	2.16 ± 0.56	1.88 ± 0.64	0.50 ± 0.76	22.8 ± 9.9	20.88 ± 1.71	119.5 ± 12.8
2003	147	10.92 ± 5.18	2.04 ± 0.63	1.55 ± 0.78	0.55 ± 0.66	26.6 ± 6.7	20.94 ± 1.74	123.3 ± 12.9
2004	109	9.96 ± 5.30	2.41 ± 0.53	2.08 ± 0.79	1.00 ± 0.82	23.4 ± 6.6	20.93 ± 1.68	125.5 ± 14.1
b ± SE		0.0243 ± 0.017	−0.0045 ± 0.0025	0.0031 ± 0.0053	−0.010 ± 0.008	−0.484 ± 0.031	−0.010 ± 0.007	−0.365 ± 0.070
n		3,787	3,389	1,179	525	3,747	2,528	2,428
P		<0.001	0.074	0.56	0.19	<0.001	0.14	<0.001

APPENDIX TABLE 3. Autocorrelations in variables indexing reproductive performance and individual quality.

Variable	Autocorrelation coefficient	*n*	*P*
Laying date	0.270	775	<0.001
Resid-laydate	0.246	427	<0.001
BCI	0.403	427	<0.001
Relative productivity	0.139	19	0.56

APPENDIX TABLE 4. Results of logistic analysis of covariance for the dependence of reproductive performance and factors indexing individual quality on survival from 1991–1994 to 1995–1997.

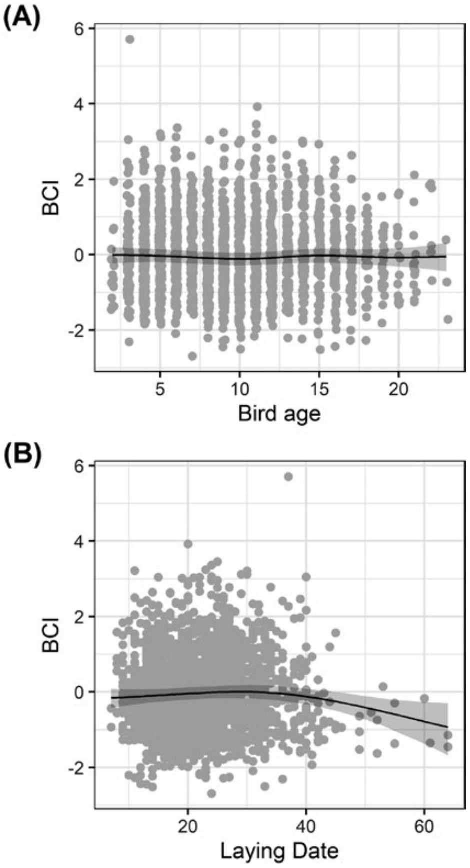
		Survival		Year effect (1991 < 1992–1994)	
Variable	*n*	Wald χ^2	*P*	Wald χ^2	*P*
Laying date	763	4.81	0.029	6.07	0.014
Residual laydate	481	0.06	0.81	3.78	0.052
BCI	481	0.66	0.42	4.05	0.044
Relative productivity	89	0.03	0.86	3.85	0.056

APPENDIX TABLE 5. Mean laying dates in each age class, averaged over the 19-yr study period, 1986–2004. Relative laying date is the difference between the laying date for each nest and the mean over all age classes in the same year. The age classes are divided into 9 groups with similar mean values for relative laying date.

	Relative laying date			
Age (yr)	Mean	SE	*n*	Group mean
2	29.0	0.36	30	29
3	15.4	0.52	204	15
4	9.0	0.43	297	9
5	5.0	0.41	335	5
6	2.8	0.44	290	3
7	0.9	0.44	293	1
8	0.8	0.38	397	1
9	0.2	0.41	334	1
10	−1.2	0.42	317	−1
11	−1.3	0.46	260	−1
12	−2.2	0.52	209	−2
13	−2.5	0.56	179	−2
14	−1.8	0.56	175	−2
15	−3.3	0.65	133	−3
16	−3.2	0.68	122	−3
17	−3.5	0.81	86	−3
18	−3.1	0.92	66	−3
19	−4.0	1.08	48	−3
20	−3.6	1.36	30	−3
21	−0.4	1.26	18	−3
22	−5.4	2.25	11	−3
23	−5.4	2.36	10	−3
24	−2.8	2.64	9	−3
25	−1.8	3.34	5	−3
26	−2.0		1	−3
28	−17.0		1	−3

APPENDIX TABLE 6. Tenth percentile laying dates for all nests in each year. Laying dates are scaled such that May 1 = 1.

Year	10th percentile laying date
1986	12
1987	13
1988	14
1989	14
1990	16
1991	25
1992	22
1993	16
1994	14
1995	13
1996	18
1997	16
1998	11
1999	9
2000	13
2001	10
2002	14
2003	20
2004	15



APPENDIX FIGURE 3. Dependence of body condition index (BCI) on (**A**) age and (**B**) laying date, based on a GAMM (see Appendix C). The plotted values of BCI are standardized relative to their mean of 127.7 g and SD of 14.3 g. The shaded areas show 95% CI.

APPENDIX D

Assignment of Cutoff Dates to Exclude Renestings

The peak period of nesting in each year is defined as the period until 25 days after the date on which the 10th percentile nest was initiated; in most years, this date coincided with a minimum frequency of initiations of new nests. Nests initiated after this date included some re-layings by older pairs that had failed earlier. To exclude such re-layings, we defined a "cutoff date" for each age class as 25 days after the 10th percentile date for birds of the same age in each year. The 10th percentile is used as the basis for these definitions rather than the earliest nest because the distribution of the early nests is sometimes irregular, and because the earliest nest may not have been located in some years. To estimate the 10th percentile date for each age and year, we used the average distribution of laying dates for birds of each age class over the 19-yr period of study (Appendix Table 5). The age classes fell into 9 groups, with group means for relative laying date ranging from +29 days for birds aged 2 yr to −3 days for birds aged ≥12 yr (Appendix Table 5). The 10th percentile dates for each year are listed in Appendix Table 6. The cutoff date for each age class in each year was obtained by adding the mean relative laying date for the age class group from Appendix Table 5 to the 10th percentile laying date for the year from Appendix Table 6. We excluded from subsequent analyses all nests initiated after the cutoff date for their age class and year ($n = 108$, 2.3% of all cases).

APPENDIX E

Mate Age and Mate Retention

In 448 cases, the ages of both mates were known (in ~1,300 other cases, mates of known-aged birds were trapped but were of unknown age, either banded when first encountered as breeding adults or unbanded). There were too few cases where the same known-aged pair was studied in more than one year to run GAMMs with bird identity (pair identity) as a random factor, so we analyzed this dataset using year as the only random factor. Bird age and mate age were strongly correlated (Pearson $r = 0.614$, $n = 448$), so we used residual mate age (the residual from a linear regression of mate age on bird age) as a fixed factor. This variable indexes the degree to which the mate was older (or younger) than the focal bird.

Within the reduced sample of nests for which the ages of both mates were known, addition of residual mate age to the GAMM for laying date improved the fit (explained deviance increasing from 69.8% to 71.9%; mean effect = −0.41 days yr^{-1}, CI: −0.33 to −0.50 days yr^{-1}; $n = 321$), but $\Delta\mathrm{AIC}_c$ was only 2.0. Addition of residual mate age to the GAMM for productivity did not improve the fit (explained deviance increasing from 15.8% to 16.1%; mean effect = −0.014 yr^{-1}, CI: −0.044 to 0.027 yr^{-1}; $n = 95$); $\Delta\mathrm{AIC}_c$ was 3.3.

To estimate the dependence of mate retention on age, we compiled all cases in which a bird of known age a was encountered with a banded mate (mate A) and was encountered again with a banded mate (mate B) in the next year; we included cases in which birds A and/or B were of unknown age. If a bird was encountered 3 or more times with a banded mate, only contiguous pairs of years were used. We defined a mate retention function $m(a)$ which took the value 1 if A = B and 0 if A ≠ B. Logistic regression showed that $m(a)$ did not depend on the age of the focal bird a at the beginning of the interval. The regression equation was:

$$\mathrm{logit}\,(m) = -\,(0.71 \pm 0.27) \quad - \quad (0.018 \pm 0.020)\,a.$$

The dependence on a was not significant (Wald $\chi^2 = 0.79$, $P = 0.38$). The mean value of m across all ages was 0.675 ($n = 255$).

APPENDIX F

Tests for Selective Survival of High-performing Birds

To test for selective survival of high-performing birds, our data were insufficient to run a formal capture–mark–recapture analysis across all years with breeding performance as a covariate. Instead, we used 2 other approaches. First, we calculated autocorrelations in the laying date, resid-laydate, BCI, and productivity variables, to test whether breeding performance was consistent within individuals among years. To control for differences among years, we defined relative productivity as the difference between the productivity of the focal bird and the mean productivity over all pairs in the same year. Second, we analyzed data for birds that were encountered as breeders in 1991–1994 and were recaptured in 1995–1997, the years of most systematic resampling. In each year from 1995 to 1997, ~45% of the terns breeding at Bird Island were trapped using stratified random sampling (Nisbet and Cam 2002), so that the probability of detection of birds that survived to 1995 or later (assuming annual survival of 0.88 yr^{-1}; Breton et al. 2014) would have been ~0.80. We used logistic analysis of covariance to test whether birds with earlier laying dates, higher BCI, or higher productivity in 1991–1994 were more likely to have survived to 1995 or later, with initial year (1991–1994) as a covariate.

Autocorrelations were significant and positive for laying date, resid-laydate, and BCI, but not significant for relative productivity, although sample size was small in the last case (Appendix Table 3). Birds with earlier laying dates in 1991–1994 were more likely to survive to 1995 or later, but there were no differences for resid-laydate, BCI, or relative productivity (Appendix Table 4). In all these comparisons, birds encountered in 1991 were more likely to survive to 1995 or later than those encountered in 1992–1994 (Appendix Table 4); there were no differences among years for birds encountered in 1992–1994 (all $P > 0.27$, details not shown).

APPENDIX G

Chick Adoptions

Adoption of chicks was fairly frequent at Bird Island during the study period (Friar 2004, Nisbet et al. 2020a). Adoptions were usually initiated by chicks ≤2 days old, which spontaneously left their natal broods and sought parental care from other nearby pairs. In years when this was studied in detail, one or more chicks from 10–20% of nests sought adoption in this way; about half of these were accepted by and received parental care from other pairs and about one-third of these were raised to fledging (I. C. T. Nisbet and M. S. Friar personal observations). Parents that adopted chicks usually lost one or more of their own chicks to starvation, so that adoption can be interpreted as a form of brood parasitism, in which the inclusive fitness of donor parents (and their chicks, including siblings of the adoptees) increases while that of adopting parents and their chicks decreases (Saino et al. 1994, Friar 2004). Documentation of adoptions requires intensive studies, including individual marking of chicks and adults at multiple nests and lengthy observations to verify which parents cared for each marked chick; we only attempted this in a few years and did not necessarily record all cases even in those years. Accordingly, in this paper we define fledging success and productivity as the proportion and number of chicks raised to fledging from eggs laid by the birds attending the clutch at the time of laying (assumed to be the genetic parents), regardless of whether the chicks were raised by their genetic parents or by other pairs, and regardless of whether the genetic parents adopted unrelated chicks from other broods. Hence, these indices of breeding performance measure the contribution to the inclusive fitness of the genetic parents, rather than the ability of these parents to raise chicks in their care to fledging.