1   **Individual variation and the source-sink group dynamics of extra-group paternity in a social mammal**

2   Running title: Extra-group paternity in badgers

3

4   Paula H. Marjamäki*[1,2], Hannah L. Dugdale[3], Deborah A. Dawson[2], Robbie A. McDonald[4], Richard Delahay[5],

5   Terry Burke[2], Alastair J. Wilson[1]

6

7   [1]Centre for Ecology and Conservation, University of Exeter, Penryn Campus, Penryn, Cornwall, TR10 9FE, UK

8

9   [2]NERC Biomolecular Analysis Facility, Department of Animal and Plant Sciences, University of Sheffield,

10   Western Bank, Sheffield, S10 2TN, UK

11

12   [3]School of Biology, Faculty of Biological Sciences, University of Leeds, Leeds, LS2 9JT, UK

13

14   [4] Environment and Sustainability Institute, University of Exeter, Penryn Campus, Penryn, Cornwall, TR10

15   9FE, UK

16

17   [5]National Wildlife Management Centre, Animal and Plant Health Agency, Woodchester Park,

18   Gloucestershire, GL10 3UJ, UK

19

20   *Corresponding author:

21   Paula H. Marjamäki
22   Centre for Ecology and Conservation, University of Exeter, Penryn Campus, Penryn, Cornwall, TR10 9FE, UK
23   Tel: 07527913031
24   Email: pm343@exeter.ac.uk
25

26

27

28

29

30   **Abstract**

31   Movement of individuals, or their genes, can influence eco-evolutionary processes in structured

32   populations. We have limited understanding of the extent to which spatial behaviour varies among groups

33   and individuals within populations. Here we use genetic pedigree reconstruction in a long-term study of

34   European badgers (*Meles meles*) to characterise the extent of extra-group paternity, occurring as a

35   consequence of breeding excursions, and to test hypothesised drivers of variation at multiple levels. We

36   jointly estimate parentage and paternity distance (PD; distance between a cub's natal and its father's social

37   group), and test whether population density and sex ratio influence mean annual PD. We also model cub-

38   level PD and extra-group paternity (EGP) to test for variation among social groups and parental individuals.

39   Mean PD varied among years but was not explained by population density or sex ratio. However, cub-level

40   analysis shows strong effects of social group, and parental identities, with some parental individuals being

41   consistently more likely to produce cubs with extra-group partners. Group effects were partially explained

42   by local sex ratio. There was also a strong negative correlation between maternal and paternal social group

43   effects on cub paternity distance, indicating source-sink dynamics. Our analyses of paternity distance and

44   EGP indicate variation in extra-group mating at multiple levels – among years, social groups and individuals.

45   The latter in particular is a phenomenon seldom documented and suggests that gene flow among groups

46   may be disproportionately mediated by a non-random subset of adults, emphasising the importance of the

47   individual in driving eco-evolutionary dynamics.

48

49   Keywords: extra-group paternity, individual variation, *Meles meles*, parentage assignment, source-sink

50   dynamics

Introduction

Movement of individuals and/or gametes influences the dynamics, persistence and genetic diversity of spatially structured populations (Ronce, 2007). Understanding movement is therefore crucial for wildlife conservation and management as it can determine species distributions (Holt, 2003), impact the vulnerability of populations to extinction (Thomas, 2000) and play an important role in the transmission of infections (Pope et al., 2007). Behaviours linked to 'dispersal', in the broadest sense of any movement with potential consequences for gene flow (Ronce 2007), are widely viewed as adaptive, allowing individuals to escape from locally intense competition for resources or mates (Daniels & Walters, 2000; Matthysen, 2005), seek good or compatible genes in potential mating partners (Hamilton 1990; Zeh & Zeh 1996), or avoid inbreeding by leaving the vicinity of related individuals (Greenwood, 1980). However, as such movements carry risks as well as benefits, associated behaviours are likely to have evolved under the influence of multiple interacting factors that ultimately shape the balance of costs and benefits (Bowler & Benton 2005; Ronce 2007).

Some of the factors influencing the costs and benefits of movement and dispersal are well documented. For instance, sex (Clarke et al. 1997; Beirinckx et al. 2006; Rabasa & Gutie 2007), age (Dale et al. 2005; Bowler & Benton 2009; Kentie et al. 2014), and density (e.g. Matthysen 2005; Nowicki & Vrabec 2011) are common drivers of variation in many taxa, although density effects can themselves be scale-dependent (e.g. Marjamäki et al. 2013). However, in addition to demographic and ecological effects, it is also becoming apparent that populations can harbor among-individual variation in the tendency to disperse. Our understanding of what drives this variation within animal populations remains limited, although social interactions and behavioral differences (e.g. "personality" variation in exploratory tendency) likely play an important role (e.g. Cote et al. 2010; Patrick et al. 2012; Weiß et al. 2016).

In this study, we employ an indirect approach to test for and investigate sources of variation in breeding excursions in a population of European badgers (*Meles meles*) in southwest England. Temporary excursions relating to mate acquisition are common in many populations but, while they will have important consequences for fine scale gene flow and genetic structure (e.g. among groups), temporary and short-term excursions can be difficult to observe directly. Nonetheless, in the absence of direct

78   observation of movement, indirect inferences on breeding excursions can be made from genetic data. This

79   can be done, for example, by characterising population genetic structure (or lack thereof; Wilson et al.

80   2004), or by detecting extra-pair or extra-group paternity (hereafter 'EGP'), which is commonly seen in

81   birds and mammals (Griffith et al. 2002; Isvaran & Clutton-Brock 2007). Combined with genetic pedigree

82   analysis, the latter approach allows identification of those individuals engaging in, as well as resulting from,

83   extra-group matings, enabling the drivers of among-individual variation to be investigated.

84   Badgers are a facultatively social species and form social groups at high densities through

85   retention of offspring in natal groups (Kruuk & Parish 1982; da Silva et al. 1994). These social groups,

86   ranging from 1 to 22 individuals of mixed age and sex, form discrete, defended territories containing

87   several communal setts (underground dens). Badgers have a polygynandrous mating system where as

88   many as seven males and females might breed within a social group annually (Dugdale et al. 2007). While

89   within-population movement is common (e.g., detected in 44% of individuals studied by Rogers et al.

90   1998), the majority of movements between social groups are temporary, with short-term movements

91   tending to be predominantly between neighboring social groups (Rogers et al., 1998). High rates of EGP (up

92   to 50% reported in high-density populations; Carpenter et al. 2005; Dugdale et al. 2007) are also consistent

93   with an important role for breeding excursions in mediating gene flow, though whether EGP is mediated

94   through transient contact between individuals, or temporary integration of individuals into social groups

95   (or both) is not yet clear.

96   We use a long-term dataset on individually marked badgers from Woodchester Park

97   (Gloucestershire, England) to reconstruct a genetic pedigree and indirectly estimate breeding excursions.

98   We build on a previous parentage analysis of the population (Carpenter et al. 2005) to reconstruct a

99   pedigree using a larger sample, more markers and more powerful parentage assignment methods.

100   Crucially, for current purposes we adopt a Bayesian approach to pedigree analysis, which allows us to make

101   better use of spatial and group membership information to improve the number of assigned relationships

102   and our confidence in them (Hadfield et al. 2006). From this we simultaneously estimate both the pedigree

103   structure and the mean distance between the father's social group and the cub's natal group (hereafter

104   'paternity distance') for each annual cohort. We first ask whether paternity distance varies among years as

4

105   a function of population density and/or sex ratio, before using assigned parent-offspring relationships to

106   test for among-individual (parent) variation in extra-group mating. Finally, noting that from a cub's

107   perspective, EGP and non-zero paternity distance may reflect temporary excursions by either parent, we

108   ask whether among-parent variation can be explained by known predictors of breeding behavior in other

109   systems, including intrinsic factors (e.g. age, body mass) and social group properties.

110

111

112   **Methods**

113   **Study population & sampling**

114   The badger population at Woodchester Park (51°42'35"N 2°16'42"W), Gloucestershire, UK, has

115   been subject to an ongoing mark-recapture study since 1976. The study area is approximately 11 km^2 and

116   consists of a steep-sided, wooded valley surrounded by farmland. Here we utilize data from a 30-year

117   period from 1985 to 2014, for which badgers were trapped and sampled up to four times a year. Steel

118   mesh box traps were deployed at active badger setts and set to catch for two consecutive nights after a

119   period of 4-8 days of pre-baiting with peanuts. Trapped badgers were anaesthetized (de Leeuw 2004) prior

120   to examination and at first capture each individual received a unique identifier tattoo on their abdomen.

121   Capture location, sex, age (if birth year known) or age class (adult, yearling, cub, based on size and tooth

122   wear) and body weight were recorded (Delahay et al. 2013). Approximately 20-30 guard hairs were plucked

123   and stored in 80% ethanol for microsatellite genotyping. After a recovery period, all badgers are released at

124   the point of capture. The total trapping dataset is comprised of over 15,000 captures for 3,283 individuals.

125   While most badgers are first caught as cubs or yearlings, 19% were first captured as adults and likely

126   represent a minimum estimate of immigration into the population. Social group territorial boundaries were

127   determined for each year of the study by bait marking (Delahay et al. 2000b). A total of 45 defined social

128   groups were counted throughout the study period, but from 1996 onwards sampling was focussed on 20—

129   25 groups only. Thus, the variation in the number of social groups reflects variation in both sampling effort

130   through time and the configuration of social groups, which occasionally undergo fissions and fusions

131 (though territories are largely stable over time; Delahay et al. 2000a; Robertson et al. 2014). All work was

132 carried out under licence from the UK Home Office and from Natural England.

133

134 **DNA extraction & genotyping**

135 Microsatellite data used for parentage analyses have been produced as part of the ongoing

136 Woodchester Park study. For current purposes, we used existing published data (Carpenter et al. 2005)

137 coupled with de novo genotyping at 6 loci described in Carpenter et al. (2003) and Lopez-Giraldez et al.

138 (2007). In brief, individuals trapped between 1986 and 2002 have been genotyped with DNA extraction

139 from hair samples according to protocols outlined in Carpenter et al. (2005), while samples between 2003

140 and 2014 were genotyped at the NERC Biomolecular Analysis Facility (University of Sheffield, UK) in batches

141 across several time periods using the ammonium acetate extraction method described in Richardson et al.

142 (2001). A minimum of 5 hairs with visible roots were used per individual.

143 Individuals have been genotyped at between 16 and 22 autosomal microsatellite loci, with slightly

144 different, but overlapping subsets of markers used over the course of the project. We used a 2-µl Qiagen

145 Multiplex PCR reaction (Qiagen Inc., Valencia, USA) and fluorescently-labelled primer sets, before

146 separation of the amplicons on a 48-capillary ABI 3730 DNA Analyzer using Prism set D and a ROX size

147 standard and genotype scoring using GENEMAPPER 3.7. Samples described in Carpenter et al. (2005) were

148 genotyped at 16 loci (Mel 101-117; as described in Carpenter et al. 2003). An additional 6 loci were added

149 to subsequent genotyping efforts (Mel 1, 10, 12, 14, 15 & 116; Carpenter et al. 2003, Lopez-Giraldez et al.

150 2007) though for 209 individuals born (or captured for the first time) after 2011, markers Mel 15 and 106

151 were not used. As genotyping has been done in batches over a number of years, samples have been cross-

152 validated by retyping subsets of previously genotyped individuals (min. 15% of samples). This was used to

153 calibrate allele sizes at each locus to ensure consistent scoring across time periods and different

154 sequencers. After scoring genotypes, we tested for deviations from Hardy-Weinberg equilibrium (HWE) and

155 linkage equilibrium (LD) for pairs of loci using 40 unrelated individuals (based on ML-Relate relatedness

156 estimates <0.125) using Genepop 4.4.3 (Raymond & Rousset 1995). P-values for LD tests were corrected to

157 account for multiple tests (false discovery rate; Benjamini & Hochberg 1995). No deviation from HWE (k =

158 22, alpha = 0.05) or LD (LD: k = 231, alpha = 0.05, adjusted p = 0.05-0.0002) were found. Null allele

159 frequencies were estimated using CERVUS 3.0.7 (Marshall et al. 1998) and were <0.1 for all loci. Therefore,

160 all loci were retained.

161   We also estimated mean allelic dropout (e1) and false allele rates (or stochastic sampling error, e2),

162 using a random subset of individuals that were re-genotyped and analysed using PEDANT 1.0 (Johnson &

163 Haydon, 2007) (Table S1). Overall, genotypes were available for 2,204 (out of 2,811) trapped individuals, at

164 a mean (±standard deviation) of 16.1 (±5.1) loci per individual. Across loci the mean observed and expected

165 heterozygosity were 0.56 (SD 0.15) and 0.61 (SD 0.13), respectively, and the mean number of alleles per

166 locus was 4.85 (SD 1.47).

167

168 **Parentage analysis**

169 We conducted Bayesian parentage analysis for 1768 genotyped cubs trapped between 1986 and 2014

170 inclusive, using MasterBayes 2.54 (Hadfield et al., 2006) in R 3.3.0 (R Development Core Team 2016).

171 Relative to most wild birds and mammals in which molecular pedigree reconstruction has been applied,

172 badgers present a particular challenge in that they are largely nocturnal and so difficult to observe.

173 Furthermore, cubs remain underground for the first 12 weeks of life (Roper 2010), and alloparental care

174 may occur at the sett (Dugdale et al. 2010). As such, while maternal identities can often be (reliably)

175 inferred from observation in other species, this is not the case in badgers. In the absence of any known

176 parents, life-history, spatial and genetic data were used simultaneously to assign paternity and maternity

177 jointly for each cohort of cubs (n = 29) and estimate mean annual paternity distance. The final pedigree

178 used in downstream analyses was then compiled based on parental assignments that met a minimum

179 confidence threshold of 80%. For comparison, we also compiled a pedigree structure according to a stricter

180 95% confidence threshold.

181

182 *Definition of candidate parents and use of spatial data*

183 Parentage assignments were run for each annual cub cohort (n=29). Although neither parent can be

184 determined by observation we follow the approach used in other systems (e.g. Walling et al. 2010; Nielsen

et al. 2012) of applying a biologically informed set of criteria to define a non-excluded list of candidate parents for each cub. For each cohort, candidate mothers were restricted to females aged ≥2 years present in the cub's natal group (i.e. the group first captured in) in the year of birth, as females are sexually mature as yearlings and, due to delayed implantation (Yamaguchi et al. 2006), can first give birth as two-year olds. Males were considered candidate fathers (regardless of social group) if they were alive and ≥1 year of age 12 months before the cub was born, to account for delayed implantation. Individuals were designated as belonging to a social group if they were caught within the territory of that group. Individuals recorded in multiple social groups were assigned joint membership to each; in years where individuals were not caught (but were known to be alive from subsequent captures), they were assigned to the social group(s) they were recorded in the preceding year. Only individuals caught as cubs or yearlings (i.e. those with known birth year) were included as offspring in parentage analysis, while badgers first caught as adults are likely to be immigrants and were included only as candidate parents. Since age data were incomplete for badgers that were not caught as cubs or yearlings (distinguishable from adults by size and tooth wear), we assumed adults of unknown age to be 2 years of age at first capture to prevent blanket exclusion from the set of candidate parents (note, this was for parentage assignment only, and assumed ages were not used in subsequent analyses described below). Similarly, where time of death was unknown, individuals were treated as being alive (for purposes of defining status as a potential candidate parent) for 1 year (cubs; Dugdale et al. 2007) or 3 years (adults; Carpenter et al. 2005) after their last capture. Individuals with missing sex or social group data were excluded.

In addition to microsatellite data, our parentage analyses also utilised geographical location data (main sett coordinates for each social group) for all offspring and candidate fathers. Inclusion of non-genetic data is expected to improve assignment where it provides additional information about the likelihood of parentage (Hadfield et al. 2006). For most cohorts (see below) we therefore used (Euclidean) "male distance" between the main sett of the candidate father's social group and that of the cub's natal group as a predictor of paternity, which yielded an estimate for each cohort (or year) of the mean paternity distance, i.e. distance between the main sett of the assigned father's social group and that in which the cub was born. Thus, paternity distance and parentage are jointly estimated from the data in a single analysis

212 (i.e. it is not the case that distance effects on paternity likelihood are first estimated and imposed in a

213 subsequent parentage assignment). Finally, we note that, while more complete genetic sampling of the

214 population should result in greater parentage assignment success (all else being equal), the number of

215 unsampled parents is estimated in a MasterBayes analysis, not specified *a priori* as an input parameter (as

216 in some likelihood-based methods of parentage assignment). Here we have limited knowledge of the

217 completeness of genetic sampling but certainly trapping does not sample all animals present on any given

218 occasion. Quarterly recapture rates (i.e. across trapping sessions) are known to vary greatly across years,

219 from 0.15-0.73 for females and from 0.20-0.78 for males (Graham et al. 2013). Approximately 19% of

220 individuals are first trapped as adults, providing an upper bound estimate for the proportion of immigrants

221 to the study area.

222

223 *Parentage assignment settings and diagnostics*

224 Markov chains were run separately for each year (i.e. cub cohort) for 2 million iterations, with a thinning

225 rate of 100 and burn-in period of 500,000. Mismatch tolerance between cub and candidate parent was set

226 to one. Tuning parameters were specified for each cohort to ensure that Metropolis–Hastings acceptance

227 rates were within acceptable limits (0.2-0.5; Hadfield 2014). Per locus genotyping error (e1 and e2; Table

228 S1) and allele frequencies calculated based on the full dataset were provided in the model specifications (as

229 direct estimation of error rates by MasterBayes from the data, though possible in principle, is particularly

230 computationally demanding; Hadfield 2012). The presence of unsampled males (per population) and

231 females (per social group) was also allowed for each cohort. Successive samples from the posterior

232 distribution had low autocorrelation (r < 0.10) for estimates of unsampled males and paternity distance.

233 Autocorrelation for unsampled females remained high (>0.10) for several cohorts, however, parentage

234 assignments at ≥80% confidence for these cohorts did not differ when a fixed number of unsampled

235 females (one per social group) was used, therefore all cohorts were retained.

236 In six of the 29 cohorts (1988, 1993, 2001, 2009, 2013 and 2014) inclusion of male distance as a

237 predictor caused problems for the parentage assignment algorithm that we were unable to resolve. The

238 reasons for this remain unknown but could include, for instance, undetected outliers or errors in the spatial

9

239   data. For these cohorts, parentage assignment was therefore estimated without male distance as a

240   predictor meaning no direct estimate of mean paternity distance was obtained. As including the distance

241   variable is expected to increase confidence in assignments (Hadfield 2012), excluding this variable from

242   pedigree models could affect the resulting parent assignments. In order to account for this, we reran a

243   subset of cohorts (including 339 cubs) without male distance and compared assignments with and without

244   paternity distance estimation. As expected, excluding male distance generally reduced the confidence

245   assigned to a cub's most likely father, with the result that putative paternities were not assigned in 30

246   instances, when they had been with models utilising male distance. However, changes in most likely father

247   were only observed for four cubs (out of 339). In all four cases, most likely candidate fathers failed to meet

248   the 80% confidence threshold for assignment regardless of whether the male distance variable was

249   included. Therefore, based on these comparisons, we expect fewer paternities will have been assigned for

250   the six cohorts where the distance variance could not be included, but consider it unlikely that the identity

251   of the most likely father is sensitive to inclusion of male distance in many instances.

252

253   **Analysis of breeding excursion proxies**

254   We used the results of our pedigree analysis to extract and model variation in three response

255   variables relating to extra-group paternity. First, we modelled among-cohort variation in mean paternity

256   distance as estimated directly by MasterBayes (subsequently denoted PD_c). Second, for each cub with an

257   assigned father, we extracted the individual paternity distance (denoted PD_i), and also defined a binary EGP

258   variable (denoted EGP_i) according to whether the assigned father was from within (0) or outside (1) the

259   cub's natal group. If a cub was assigned both within- and extra-group paternity by the same father (e.g.

260   where a father was recorded in multiple social groups within a year), the cub was assumed to be within-

261   group offspring. Both PD_i and EGP_i are defined for the cub (i) and non-zero values therefore reflect

262   movements by the mother and/or the father beyond its own social group. We also note that these

263   individual-level estimates are necessarily derived from an estimated pedigree and thus carry over error

264   associated with parentage assignments to downstream analyses that is not readily accounted for. In this

265   respect, we also note an unavoidable trade-off, regarding analyses of PD_i and EGP_i, between using

266   assignments made at 80% confidence (increased samples size but higher error rate) or 95% confidence

267   (reduced sample size but lower error rate). Here results from analyses are presented using the lower

268   threshold but parallel analyses based on 95% confidence can be found in supplemental materials (Tables

269   S6-S8). Overall, qualitative conclusions are consistent between analyses based on the two thresholds. Note

270   however that, since MasterBayes estimates a full posterior for PD_c, uncertainty in the annual mean

271   paternity distances could be readily accounted for in our analysis of among-cohort variation.

272

273   *Among-cohort variation in annual mean paternity distance*

274   Our MasterBayes analyses generated estimated posterior distributions (15,000 values per cohort) of PD_c for

275   23 cohorts caught between 1986 and 2014 (Figure 1). As noted above, in six years (1988, 1993, 2001, 2009,

276   2013, 2014) inclusion of spatial data in the pedigree assignment step proved problematic so no estimates of

277   PD_c are available. Using a simple multiple regression model of PD_c we tested whether total population size

278   or population sex ratio, determined by dividing the number of males by total population size (as defined

279   below), explained variation in mean paternity distance. We also included a (linear) effect of year to test for

280   any systematic trend in PD_c across the study timeline. All three variables were mean centered to ease

281   interpretation of the intercept (i.e. as predicted PD_c at mean population size, sex ratio and year). Because

282   sampling effort for some social groups varied across years, proxies of total population size and population

283   sex ratio values for each year were estimated using the POPAN model in the program MARK 8.2 (White &

284   Burnham 1999) using capture data from 20 "core" social groups with consistent trapping efforts across all

285   years. Graphical representation of annual mean estimates for population size and numbers of males and

286   females can be found in Figure 1b. Badgers with missing sex information (n=2) were excluded from this

287   analysis. In order to integrate across uncertainty in annual mean paternity distance estimation, our

288   regression model was applied to the full posterior distributions of PD_c for each cohort, allowing estimation

289   of 95% credible intervals (CI) for the partial regression coefficients. These were considered significant if 95%

290   CI did not span zero.

291

292

293

Among-individual and among-group variation in paternity distance and extra-group paternity

Using the program ASReml 3.0 (VSN International Ltd., Hemel Hempstead, UK), we fitted mixed

effects models of PD_i (i.e. Euclidean paternity distance measured in meters), and EGP_i, a binary variable

assigning the offspring of each male as either within (0) or extra (1) group. For both response variables, a

Gaussian error structure was assumed but PD_i was natural log-transformed prior to analysis to reduce

positive skew in residuals. While noting that the Gaussian assumption cannot be strictly true for bounded

(ln PD_i) or binary (EGP_i) response variables, inspection of model residuals showed it to be a reasonable

approximation here (Figure S2). We therefore chose this approach as being more pragmatic than, for

instance, Bayesian implementation of generalised mixed models as it more readily allows inference on, and

modelling of hypothesized covariance between, random effects (see below). Both variables were then

scaled to standard deviation units (SDU) to ease interpretation of results.

For both response variables, models included fixed explanatory variables of maternal age, maternal

body mass, maternal group size, and maternal social group sex ratio (as linear effects) and the

corresponding paternal variables. Social group sizes (mean 6.4 SD ±3.6) reflect numbers of resident

yearlings and adults (i.e. reproductively active individuals) in the cub's conception year, where group

residency is determined from capture records each year following Vicente et al. (2007). Social group sex

ratios are calculated as the number of males divided by the total number of adult group members,

representing the proportion of males in each group (mean 0.4 SD ±0.2). These measures exclude cubs and

transient non-residents (based on criteria used by Vicente et al. 2007) caught within social group

boundaries, but represent a baseline measure for the density of potential breeders encountered by

individuals in their social group. Body mass was included to test for size-dependence of extra-group

paternity and for individuals with more than one weight measurement within a year, the mean of these

was used. Note that we also fitted the models using a standardised measure of body condition, the scaled

mass index (SMI; Peig & Green 2009), in place of body mass. In principle, this might better account for

sexual dimorphism and seasonal variation in body mass (Beirne et al 2015; Peig & Green, 2010). However,

in practice, qualitative conclusions of the analyses were unaltered, and since use of SMI in place of body

320   mass resulted in a 16% reduction in sample size, only the results of analyses using body mass are presented

321   here (results for SMI analysis can be found in Tables S3-S5). Significance of fixed effects was determined

322   using conditional Wald F-tests implemented in ASReml (with denominator degrees of freedom calculated

323   following Kenward and Roger 1997).

324   Year (as a factor), maternal and paternal identities and maternal and paternal social group IDs were

325   included as random effects in the models. This allowed us to partition variance in PD_i and EGP_i to assess the

326   relative importance of individual and group level effects (conditional on fixed effects). We make the

327   standard assumptions that random effects are normally distributed with means of zero and variances to be

328   estimated. For ease of interpretation, variance components were standardized to intraclass correlations

329   (ICC) by dividing by phenotypic variance (determined as the sum of all variance components). ICC are thus

330   interpretable as individual and group repeatabilities (R) for random effects relating to parental individuals

331   and their social groups (Nakagawa & Schielzeth, 2010). In addition, we explicitly modelled a covariance

332   term between the maternal and paternal social group identity effects. The strength and sign of this

333   relationship is biologically informative since, for instance, if groups vary in EGP in a non sex-specific way we

334   predict a positive covariance. Conversely, since cub natal and maternal social groups are the same, if EGP

335   follows a source-sink dynamic with respect to genetic consequences (i.e. some groups are net importers of

336   genes and some net exporters) we predict a negative relationship.

337   Statistical inference on random effects was by likelihood ratio test comparison of the full model to

338   reduced formulations in which (co)variance components arising from the tested random effects were

339   assumed absent. Twice the difference in log-likelihood between full and reduced models was assumed to

340   have a χ^2- distribution, and we conservatively (see Visscher 2006) assume the degrees of freedom (DF)

341   equal to the number of additional parameters in the full model.

342   The analyses described above were conducted using all available PD_i and EGP_i observations based

343   on the 80% confidence threshold for parentage assignment. To assess sensitivity of results to this choice of

344   confidence threshold, we repeated the analyses using only parentage assigned at 95% confidence. While

345   the higher threshold should reduce 'measurement error' in PD_i and EGP_i arising from erroneous

346   assignments, it also reduced sample size for analyses of these variables. Overall, conclusions regarding

347 individual and group-level variation remained broadly the same. Some inflation of variance components

348 occurred in models using the higher threshold, and there were also some changes to the significance of

349 fixed effects. Full results of these additional analyses are reported in the electronic supplement (Tables S6-

350 S8) and commented on, where appropriate, below.

351

352

353 **Results**

354 *Parentage analysis*

355 In total, pedigree reconstruction resulted in 617 cubs being assigned at least one parent (35% of

356 genotyped cubs included in the analyses), representing 29 cohorts and 6 generations (see Figure S1 for

357 visual representation). Out of these, 556 (89%) cubs were assigned both parents, while 23 (4%) were

358 assigned only a mother and 40 (7%) only a father. Overall, the 1,175 parental relationships (579 maternities

359 and 596 paternities) were represented by 239 fathers and 278 mothers. Among these, half-sibship sizes

360 (mean ±SD) varied from 1-11 (2.08 ±1.53) for mothers and 1-14 (2.49 ± 2.37) for fathers, with a total of 638

361 maternal and 1113 paternal sibships out of which 186 were full sibships. Additionally, 189 and 191

362 maternal grandmaternal and -paternal, as well as 155 and 161 paternal grandmaternal and -paternal links

363 were present. Based on successful maternal assignments, mean litter size was 1.24 (range 1-3), which is

364 slightly lower than previous reports for this and other populations (1.4-1.5; Carpenter et al. 2005; Dugdale

365 et al. 2007; Annavi et al. 2014). Out of 101 litters of more than one cub, 23% (compared to a previous

366 estimate of 16%; Carpenter et al. 2005) were multiple paternity litters, comprising 18 litters of n=2 and 4 of

367 n=3 contributed to by two different fathers, and one of n=3 with each cub assigned a different father.

368 Parent-offspring assignments covered 37 social groups out of the 45 represented in the full database. Based

369 on the parent-offspring assignments made, the mean rate of extra-group paternity over the 29 years was

370 37.1% (SD ±18.4). The relatively small proportion of assignments likely reflects the lack of strong prior

371 information on maternity in badgers. Certainly, this greatly reduces power, and so the number of

372 assignments, relative to paternity assignment when the mother is already known (Jones et al. 2010).

373 Incomplete sampling of candidate parents is likely to be another contributing factor. The number of

14

374   unsampled candidate parents estimated by MasterBayes varies considerably between cohorts with a

375   median (range) of 0.819 (0.359-0.628) females per group, and 20.4 (5.13-239) males in the whole study

376   area (Table S9). Out of the total parent-offspring assignments accepted at ≥80% confidence, 34% and 19%

377   were assigned with ≥90% and ≥95% confidence, respectively.

378

379   *Among-cohort variation in mean annual paternity distance*

380   Across the 23 cohorts for which spatial data could be included in the parentage assignment, point

381   estimates of PD_c obtained as the mean of the posterior distributions for each cohort varied from 173 m

382   (95% CI, 93-275 m) to 608 m (95% CI, 270-1249 m) with a mean of 354 m (SE ±19.6) across cohorts. Despite

383   relatively high uncertainty around some annual estimates, non-overlapping credible intervals for some

384   pairwise comparisons indicate significant annual variation in PD_c (Figure 1a). However, this variation was

385   not related to any of the explanatory variables (population size, sex ratio or year treated as a continuous

386   variable to characterise any trend) tested in our multiple regression model (Table 1).

387

388   *Among-individual and among-group variation in paternity distance*

389   Our mixed model analysis of PD_i indicated no significant effects of parental age, weight or group

390   size (neither maternal nor paternal variables; Table 2). Maternal social group sex ratio, on the other hand,

391   had a significant negative effect on paternity distance (Table 2), indicating that cubs from maternal social

392   groups (i.e. cub's natal group) with a higher proportion of males have lower paternity distances on average.

393   Paternal social group sex ratio showed the opposite trend, but the effect was not significant ($p>0.05$).

394   Testing the random effects provided evidence of significant among-individual variation in PD_i for both

395   mothers (among-mother repeatability, denoted $R_M = 0.16$ SE ±0.05, $\chi^2=40.29$, $p<0.001$) and fathers (among-

396   father repeatability, denoted $R_P = 0.2$ SE ±0.06, $\chi^2=35.82$, $p<0.001$) (see Figure 2). Comparison of the full

397   model fit to one in which maternal and paternal identity variance components were constrained to be

398   equal provided no significant evidence against the null hypothesis that mother and father explain equal

399   variance in cub PD_i ($\chi^2=0.38$, $p=0.5$). The random effect of year was estimated at c. 1% of the variance and

400   was not significant.

401   Parental social group identities also explained significant variation in PD_i, with group level

402   repeatabilities of $R_{MSG}=0.25$ (SE ±0.05; $\chi^2=58.2$, $p<0.001$) and $R_{PSG}=0.38$ (SE ±0.06; $\chi^2=64.5$, $p<0.001$), where

403   MSG refers to maternal, and PSG to paternal social group (Figure 2). The difference in the proportion of

404   variance in PD_i explained by PSG compared to that of MSG was marginally non-significant ($\chi^2=3.43$, $p=0.06$).

405   There was a strong negative covariance between maternal and paternal group identity effects, which

406   corresponds to a correlation (±SE) of $r_{MSG.PSG} = -0.99$ (±0.03; $\chi^2=39.3$ $p<0.001$; Figure 3c). Thus, social groups

407   in which resident females (males) are more likely to mate with males (females) from further away are the

408   same groups in which resident males (females) are less likely to mate with females (males) from further

409   away. To visualise this pattern better, and the among-group variation in PD_i generally, we extracted the

410   group level random effect predictions (best linear unbiased predictors or, BLUPs, see Table S2), which

411   represent the predicted deviation of each (maternal and paternal) social group from the mean paternity

412   distance, and overlaid them on a spatial map of the study area (Figure 3). This confirms that PSG with

413   longer-than-average paternity distances, correspond to MSG with shorter-than-average paternity

414   distances. Biologically, this is consistent with source-sink dynamics where some groups both retain resident

415   male genes as well as attracting extra-group paternity, however, under the current methodology it is not

416   possible to discern whether it is primarily driven by physical movement of males, females, or both. Note

417   that while the sources of among-group variation are unknown, we highlight that estimates here are

418   conditioned on group size and sex ratio, the latter having some effects as described above.

419

420

421   *Among-individual and among-group variation in extra-group paternity*

422   Analysis of EGP_i yielded broadly similar insights to our model of PD_i, although paternal, as well as

423   maternal, social group sex ratio had significant effects on extra-group paternity (Table 2). Similar to PD_i, the

424   effect was negative for maternal, and positive for paternal group sex ratio. Thus, there is lower extra-group

425   paternity among offspring in groups with higher male to female ratios. Other fixed effects were non-

426   significant (Table 2). Maternal and paternal ID had significant repeatabilities ($R_M = 0.15$ ±0.04, $\chi^2=40.61$,

427   $p<0.001$; $R_P = 0.17$±0.04, $\chi^2=35.34$, $p<0.001$) indicating consistent differences among individuals of both

16

428   sexes in their tendency to have offspring with extra-group partners (Figure 2). Social group level effects

429   were also significant and again almost perfectly negatively correlated ($r_{MSG.PSG}$ = -0.99 SE ±0.03; Table 3,

430   Figure 3). Differences in the amount of variance explained by maternal versus paternal identity, and MSG

431   versus PSG were not significant, while year explained only a small (and non-significant) amount of variance

432   in EGP_i (Table 3).

433

434

435   **Discussion**

436   We examined variation in breeding excursions using pedigree-derived information on extra-group

437   paternity and paternity distance in a wild population of badgers. We found evidence that cohort mean

438   paternity distance (PD_c, the mean distance between the social groups of fathers and their cubs) varied

439   among years. Contrary to our predictions, this among-cohort variation in PD_c was not explained by annual

440   variation in population size or sex ratio, nor did we see any systematic temporal trend in paternity distance

441   over the study period. However, individual (cub) level analyses showed significant among-parent (both

442   mother and father) and among-social group variance in breeding excursions, with the latter contributed to

443   (but not fully explained) by differences in group sex ratios. Below we discuss these findings in the context of

444   the wider literature, focusing on their implications for ecological and evolutionary dynamics.

445

446

447   *Among-cohort variation in average paternity distance*

448   Our point estimates of PD_c varied considerably among years, suggesting temporal variation in the

449   tendency of badgers to undertake breeding excursions. However, there was no systematic trend over time

450   and cohort variation was not explained by changes in the size or sex ratio of the Woodchester Park

451   population as a whole. A post hoc analysis of PD_i and EGP_i with population-level estimates included as

452   additional predictors also revealed no significant effects of population size or sex ratio. Year-to-year

453   variation in PD_c therefore remains unexplained at present, but could plausibly be linked to other variables

454   such as weather conditions, relatedness and neighbouring group composition, all of which are known to

455   influence movement, activity and dispersal in badgers (Annavi et al., 2014; Noonan et al., 2014), but which

456   were not investigated here. More generally, the absence of population size effects on PD_c contrasts

457   somewhat with previous studies. In badgers and other species (e.g. Møller 1991; Mougeot 2004; Annavi et

458   al. 2014), local density-dependence has been reported in rates of extra-group paternity – a pattern often

459   linked to changes in mate guarding behaviour (e.g. Møller 1991; Kokko & Rankin 2006; Isvaran & Clutton-

460   Brock 2007), though evidence for mate guarding in badgers is limited (Dugdale et al. 2007). Variation in

461   movement distance has also been linked to population density in badgers (Frantz et al. 2010; Byrne et al.

462   2014), and is sensitive to local density reductions from culling (Tuyttens et al. 2000a Tuyttens et al. 2000b;

463   Pope et al. 2007). However, we note that paternity distance is considered a proxy for movements relating

464   specifically to breeding excursions here. Certainly, the processes governing rates of breeding excursions

465   may differ from those influencing other types of movement making direct comparisons difficult.

466   There are also several other explanations for the apparent discrepancy between our results and

467   these previous findings. Firstly, it is possible that among-year density variation in the current study is not

468   sufficient to reveal a density-dependent response, as Woodchester Park has one of the highest recorded

469   densities (25 adults/km^2) of badgers throughout the species' range (Rogers et al. 1997) and the habitat may

470   be saturated. However, population fluctuation over the period of this study suggests this is not the case, as

471   population size increased in some years. Second, it is possible that the (overall) population density measure

472   used here doesn't capture variation at the correct scale to reveal density-dependence. The latter appears

473   to be the case for sex ratio, with temporal variation in population level PD_c not being predicted by

474   population sex ratio, but local (i.e. group) sex ratios contributing to spatial variation in EGP_i and PD_i defined

475   at individual (cub) level (discussed further below). However, parallel local density effects (modelled as

476   social group size effects) did not contribute to spatial variation in either EGP_i or PD_i. An additional

477   consideration is the fact that the lack of a clear density-dependent pattern could conceivably be an artefact

478   of the study scale, as high-density populations (such as Woodchester Park) typically involve sampling over

479   smaller spatial areas and may therefore miss longer distance movement (Byrne et al. 2014). Finally, we

480   note that the large proportion of unresolved parentage across the study period, as indicated by the

481 relatively low number of parentage assignments (35% cubs assigned parent(s)), may well have resulted in a

482 lack of power to distinguish density and sex ratio effects on cohort mean paternity distance.

483

484 *Among-group variation in cub PD_i and EGP_i*

485   Analysis of cub level proxies of (parental) breeding excursions revealed several important sources

486 of variation. Parental social group sex ratios influenced both EGP_i and PD_i. Although we note that the effect

487 of PSG sex ratio on PD_i was not statistically significant in the main analysis presented, it was significant

488 when we refitted our model using only those paternity distances inferred from assignments at the 95%

489 confidence threshold (see Table S6). Cubs had higher PD_i (on average) and were more likely to have an

490 extra-group father if born into less male-biased social groups. Conversely, cubs born in groups with more

491 male-biased sex ratios were more likely to be fathered by within-group males. These results are consistent

492 with earlier analysis of trapping data in Woodchester Park in which Rogers et al. (1998) concluded that

493 males preferentially move to groups with a higher proportion of females. Woodroffe et al. (1993) also

494 found that the peak of these temporary excursions coincides, for both males and females, with female

495 oestrus while in the Wytham Wood (Oxfordshire, UK) badger population, while, similar to Woodchester

496 Park, higher numbers of within-group males were associated with lower rates of EGP (Annavi et al., 2014).

497 Taken together, these results are consistent with ongoing mate guarding by males (anti-kleptogamy

498 hypothesis; Robertson et al. 2014) although they do not provide direct evidence. Although previous studies

499 have thus emphasised the role of males in breeding excursions, we stress that our indirect inferences from

500 paternity distance and extra-group paternity do not allow us to discriminate between male and female

501 movements. Temporary excursions by both sexes are possible and our results could reflect important

502 variation in female mating behavior in response to mate availability. For instance, females may be less

503 inclined to seek extra-group matings in male-biased groups if they have greater choice of partners.

504 Nevertheless, the relative importance of contributing factors (e.g. avoidance of male-male competition,

505 female choice for extra-group males, inbreeding avoidance by either sex) is not clear (although see Annavi

506 et al. 2014).

507   After accounting for sex ratio (and group size) effects, parental social group identities together

508   account for more of the remaining variance in cub PD_i and EGP_i (63% and 49%, respectively) than any other

509   variance component. Further, the strong negative correlation between maternal and paternal group

510   identity effects in both models indicates that maternal groups that predispose to high paternity distance

511   are the same as the paternal groups predisposed to low paternity distance. These social group identity

512   effects are not readily explained as a simple consequence of, for example, (relative) distances between

513   groups or edge-effects. In the former case, a positive correlation between maternal and paternal social

514   groups would be present, while, in the latter, groups at the edges of the study area would be expected to

515   have below average PD_i. This is because we expect failure to assign paternity to cubs sired by unsampled

516   males from outside the study area, such that edge effects are likely to cause downward bias in average PD_i

517   and EGP_i for peripheral maternal groups. However, no such pattern is readily apparent in our analysis (see

518   spatial maps of group effects on cub paternity distance in Figure 3).

519   Thus, while reiterating the earlier caveat that some long-distance movements may be missed by

520   our analysis, among-group variation in cub paternity distance is not readily explained as an artefact here.

521   Rather the emerging picture is one of source-sink dynamics, where some social groups are more 'attractive'

522   than others thus both retaining and drawing in male genes. From the male's point of view this could signal

523   variation in some unknown aspect of "quality" among females from different social groups, which itself

524   may be mediated by spatial variation in resource availability (e.g. food, setts) that determine habitat

525   preferences of females. Conversely, the observed pattern could reflect variation in female mating

526   preferences if 'attractive' males are spatially clustered. Spatial variation in habitat quality has previously

527   been linked to differences in group size across Woodchester Park (Delahay et al. 2006) and is certainly a

528   plausible hypothesis for explaining among-group differences 'attractiveness', although variance explained

529   by parental social group identities is estimated here conditional on a set of fixed effects including group

530   size. Furthermore, group size itself was not a significant predictor of either response variable in the main

531   analyses presented based on parentage assignments made at 80% confidence. However, using the more

532   stringent assignments threshold of 95%, group sizes did have a significant effect. Given statistical support

533   for group size effects is thus rather equivocal we draw no strong conclusions about its role. However, at

20

534 least in a qualitative sense it is worth pointing out that PD_i and EGP_i seem to increase with paternal group

535 size and decrease with maternal group size.

536 Similar variation has been recorded in great cormorants (*Phalacrocorax carbo sinensis*), where

537 Minias et al. (2016) found higher rates of extra-pair paternity in the periphery than in the centre of a

538 nesting colony. This pattern was not explained by density but by variation in mate quality, as indicated by

539 nest site location. Habitat structure has also been shown to influence rates of extra-pair paternity, for

540 instance, in blue-footed boobies (*Sula nebouxii*), by restricting movements within the colony (Ramos et al.,

541 2014). Although our results, as well as results from previous studies (Carpenter et al., 2005; Rogers et al.,

542 1998), suggest that movement in this population is focused around neighbouring social groups, with an

543 average PD_C of 358 m and a nearest neighbour distance between social group main setts of 355 m (SD 84)

544 m, habitat structure *per se* is unlikely to influence movement in this population, spatial structuring

545 (particularly of females) instead being mediated by resource availability (da Silva et al., 1994; Delahay et al.

546 2006).

547

548 *Among-individual variation in cub PD_i and EGP_i*

549 In addition to social group effects, we found that there was repeatable variation among both

550 mothers and fathers for cub PD_i and EGP_i. The most parsimonious interpretation of these results is that

551 there is among-individual variation, in both sexes, for breeding behavior. This interpretation is in line with

552 trapping-based inferences for the Woodchester Park badger population (Rogers et al., 1998), as well as

553 studies of other taxa. For instance, Whittingham et al. (2006) found the proportion of extra-pair young

554 produced to be highly repeatable for female tree swallows (*Tachycineta bicolor;* intra-class correlation, $r=$

555 0.83). In coal tits (*Parus ater*), the proportion of extra-pair young showed repeatability in both sexes among

556 the same social pairing ($r=0.33$ and 0.47 for males and females respectively; Dietrich et al. 2004).

557 Conversely, breeding excursions were found not to be a repeatable behaviour in female roe deer

558 (*Capreolus capreolus*; Debeffe et al. 2014). Among-individual differences in other dispersal and exploratory

559 behaviours have also been recorded for spiders (Bonte et al. 2009; Johnson et al. 2015), fish (Harrison et al.

560 2015), amphibians (Cosentino & Droney, 2016) and birds (Reid et al. 2011a; Patrick et al. 2012; Grist et al.

561   2014). Thus, among-individual variance in PD_i and EGP_i could be linked to both reproductive decision

562   making (i.e., individuals varying in their propensity/ability to seek or obtain extra-group matings), and more

563   general exploratory traits influencing encounter rates between badgers from different groups. Regardless,

564   a further aspect of our analysis worth noting is that similar levels of variation in cub PD_i and EGP_i were

565   explained by maternal and paternal identities. Thus, whether gene flow from breeding excursions is being

566   mediated primarily by variation in movement *per se,* or by reproductive decision making, both sexes appear

567   to have an equal impact.

568   Our analyses have not clearly identified the underlying source(s) of among-individual variance in

569   (parental) mating behaviour. Neither size nor age (of either parent) significantly predict PD_i and EGP_i in the

570   main analyses, although we note that using the 95% confidence pedigree the positive effects of paternal

571   age on both response variables are statistically significant (Table S4). This suggests that older males tend to

572   produce more extra-group offspring and make longer breeding excursions (or mate with females that do),

573   though this conclusion remains tentative. In a broader sense, among-individual variation will reflect the fact

574   that individuals experience different environmental conditions (e.g. maternal effects, food availability,

575   social status) even within groups and years (which were both modelled separately), although genetic

576   variation may also be present. Dispersal distance has been shown to be heritable in a free-living population

577   of great tits (*Parus major*; h^2= 0.15 SE ± 0.006; Korsten et al. 2013), as has EGP rate in in female, but not

578   male, song sparrows (*Melospiza melodia*; Reid et al. 2011a&b). It is, therefore, possible that the among-

579   individual variance found here has a partial genetic basis. In fact, the pedigree will facilitate testing this,

580   although it would best be achieved through quantitative genetic modelling of independently obtained

581   trapping data.

582

583   **Conclusions**

584   We have used a genetic pedigree to characterise variation in paternity distance and extra-group

585   paternity in a high-density badger population. We show there to be variation among years and social

586   groups, but also among-parental individuals (both mothers and fathers) within groups. Although effects of

587   social group sex ratio (and potentially group size and paternal age) were detected, in general this variation

588   is not readily explained by life-history and social correlates. Among-group variation appears to follow a

589   pattern of source-sink dynamics, suggesting that some social groups are more attractive to extra-group

590   partners than others, though levels of among-parental variation in our metrics were similar across the

591   sexes. Not readily explained by age or body size, it is possible that genes as well as individual-specific

592   (rather than group level) environmental factors contribute to among-individual variation although this

593   remains to be tested. Individual-level differences can have important consequences for many ecological

594   and evolutionary processes, and our results highlight the fact that individuals can vary consistently in their

595   mating behavior. Together these results emphasise the importance of including individual-level variation in

596   evolutionary models of animal movement and mating behavior, as well as management and conservation

597   measures.

598

599

600

601

602

603

604

605

606

607

608

609

610

611

612

613

Data accessibility Analyses reported in this article can be reproduced using the data provided by Marjamäki

et al. (2018).

Author contributions A.J.W., R.A.M., R.D. and P.H.M. designed research; P.H.M performed research; T.B.

provided facilities and support for the molecular work, D.A.D. assisted with genotyping and data validation;

R.D. contributed data; P.H.M. analysed data with A.J.W. and H.L.D; P.H.M. lead the write-up with input

from all authors.

643 **References**

644 Annavi, G., Newman, C., Dugdale, H. L., Buesching, C. D., Sin, Y. W., Burke, T., & Macdonald, D. W. (2014).

645 Neighbouring-group composition and within-group relatedness drive extra-group paternity rate in the

646 European badger (*Meles meles*). *Journal of Evolutionary Biology*, 27(10), 2191–203. doi: 10.1111/jeb.12473

647 Beirinckx, K., Gossum, H. Van, Lajeunesse, M. J., & Forbes, M. R. (2006). Sex biases in dispersal and

648 philopatry: insights from a meta-analysis based on capture mark recapture studies of damselflies. *Oikos*,

649 113, 539–547. doi: 10.1111/j.2006.0030-1299.14391.x

650 Beirne, C., Delahay, R., Young, A., & Young, A. (2015). Sex differences in senescence: the role of intra-sexual

651 competition in early adulthood. *Proceedings of the Royal Society B*, 282. doi: 10.1098/rspb.2015.1086

652 Benjamini, Y. & Hochberg, Y. (1995). Controlling the False Discovery Rate: A practical and powerful

653 approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1), 289–

654 300. www.jstor.org/stable/2346101

655 Bonte, D., de Clercq, N., Zwertvaegher, I. (2009). Repeatability of dispersal behaviour in a common dwarf

656 spider: evidence for different mechanisms behind short- and long-distance dispersal. *Ecological*

657 *Entomology*, 34, 271–276. doi: 10.1111/j.1365-2311.2008.01070.x

658 Bowler, D. E., & Benton, T. G. (2005). Causes and consequences of animal dispersal strategies: relating

659 individual behaviour to spatial dynamics. *Biological Reviews of the Cambridge Philosophical Society*, 80(2),

660 205–225. doi: 10.1017/s1464793104006645

661 Bowler, D. E., & Benton, T. G. (2009). Variation in dispersal mortality and dispersal propensity among

662 individuals: The effects of age, sex and resource availability. *Journal of Animal Ecology*, 78(6), 1234–1241.

663 doi: 10.1111/j.1365-2656.2009.01580.x

664 Breton, A. R., Nisbet, I. C. T., Mostello, C. S., & Hatch, J. J. (2014). Age-dependent breeding dispersal and

665 adult survival within a metapopulation of Common Terns *Sterna hirundo. Ibis*, 156(3), 534–547. doi:

666 10.1111/ibi.12161

667   Byrne, A. W., Quinn, J. L., Keeffe, J. J. O., Green, S., Sleeman, D. P., Martin, S. W., & Davenport, J. (2014).

668   Large-scale movements in European badgers: has the tail of the movement kernel been underestimated?

669   *Journal of Animal Ecology*, 83, 991–1001. doi: 10.1111/1365-2656.12197

670   Carpenter, P. J., Dawson, D. a., Greig, C., Parham, A., Cheeseman, C. L., & Burke, T. (2003). Isolation of 39

671   polymorphic microsatellite loci and the development of a fluorescently labelled marker set for the Eurasian

672   badger (*Meles meles*) (Carnivora: Mustelidae). *Molecular Ecology Notes*, 3(4), 610–615. doi:

673   10.1046/j.1471-8286.2003.00529.x

674   Carpenter, P. J., Pope, L. C., Greig, C., Dawson, D. A, Rogers, L. M., Erven, K., Wilson, G. J., Delahay, R. J.,

675   Cheeseman, C. J., Burke, T. (2005). Mating system of the Eurasian badger, *Meles meles*, in a high density

676   population. *Molecular Ecology*, 14(1), 273–84. doi: 10.1111/j.1365-294X.2004.02401.x

677   Clarke, A. L., Sæther, B., & Røskaft, E. (1997). Sex Biases in Avian Dispersal: A Reappraisal. *Oikos*, 79(3), 429–

678   438. doi: 10.2307/3546885

679   Cosentino, B. J., & Droney, D. C. (2016). Movement behaviour of woodland salamanders is repeatable and

680   varies with forest age in a fragmented landscape. *Animal Behaviour*, 121, 137–146. doi:

681   10.1016/j.anbehav.2016.08.013

682   Cote, J., Clobert, J., Brodin, T., Fogarty, S., & Sih, A. (2010). Personality-dependent dispersal:

683   characterization, ontogeny and consequences for spatially structured populations. *Philosophical*

684   *Transactions of the Royal Society B*, 365, 4065–4076. doi: 10.1098/rstb.2010.0176

685   da Silva, J., Macdonald, D. W., & Evans, P. G. H. (1994). Net costs of group living in a solitary forager, the

686   Eurasian badger (*Meles meles*). *Behavioral Ecology*, 5(2), 151–158. doi: 10.1093/beheco/5.2.151

687   Dale, S., Lunde, A., & Steifetten, O. (2005). Longer breeding dispersal than natal dispersal in the ortolan

688   bunting. *Behavioral Ecology*, 16(1), 20–24. doi: 10.1093/beheco/arh129

689   Daniels, S. J., & Walters, J. R. (2000). Between-year breeding dispersal in red-cockaded woodpeckers:

690   Multiple causes and estimated cost. *Ecology*, 81, 2473–2484. doi: 10.2307/177468

691 de Leeuw, A. N. S., Forrester, G. J., Spyvee, P. D., Brash, M. G. I., Delahay, R. J. (2004). Experimental

692 comparison of ketamine with a combination of ketamine, butorphanol and medetomidine for general

693 anaesthesia ofthe Eurasian badger (*Meles meles* L.). *The Veterinary Journal*, 167, 186–193. doi:

694 10.1016/S1090-0233(03)00113-8

695 Debeffe, L., Focardi, S., Bonenfant, C., Hewison, A. J. M., Morellet, N., Vanpé, C., Heurich M., Kjellander, P.,

696 Linnell, J. D. C., Mysterud, A., Pellerin M., Sustr, P., Urbano, F., Cagnacci, F. (2014). A one night stand?

697 Reproductive excursions of female roe deer as a breeding dispersal tactic. *Oecologia*, 176(2), 431–443. doi:

698 10.1007/s00442-014-3021-8

699 Delahay, R. J., Carter, S. P., Forrester, G. J., Mitchell, A. & Cheeseman, C. L. (2006). Habitat correlates of

700 group size, bodyweight and reproductive performance in a high-density Eurasian badger (*Meles meles*)

701 population. Journal of Zoology, 270(3), 437–447. doi: 10.1111/j.1469-7998.2006.00165.x

702 Delahay, R., Brown, J., Mallinson, P., Spyvee, P., Handoll, D., Rogers, L., & Cheeseman, C. (2000a). The use

703 of marked bait in studies of the territo- rial organization of the European badger (*Meles meles*). *Mammal*

704 *Review*, 30, 73–87.

705 Delahay, R. J., Langton, S., Smith, G. C., Clifton-Hadley, R. S., & Cheeseman, C. L. (2000b). The spatio-

706 temporal distribution of *Mycobacterium bovis* (bovine tuberculosis) infection in a high-density badger

707 population. *Journal of Animal Ecology*, 69(3), 428–441. doi: 10.1046/j.1365-2656.2000.00406.x

708 Delahay, R.J., Walker, N., Smith, G.S., Wilkinson, D., Clifton-Hadley, R.S., Cheeseman, C.L. et al. (2013).

709 Long-term temporal trends and estimated transmission rates for *Mycobacterium bovis* infection in an

710 undisturbed high-density badger (*Meles meles*) population. Epidemiology & Infection, 141, 1445–1456. doi:

711 10.1017/S0950268813000721

712 Dietrich, V., Schmoll, T. I. M., & Winkel, W. (2004). Pair identity - an important factorconcerning variation in

713 extra-pair paternity in the coal tit (*Parus ater*). *Behaviour*, 141, 817–835. doi: 10.1163/1568539042265644

714 Dobson, B. Y. F. S. (1982). Competition for mates and predominant juvenile male dispersala in mammals.

715 *Animal Behaviour*, 30, 1183–1192. doi: 10.1016/S0003-3472(82)80209-1

716   Dugdale, H. L., Macdonald, D. W., Pope, L. C., & Burke, T. (2007). Polygynandry, extra-group paternity and

717   multiple-paternity litters in European badger (*Meles meles*) social groups. *Molecular Ecology*, 16(24), 5294–

718   306. doi: 10.1111/j.1365-294X.2007.03571.x

719   Dugdale HL, Ellwood SA, Macdonald DW (2010) Alloparental behaviour and long-term costs of mothers

720   tolerating other group-members in a plurally breeding mammal. *Animal Behaviour* 80, 721–735.

721   Frantz, A. C., Do Linh San, E., Pope, L. C., & Burke, T. (2010). Using genetic methods to investigate dispersal

722   in two badger (*Meles meles*) populations with different ecological characteristics. *Heredity*, 104(5), 493–

723   501. doi: 10.1111/j.1365-2656.2008.01415.x

724   Greenwood, P. J. (1980). Mating systems, philopatry and dispersal in birds and mammals. *Animal*

725   *Behaviour*, 28, 1140–1162. doi: 10.1016/S0003-3472(80)80103-5

726   Greenwood, P. J. and Harvey, P. H. (1982). The natal and breeding dispersal of birds. *Annual Review of*

727   *Ecology and Systematics*, 13, 1–21. doi: 10.1146/annurev.es.13.110182.000245

728   Griffith, Simon C., Owens, Ian P. F., Thuman, K. A. (2002). Extra pair paternity in birds: a review of

729   interspecific variation and adaptive function. *Molecular Ecology*, 11, 2195–2212. doi: 10.1046/j.1365-

730   294X.2002.01613.x

731   Grist, H., Daunt, F., Wanless, S., Nelson, E. J., Harris, M. P., Newell, M., Burthe, S., Reid, J. M. (2014). Site

732   fidelity and individual variation in winter location in partially migratory European shags. *PLoS ONE*, 9(6). doi:

733   10.1371/journal.pone.0098562

734   Hadfield, J.D. (2012) MasterBayes: Maximum Likelihood and Markov chain Monte Carlo methods for

735   pedigree reconstruction, analysis and simulation. Accessed 01/11/2014

736   https://cran.rproject.org/web/packages/MasterBayes/vignettes/Tutorial.pdf

737   Hadfield, J. D., Richardson, D. S., & Burke, T. (2006). Towards unbiased parentage assignment: combining

738   genetic, behavioural and spatial data in a Bayesian framework. *Molecular Ecology*, 15, 3715–3730. doi:

739   10.1111/j.1365-294X.2006.03050.x

740   Hamilton, W. D. (1990). Mate Choice Near or Far. *American Zoologist*, 30, 341–352. doi:

741   10.1093/icb/30.2.341

742   Harrison, P. M., Lee, F., Gutowsky, G., Martins, E. G., & Patterson, D. A. (2015). Personality-dependent

743   spatial ecology occurs independently from dispersal in wild burbot (*Lota lota*). *Behavioural Ecology*, 26,

744   483–492. doi: 10.1093/beheco/aru216

745   Holt, R. D. (2003). On the evolutionary ecology of species' ranges. *Evolutionary Ecology Research*, 5, 159–

746   178. doi: 10.1139/A09-014

747   Huck, M., Frantz, A. C., Dawson, D. A., Burke, T., & Roper, T. J. (2008). Low genetic variability, female-biased

748   dispersal and high movement rates in an urban population of Eurasian badgers *Meles meles*. *Journal of

749   Animal Ecology*, 77(5), 905–915. doi: 10.1111/j.1365-2656.2008.01415.x

750   Isvaran, K., & Clutton-Brock, T. (2007). Ecological correlates of extra-group paternity in mammals.

751   *Proceedings of the Royal Society B: Biological Sciences*, 274, 219–224. doi: 10.1098/rspb.2006.3723

752   Johnson, J. C., Halpin, R., Stevens, D., Vannan, A., Lam, J., & Bratsch, K. (2015). Individual variation in

753   ballooning dispersal by black widow spiderlings: The effects of family and social rearing. *Current Zoology*,

754   61(3), 520–528. doi: 10.1093/czoolo/61.3.520

755   Johnson, P. C. D., & Haydon, D. T. (2007). Maximum-likelihood estimation of allelic dropout and false allele

756   error rates from microsatellite genotypes in the absence of reference data. *Genetics*, 175(2), 827–842. doi:

757   10.1534/genetics.106.064618

758   Kentie, R., Both, C., Hooijmeijer, J. C. E. W., & Piersma, T. (2014). Age-dependent dispersal and habitat

759   choice in black-tailed godwits *Limosa limosa limosa* across a mosaic of traditional and modern grassland

760   habitats. *Journal of Avian Biology*, 45(4), 396–405. doi: 10.1111/jav.00273

761   Kenward, M. G. & Roger, J. H. (1997). The precision of fixed effects estimates from restricted maximum

762   likelihood. *Biometrics*, 53, 983–997.

763   Kokko, H. & Rankin, D. J., (2006). Lonely hearts or sex in the city? Density-dependent effects in mating

764   systems. *Phiosophical. Transactions of the Royal Society B*. 361, 319–334 doi:10.1098/rstb.2005.1784

765   Korsten, P., van Overveld, T., Adriaensen, F. & Matthysen, E. (2013). Genetic integration of local dispersal

766   and exploratory behaviour in a wild bird. *Nature Communications*, 4, 1–7. doi: 10.1038/ncomms3362

767   Kruuk, Hans & Parish, T. (1982). Factors affecting population density, group size and territory size of the

768   European badger, *Meles meles*. *Journal of Zoology*, 196, 31–39. doi: 10.1111/j.1469-7998.1982.tb03492.x

769   López-Giráldez, F., Marmi, J., Domingo-Roura, X. (2007). High incidence of nonslippage mechanisms

770   generating variability and complexity in Eurasian badger microsatellites. *Journal of Heredity*,98(6), 620-628.

771   doi: 10.1093/jhered/esm068

772   Macdonald, D. W., Newman, C., Buesching, C. D., & Johnson, P. J. (2008). Male-biased movement in a high-

773   density population of the Eurasian Badger (*Meles meles*). *Journal of Mammalogy*, 89(5), 1077–1086. doi:

774   10.1644/07-MAMM-A-185.1

775   Marjamäki, P. H., Contasti, A. L., Coulson, T. N., McLoughlin P. D. (2013). Local density and group size

776   interacts with age and sex to determine direction and rate of social dispersal in a polygynous mammal.

777   *Ecology and Evolution*, 3(9), 3073–3082. doi: 10.1002/ece3.694

778   Marjamäki, P. H., Dugdale, H. L.,Dawson, D. A., McDonald, R. A., Delahay, R., Burke, T., Wilson, A. J. (2018).

779   Data from: Individual variation and the source-sink group dynamics of extra-group paternity in a social

780   mammal. *Behavioral Ecology*. https://doi.org/10.5061/dryad.9n5m91v.

781

782   Marshall, T. C., Slate, J., Kruuk, L. E B, Pemberton, J. M. (1998). Statistical confidence for likelihood-

783   based paternity inference in natural populations. *Molecular Ecology*, 7(5), 639-655. doi:10.1046/j.1365-

784   294x.1998.00374.x

785   Matthysen, E. (2005). Density-dependent dispersal in birds and mammals. *Ecography*, 28, 403–416. doi:

786   10.1111/j.0906-7590.2005.04073.x

787 Minias, P., Wojczulanis-Jakubas, K., Rutkowski, R., Kaczmarek, K., & Janiszewski, T. (2016). Spatial patterns

788 of extra-pair paternity in a waterbird colony: separating the effects of nesting density and nest site location.

789 *Behavioral Ecology and Sociobiology*, 70(3), 369–376. doi: 10.1007/s00265-015-2056-0

790 Mougeot, F. (2004). Breeding density, cuckoldry risk and copulation behaviour during the fertile period in

791 raptors: a comparative analysis. *Animal Behaviour* 67, 1067–1076. 10.1016/j.anbehav.2003.10.011

792 Møller, A.P. (1991) Density-dependent Extra-pair Copulations in the Swallow *Hirundo rustica*. *Ethology*, 87,

793 316-329. 10.1111/j.1439-0310.1991.tb00255.x

794 Nakagawa, S., & Schielzeth, H. (2010). Repeatability for Gaussian and non-Gaussian data: a practical guide

795 for biologists. *Biological Reviews*, 85, 935–956. doi: 10.1111/j.1469-185X.2010.00141.x

796 Noonan, M. J., Markham, A., Newman, C., Trigoni, N., Buesching, C. D., Ellwood, S. a., & Macdonald, D. W.

797 (2014). Climate and the individual: Inter-annual variation in the autumnal activity of the European badger

798 (*Meles meles*). *PLoS ONE*, 9(1), 1–9. doi: 10.1371/journal.pone.0083156

799 Nielsen, J. F., English, Goodall-Copestake, W. P., Wang, J., Walling, C., Bateman, A. W., Flower, T. P.,

800 Sutcliffe, R. L., Samson, J., Thavarajah, N. K., Kruuk, L. E. B., Clutton-Brock, T. H., Pemberton, J. M. (2012).

801 Inbreeding and inbreeding depression of early life traits in a cooperative mammal. *Molecular Ecology*, 21,

802 11, 2788-804.

803 Nowicki, P., & Vrabec, V. (2011). Evidence for positive density-dependent emigration in butterfly

804 metapopulations. *Oecologia*, 167, 657–665. doi: 10.1007/s00442-011-2025-x

805 Patrick, S. C., Chapman, J. R., Dugdale, H. L., Quinn, J. L., & Sheldon, B. C. (2012). Promiscuity, paternity and

806 personality in the great tit. *Proceedings of the Royal Society B*, 279(1734), 1724–30. doi:

807 10.1098/rspb.2011.1820

808 Peig J. & Green A.J. (2009). New perspectives for estimating body condition from mass/length data: the

809 scaled mass index as an alternative method. *Oikos*, 118, 1883–1891. doi: 10.1111/j.1600-

810 0706.2009.17643.x

811   Peig, J., & Green, A. J. (2010). The paradigm of body condition: a critical reappraisal of current methods

812   based on mass and length. *Functional Ecology*, 24, 1323–1332. doi: 10.1111/j.1365-2435.2010.01751.x

813   Pope, L. C., Butlin, R. K., Wilson, G. J., Woodroffe, R., Erven, K., Conyers, C. M., Franklin, T., Delahay, R. J.,

814   Cheeseman, C. L., Burke, T. (2007). Genetic evidence that culling increases badger movement: implications

815   for the spread of bovine tuberculosis. *Molecular Ecology*, 16(23), 4919–29. doi: 10.1111/j.1365-

816   294X.2007.03553.x

817   R Core Team (2016). R: A language and environment for statistical computing. R Foundation for

818   Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

819   Rabasa, S. G., & Gutie, D. (2007). Metapopulation structure and habitat quality in modelling dispersal in the

820   butterfly *Iolana iolas*. *Oikos*, 116, 793–806. doi: 10.1111/j.2007.0030-1299.15788.x

821   Ramos, A. G., Nunziata, S. O., Lance, S. L., Rodríguez, C., Faircloth, B. C., Adair, P., & Drummond, H. (2014).

822   Habitat structure and colony structure constrain extrapair paternity in a colonial bird. *Animal Behaviour*, 95,

823   121–127. doi: 10.1016/j.anbehav.2014.07.003

824   Reid, J. M., Arcese, P., Sardell, R. J., & Keller, L. F. (2011a). Additive genetic variance, heritability , and

825   inbreeding depression in male extra-pair reproductive success. *American Naturalist*, 177(2), 177–187. doi:

826   10.1086/657977

827   Reid, J. M., Arcese, P., Sardell, R. J., & Keller, L. F. (2011b). Heritability of female extra-pair paternity rate in

828   song sparrows (*Melospiza melodia*). *Proceedings of the Royal Society B*, 278, 1114–1120. doi:

829   10.1098/rspb.2010.1704

830   Richardson, D. S., Jury, F. L., Blaakmeer, K., Komdeur, J., & Burke, T. (2001). Parentage assignment and

831   extra-group paternity in a cooperative breeder: The Seychelles warbler (*Acrocephalus sechellensis*).

832   *Molecular Ecology*, 10(9), 2263–2273. doi: 10.1046/j.0962-1083.2001.01355.x

833   Robertson, A., Palphramand, K. L., Carter, S. P., & Delahay, R. J. (2014). Group size correlates with territory

834   size in European badgers: implications for the resource dispersion hypothesis? *Oikos*, 124(4), 507-514. doi:

835   10.1111/oik.01459

836   Rogers, L. M., & Cheeseman, C L, Mallison, P. J. (1997). The demography of a high-density badger (*Meles*

837   *meles*) population in the west of England. *Journal of Zoology*, 242, 705–728. doi: 10.1111/j.1469-

838   7998.1997.tb05821.x

839   Rogers, L. M., Delahay, R., Cheeseman, C. L., Langton, S., Smith, G. C., & Clifton-Hadley, R. S. (1998).

840   Movement of badgers (*Meles meles*) in a high-density population: individual, population and disease

841   effects. *Proceedings of the Royal Society B*, 265(1403), 1269–1276. doi: 10.1098/rspb.1998.0429

842   Ronce, O. (2007). How does it feel to be like a rolling stone? Ten questions about dispersal evolution.

843   *Annual Review of Ecology, Evolution, and Systematics*, 38, 231–253. doi:

844   10.1146/annurev.ecolsys.38.091206.095611

845   Roper, T.J. (2010). Badger. HarperCollins, London.

846   Thomas, C. D. (2000). Dispersal and extinction in fragmented landscapes. *Proceedings of the Royal Society*

847   *B*, 267, 139–145. doi: 10.1098/rspb.2000.0978

848   Tuyttens, F. A. M., Macdonald, D. W., & Long, B. (2000a). Spatial perturbation caused by a badger (*Meles*

849   *meles*) culling operation: implications for the function of territoriality and the control of bovine tuberculosis

850   (*Mycobacterium bovis*). *Journal of Animal Ecology*, 69(69), 815–828. doi: 10.1046/j.1365-

851   2656.2000.00437.x

852   Tuyttens, F. A. M., Macdonald, D. W., Rogers, L. M., Cheeseman, C. L., & Roddam, A. W. (2000b).

853   Comparative study on the consequences of culling badgers (*Meles meles*) on biometrics, population

854   dynamics and movement. *Journal of Animal Ecology*, 69, 567–580. doi: 10.1046/j.1365-2656.2000.00419.x

855   Vicente, J., Delahay, R. J., Walker, N. J., Cheeseman, C. L. (2007). Social organization and movement

856   influence the incidence of bovine tuberculosis in an undisturbed high-density badger *Meles meles*

857   population. *Journal of Animal Ecology*, 76, 348-360. doi: 10.1016/j.anbehav.2015.05.020

858   Walling, C. a, Pemberton, J. M., Hadfield, J. D., & Kruuk, L. E. B. (2010). Comparing parentage inference

859   software: reanalysis of a red deer pedigree. *Molecular Ecology*, 19(9), 1914–28. doi: 10.1111/j.1365-

860   294X.2010.04604.x

861   Weiß, B. M., Kulik, L., Ruiz-Lambides, A. V., Widdig, (2016). Individual dispersal decisions affect fitness via

862   maternal rank effects in male rhesus macaques. *Scientific Reports*, 6, 32212. doi: 10.1038/srep32212

863   Wilson A. J., Gislason, D., Skulason, S., Snorrason, S., Adams, C. E., Alexander, G., Danzmann, R. G.,

864   Ferguson, M. M. (2004). Population genetic structure of Arctic Charr, *Salvelinus alpinus*, from northwest

865   Europe on large and small spatial scales. *Molecular Ecology*, 13, 1129–1142. doi: 10.1111/j.1365-

866   294X.2004.02149.x

867   White, G. C. & Burnham, K. P. (1999). Program MARK: survival estimation from populations of marked

868   animals. *Bird Study*, 46, 120–139. doi: 10.1080/00063659909477239

869   Whittingham, L. A., Dunn, Peter O., Stapleton, M. K. (2006). Repeatability of extra-pair mating in tree

870   swallows. *Molecular Ecology*, 15, 841-849. doi: 10.1111/j.1365-294X.2006.02808.x

871   Woodroffe, R, Macdonald, D W, da Silva J (1993). Dispersal and philopatry in the European badger, *Meles*

872   *meles*. *Journal of Zoology*. 237, 227-239. doi: 10.1111/j.1469-7998.1995.tb02760.x.

873   Yamaguchi, Nobuyuki, Dugdale, Hannah L., MacDonald, D. W. (2006). Female receptivity, embryonic

874   diapause, and superfetation in the european badger (*Meles meles*): Implications for the reproductive tactics

875   of males and females. *The Quarterly Review of Biology*, 81(1), 33–48. doi: 10.1086/503923

876   Zeh, J. A. & Zeh, D. W. (1996). The evolution of polyandry I: intragenomic conflict and genetic

877   incompatibility. *Proceedings of the Royal Society B*, 263, 1711–1717. doi: 10.1098/rspb.1996.0250

878
879
880
881
882
883

884

885

886

887

888

889

890

Figure 1. **Top:** Annual modal paternity distance (PD_c) estimated for each of 23 cohorts by MasterBayes (Hadfield et al. 2006) during pedigree reconstruction. Lines represent 95% credible intervals. Numbers above points represent the number of cubs assigned parentage in each year. **Bottom:** Total population size and number of males and females estimated in program MARK for each year of the study, based on 20 core social groups with consistent capture records. Bars represent standard errors.

896

Figure 2. Estimated intra-class correlations (i.e. proportion of total phenotypic variance calculated by dividing each component by the sum of all variance components) for each random effect in models of PD_i and EGP_i. Bars represent standard errors. M and P denote maternal and paternal individuals, while MSG and PSG denote the corresponding maternal and paternal social groups.

901

902

Figure 3. Spatial representation of **a)** maternal and **b)** paternal social group effects and **c)** the relationship between them. Effects are predicted from the mixed model of log-transformed PD_i (see main text) using best linear unbiased prediction (BLUP) while the spatial configuration of social group territories illustrated is derived from a bait marking survey in 1993 (when the maximum number of social groups were present). Six social groups included in current analyses are not shown on panels a) or b) due to missing bait-marking data, while grey shaded territories correspond to groups with no parentage assigned. Error bars in panel c) denote ± standard error and the regression line (red) slope is calculated directly from the model (co)variance estimates as $COV_{MSG.PSG}/V_{MSG.}$ MSG and PSG denote maternal and paternal social groups.

911

912

913

914

915

916

917

918

919

920

921

922

923

924

925

926

927

928 Table 1. Estimated effects of population size, sex ratio and cohort (year) on modal annual paternity
929 distance (PD_c). Estimates are from multiple regression with uncertainty integrated over the full posteriors
930 of annual PD_c (see main text). Predictors were mean centred for analysis.

	Estimate	95% credible interval
Intercept	332.43	319.90- 382.60
Population size[†]	0.36	-0.67 – 1.15
Sex ratio[‡]	-331.43	-1706.30 – 1743.66
Year	0.44	-7.81 – 4.74

931

932 [†] annual estimate of the number of badgers in Woodchester Park, based on 20 "core" social groups with consistent
933 capture records
934 [‡] calculated from annual population size estimates as the number of males divided by total population
935

936

937

938

939

940

941

942

943

944

945

946

947

948

949

	Log(PD$_i$)				EGP$_i$			
	Estimate (SE)	DF	F	P	Estimate (SE)	DF	F	P
Intercept	-0.72 (0.15)	1, 214.7	24.56	**<0.001**	0.72 (0.14)	1, 226.9	74.97	**<0.001**
Age$_M$	-0.45 (0.15)	1, 533.1	0.09	0.76	-0.52 (0.15)	1, 534.0	0.12	0.73
Body mass$_M$†	-0.61 (0.13)	1, 302.8	0.22	0.63	-0.66 (0.13)	1, 304.1	0.26	0.61
Group_size$_{MSG}$	0.94 (0.18)	1, 456.9	0.28	0.59	0.96 (0.18)	1, 443.0	0.29	0.59
Sex_ratio$_{MSG}$‡	-0.74 (0.22)	1, 531.5	10.97	**<0.001**	-0.82 (0.22)	1, 524.2	13.55	**<0.001**
Age$_P$	0.28 (0.2)	1, 516.7	2.11	0.15	0.30 (0.2)	1, 517.3	2.4	0.12
Body mass$_P$†	-0.59 (0.12)	1, 213.4	0.25	0.62	-0.56 (0.19)	1, 215.0	0.23	0.64
Group.Size$_{PSG}$	-0.12 (0.18)	1, 537.4	0.44	0.50	-0.12 (0.18)	1, 531.9	0.43	0.51
Sex_ratio$_{PSG}$‡	0.43 (0.24)	1, 538.1	3.21	0.08	0.50 (0.24)	1, 536.0	4.48	**0.04**

Table 2. Estimated fixed effect coefficients (standard error) and Wald F-tests from mixed models of log-
transformed PD$_i$ and EGP$_i$ (see main text for details). Response variables were standardised into standard
deviation units (SDU) prior to analysis. M and P denote maternal and paternal individuals, while MSG and
PSG denote the corresponding maternal and paternal social groups. DF stands for degrees of freedom.

Full models fitted for each response were y ~ μ + Age$_M$ + Body_Mass$_M$ + Group_size$_{MSG}$ + Sex_ratio$_{MSG}$ + Age$_P$ +
Body_Mass$_P$ + Group_size$_{PSG}$ + Sex_ratio$_{PSG}$ + *M* + *P* + *MSG* + *PSG* + *Year* where italic font denotes random effects and y
is either log(PD$_i$) or EGP$_i$

† mean body mass for parental individuals with multiple weight measurements within year of cub's birth

‡calculated as number of males divided by group size where group size is males plus females

974

975

976

977

978

979 Table 3. Estimated (co)variance components (standard error) associated with random effects in mixed
980 models of EGP_i and log-transformed PD_i. Statistical inference of random effects is by likelihood ratio test
981 results (see main text for details). M and P denote maternal and paternal individuals, while MSG and PSG
982 denote the corresponding maternal and paternal social groups.

983

	$\log(PD_i)$				EGP_i			
	Variance (SE)	df	χ^2_1	P	Variance (SE)	χ^2_1	df	P
V_{year}	0.02 (0.02)	1	3.22	0.07	0.02 (0.03)	2.83	1	0.09
$V_M{}^\dagger$	0.26 (0.05)	1	40.29	**<0.001**	0.26 (0.06)	40.61	1	**<0.001**
$V_P{}^\dagger$	0.31 (0.06)	1	35.82	**<0.001**	0.31 (0.06)	35.34	1	**<0.001**
$V_{MSG}{}^\ddagger$	0.39 (0.15)	2	58.16	**<0.001**	0.34 (0.13)	55.00	2	**<0.001**
$V_{PSG}{}^\ddagger$	0.59 (0.21)	2	64.54	**<0.001**	0.54 (0.19)	62.91	2	**<0.001**
$COV_{MSG,PSG}$	*-0.48 (0.17)*	*1*	*39.33*	***<0.001***	*-0.43 (0.15)*	*36.84*	*1*	***<0.001***
V_R	0.32 (0.04)	-	-	-	0.32 (0.04)	-	-	-

984 † not significantly different from each other (logLRT, PD_i: $\chi^2 = 0.38$, p=0.5 EGP_i: $\chi^2 = 0.28$, p=0.6)
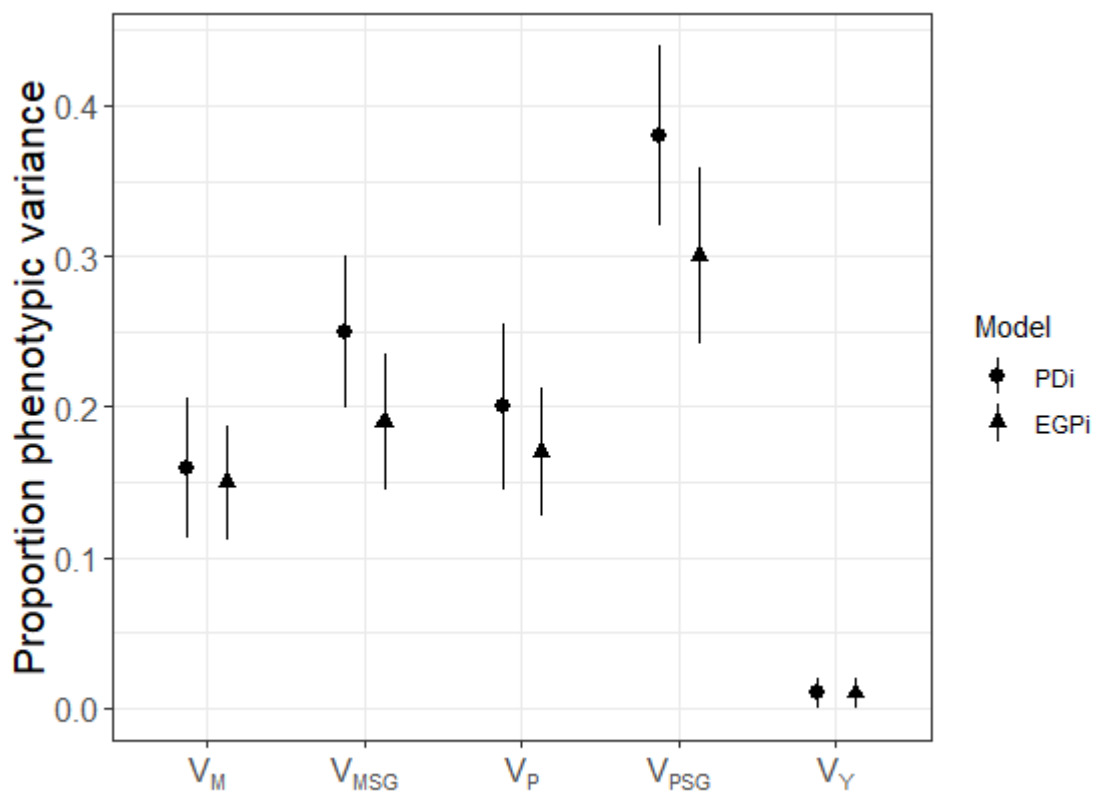
985 ‡ not significantly different from each other (logLRT, PD_i: $\chi^2 = 3.43$, p=0.06, EGP_i: $\chi^2 = 3.68$, p=0.06)

986

987

988

989

990

991

992
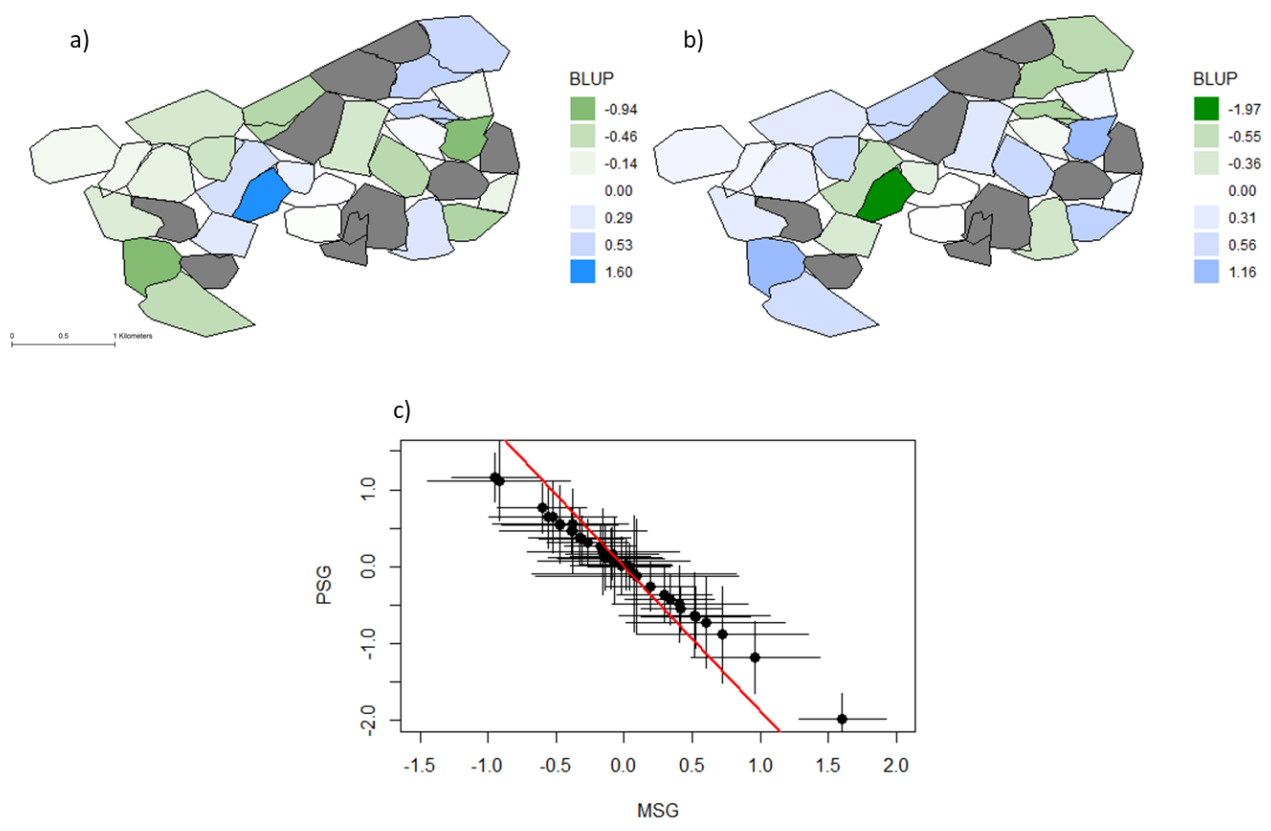
993

994

995

996

997

998
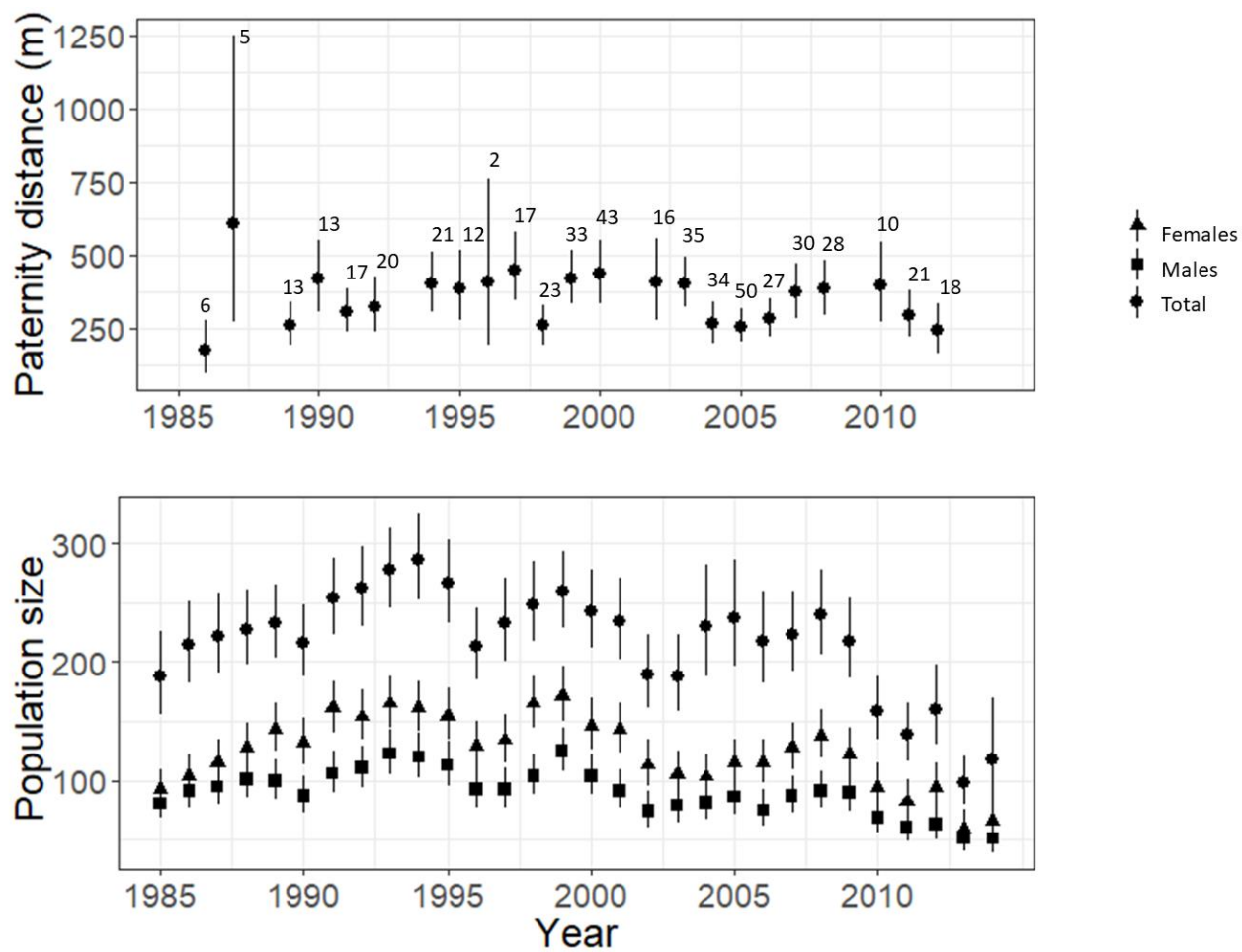
1003

Figure 1.

1005

1006

1007

1008

1009   **Figure 2.**

1010

Figure 3.

1013

1014

1015

1016

This file contains supplementary tables for the article:

Individual variation and the source-sink group dynamics of extra-group paternity in a social mammal

Paula H. Marjamäki*[1,2], Hannah L. Dugdale[3], Deborah A. Dawson[2], Robbie A. McDonald[4], Richard Delahay[5], Terry Burke[2], Alastair J. Wilson[1]

[1]Centre for Ecology and Conservation, University of Exeter, Penryn Campus, Penryn, Cornwall, TR10 9FE, UK
[2]NERC Biomolecular Analysis Facility, Department of Animal and Plant Sciences, University of Sheffield, Western Bank, Sheffield, S10 2TN, UK
[3]School of Biology, Faculty of Biological Sciences, University of Leeds, Leeds, LS2 9JT, UK
[4] Environment and Sustainability Institute, University of Exeter, Penryn Campus, Penryn, Cornwall, TR10 9FE, UK
[5]National Wildlife Management Centre, Animal and Plant Health Agency, Woodchester Park, Gloucestershire, GL10 3UJ, UK

*Corresponding author:

Paula H. Marjamäki
Centre for Ecology and Conservation, University of Exeter, Penryn Campus, Penryn, Cornwall, TR10 9FE, UK
Tel: 07527913031
Email: pm343@exeter.ac.uk

1062 Table S1. Per locus mean allelic dropout (e1) and false allele or stochastic sampling error rates (e2),
1063 estimated using PEDANT 1.0 (Johnson & Haydon 2007) using 209 individuals for which repeat genotypes
1064 were available. Loci for which estimated error was zero, and those for which estimation was not possible
1065 (Mel15 & 106) due to lack of repeat genotypes, the default rate of 0.005 was used (Hadfield 2014).

1066

1067

Locus	E1	95% CI	E2	95% CI2
Mel1	0.03	0.005-0.08	0.006	0.0006-0.02
Mel10	0.03	0.005-0.1	0	0-0.02
Mel12	0.1	0.07-0.2	0.07	0.05-0.1
Mel14	0.02	0.006-0.04	0.03	0.01-0.04
Mel15	0.005	-	0.005	-
Mel101	0.1	0.03-0.2	0.02	0.002-0.06
Mel102	0.02	0.006-0.05	0	0-0.009
Mel103	0.02	0.0009-0.07	0.03	0.006-0.06
Mel104	0.03	0.008-0.08	0.01	0.001-0.04
Mel105	0.03	0.01-0.05	0.05	0.03-0.07
Mel106	0.005	-	0.005	-
Mel107	0.01	0.002-0.05	0	0-0.007
Mel108	0.01	0.003-0.04	0	0-0.007
Mel109	0.07	0.04-0.1	0.08	0.05-0.1
Mel110	0.02	0.003-0.05	0.004	0.00008-0.02
Mel111	0.08	0.04-0.1	0.04	0.01-0.07
Mel112	0.006	0-0.03	0.003	0.00006-0.02
Mel113	0.06	0.02-0.1	0.02	0.005-0.06
Mel114	0.05	0.004-0.2	0	0-0.06
Mel115	0.02	0.004-0.04	0.006	0.0005-0.02
Mel116	0.1	0.05-0.3	0.2	0.002-0.07
Mel117	0.009	0.001-0.03	0	0-0.01

Table S2. Best linear unbiased predictor (BLUP) values (represent the predicted deviation of each (maternal and paternal) social group from the mean paternity distance) and standard errors for each maternal (M) and paternal (P) social group extracted from the PD$_i$ model. Values represent the predicted deviation of each social group from the mean. Groups with missing data had no parentage assignments, therefore BLUPs were not estimated. Results are on the log-transformed scale with untransformed PD in meters.

Social group	BLUP$_P$ (SE)	BLUP$_M$ (SE)
Arthurs	0.31 (0.29)	-0.27 (0.25)
Atcombe West	-0.09 (0.75)	0.07 (0.62)
Atcombe Corner	1.12 (0.53)	-0.92 (0.43)
Bamboo	0.08 (0.56)	-0.07 (0.46)
Beech	0.26 (0.26)	-0.17 (0.22)
Bungalow	-0.63 (0.55)	0.52 (0.45)
Cedar	0.35 (0.31)	-0.31 (0.26)
Cole Park	-0.88 (0.63)	0.72 (0.52)
Colliers Wood	0.03 (0.3)	0.01 (0.28)
Convent	-	-
Dark Wood	-	-0.38(0.45)
Dingle	-0.73 (0.53)	0.59 (0.43)
Field Farm	0.55 (0.50)	-0.47 (0.42)
Gully	-	-
Hedge	-0.36 (0.35)	0.30 (0.29)
Hogarths	-	-
Holly Wood	0.41 (0.41)	0.41 (0.41)
Honeywell	0.65 (0.41)	-0.56 (0.34)
Inchbrook	0.12 (0.42)	-0.14 (0.35)
Jacks Mirey	1.16 (0.32)	-0.95 (0.27)
Kennel	-0.003 (0.30)	0.036 (0.25)
Larch	0.14 (0.29)	-0.10 (0.25)
Listers	-0.73 (0.59)	0.59 (0.48)
Nettle	0.64 (0.47)	-0.52 (0.39)
Old Oak	0.38 (0.37)	-0.32 (0.31)
Park Mill	0.11 (0.39)	-0.09 (0.33)
Peglars	0.02 (0.37)	-0.02 (0.31)
Septic Tank	0.56 (0.28)	-0.38 (0.24)
Thistle Wood Bank	-	-
Top Sett	-1.97 (0.32)	1.60 (0.26)
West	0.17 (0.34)	-0.09 (0.28)
Windsor Edge	0.76 (0.33)	-0.60 (0.28)
Wood Farm	-0.42 (0.33)	0.34 (0.27)
Wych Elm	-0.25 (0.32)	0.19(0.26)
Yew	-0.55 (0.29)	0.42 (0.25)

Table S3. Reanalysis of PDi and EGPi using standardised body mass index (SMI) in place of body mass. Response variables were standardised into standard deviation units (SDU) prior to analysis. M and P denote maternal and paternal individuals, while MSG and PSG denote the corresponding maternal and paternal social groups.

	$\text{Log(PD}_i)$				EGP_i			
	Estimate (SE)	DF	F	P	Estimate (SE)	DF	F	P
Intercept	0.73 (0.16)	1, 284.6	20.47	**<0.001**	0.74 (0.16)	1, 297.7	21.01	**<0.001**
Age_M	-0.009 (0.01)	1, 539.9	0.55	0.46	0.01 (0.01)	1, 543.3	0.64	0.42
$\text{SMI}_M{}^†$	0.009 (0.01)	1, 333.3	0.78	0.38	0.009 (0.01)	1, 336.9	0.69	0.41
Group_size_{MSG}	0.007 (0.02)	1, 461.4	0.19	0.66	0.008 (0.02)	1, 446.9	0.20	0.66
Sex_ratio_{MSG} ‡	-0.71 (0.22)	1, 533.4	10.28	**<0.001**	-0.79 (0.22)	1, 526.5	12.74	**<0.001**
Age_P	0.03 (0.02)	1, 506.6	2.24	0.14	0.03 (0.02)	1, 507.9	2.54	0.11
$\text{SMI}_P{}^†$	-0.02 (0.01)	1, 247.7	1.15	0.29	-0.02 (0.01)	1, 249.9	1.11	0.29
Group.Size_{PSG}	-0.02 (0.02)	1, 538.2	0.69	0.41	-0.02 (0.02)	1, 532.4	0.68	**0.04**
Sex_ratio_{PSG} ‡	0.42 (0.24)	1, 539.6	3.06	0.08	0.50 (0.24)	1, 537.1	4.29	**<0.001**

†mean body mass for parental individuals with multiple weight measurements within cub's birth year

‡calculated as number of males divided by group size where group size is males plus females

Full models fitted for each response were y ~ μ + Age_M + SMI_M + Group_size_{MSG} + Sex_ratio_{MSG} + Age_P + SMI_P + Group_size_{PSG} + Sex_ratio_{PSG} + *M* + *P* + *MSG* + *PSG* + *Year* where italic font denotes random effects and y is either $\text{log(PD}_i)$ or EGP_i

1134

1135 Table S4. Estimated (co)variance components (standard error) associated with random effects in mixed
1136 models of EGP_i and log-transformed PD_i, reanalysed using using standardised body mass index (SMI) in
1137 place of body mass. Statistical inference of random effects is by likelihood ratio test results (see main text
1138 for details). M and P denote maternal and paternal individuals, while MSG and PSG denote the
1139 corresponding maternal and paternal social groups.

1140

	log(PD$_i$)				EGP$_i$			
	Variance (SE)	df	χ^2_1	P	Variance (SE)	df	χ^2_1	P
V_{year}	0.06 (0.02)	1	3.76	**0.05**	0.02 (0.01)	1	3.20	0.07
$V_M{}^\dagger$	0.25 (0.05)	1	40.74	**<0.001**	0.26 (0.05)	1	40.91	**<0.001**
$V_P{}^\dagger$	0.31 (0.06)	1	35.22	**<0.001**	0.31 (0.06)	1	34.71	**<0.001**
$V_{MSG}{}^\ddagger$	0.41 (0.15)	1	20.64	**<0.001**	0.35 (0.13)	1	19.92	**<0.001**
$V_{PSG}{}^\ddagger$	0.60 (0.21)	1	26.57	**<0.001**	0.54 (0.19)	1	27.5	**<0.001**
COV$_{MSG,PSG}$	*-0.49 (0.17)*	*1*	*39.84*	**<0.001**	*-0.44 (0.15)*	*1*	*37.05*	**<0.001**
V_R	0.32 (0.04)	-	-	-	0.32 (0.03)	-	-	-

1141 † not significantly different from each other (logLRT, PD$_i$: χ^2_1 = 0.22, p=0. 64; EGP$_i$: χ^2_1 = 0.30, p=0.59)

1142 ‡ not significantly different from each other (logLRT, PD$_i$: χ^2_1 = 3.73, p= 0.05; EGP$_i$: χ^2_1 = 3.69, p=0.05)

1143

1144 Table S5. Repeatabilities (R) of variance components from reanalyses of EGP_i and log-transformed PD_i,
1145 reanalysed using standardised body mass index (SMI) in place of body mass. R calculated as variance
1146 component/sum of all variance components. Values for *COR$_{MSG,PSG}$* are correlation coefficients. M and P
1147 denote maternal and paternal individuals, while MSG and PSG denote the corresponding maternal and
1148 paternal social groups.

1149

	log(PD$_i$) R (SE)	EGP$_i$ R (SE)
R_{year}	0.01(0.008)	0.009 (0.008)
$R_M{}^a$	0.13 (0.04)	0.14 (0.04)
$R_P{}^a$	0.16 (0.04)	0.17 (0.04)
$R_{MSG}{}^b$	0.22 (0.05)	0.20 (0.05)
$R_{PSG}{}^b$	0.31 (0.06)	0.30 (0.06)
COR$_{MSG,PSG}$	*-0.99 (0.03)*	*-0.99 (0.03)*
R_R	0.17 (0.04)	0.18 (0.04)

1150

1151

1152

1153

1154

1155

Table S6. Estimated fixed effect coefficients (standard error) and Wald F-tests from mixed models from reanalysis on log-PD_i and EGP_i using the 95% confidence pedigree, where only those parent assignments that met a 95% confidence threshold were included. Response variables were standardised into standard deviation units (SDU) prior to analysis. M and P denote maternal and paternal individuals, while MSG and

	Log(PD_i)				EGP_i			
	Estimate (SE)	DF	F	P	Estimate (SE)	DF	F	P
Intercept	0.62 (0.23)	1, 94.3	7.22	**<0.01**	0.64 (0.23)	1, 92.7	7.51	**<0.01**
Age_M	-0.01 (0.008)	1, 115.9	0.01	0.91	0.002 (0.007)	1, 115.5	0.09	0.76
Body mass$_M$†	-0.04 (0.026)	1, 219.1	2.01	0.16	-0.04 (0.03)	1, 220.0	2.09	0.15
Group_size$_{MSG}$	0.12 (0.01)	1, 96.0	127.59	**<0.001**	0.13 (0.01)	1, 88.3	185.62	**<0.001**
Sex_ratio$_{MSG}$ ‡	-3.29 (0.14)	1, 142.5	535.52	**<0.001**	-3.65 (0.13)	1, 135.1	805.31	**<0.001**
Age_P	0.03 (0.008)	1, 112.0	19.16	**<0.001**	0.03 (0.007)	1, 112.9	18.67	**<0.001**
Body mass$_P$†	-0.02 (0.03)	1, 156.8	0.71	0.40	-0.02 (0.03)	1, 160.7	0.66	0.42
Group.Size$_{PSG}$	-0.08 (0.01)	1, 97.0	28.90	**<0.001**	-0.09 (0.01)	1, 89.4	52.63	**<0.001**
Sex_ratio$_{PSG}$ ‡	2.84 (0.17)	1, 160.2	287.70	**<0.001**	3.22 (0.15)	1, 149.8	446.25	**<0.001**

PSG denote the corresponding maternal and paternal social groups.

1161

†mean body mass for parental individuals with multiple weight measurements within cub's birth year

‡calculated as number of males divided by group size where group size is males plus females

Full models fitted for each response were y ~ μ + Age_M + Body_Mass$_M$ + Group_size$_{MSG}$ + Sex_ratio$_{MSG}$ + Age_P + Body_Mass$_P$ + Group_size$_{PSG}$ + Sex_ratio$_{PSG}$ + *M* + *P* + *MSG* + *PSG* + *Year* where italic font denotes random effects and y is either log(PD_i) or EGP_i

1167

1168

1169

1170

1171

1172

1173

1174

1175

1176

1177

1178

1179

1180

1181

1182

1183

1184

1185 Table S7. Estimated (co)variance components (standard error) associated with random effects in mixed

1186 models of EGP_i and log-transformed PD_i, reanalysed using 95% confidence pedigree. Statistical inference of

1187 random effects is by likelihood ratio test results (see main text for details). M and P denote maternal and

1188 paternal individuals, while MSG and PSG denote the corresponding maternal and paternal social groups.

1189

	log(PD_i)				EGP_i			
	Variance (SE)	df	χ^2_1	P	Variance (SE)	df	χ^2_1	P
V_{year}	0.04 (0.02)	1	45.88	**<0.001**	0.04 (0.02)	1	60.89	**<0.001**
V_M[†]	1.86 (0.23)	1	115.12	**<0.001**	1.94 (0.24)	1	131.15	**<0.001**
V_P[†]	1.80 (0.25)	1	98.34	**<0.001**	1.93 (0.26)	1	114.94	**<0.001**
V_{MSG}[‡]	2.21 (0.71)	1	41.93	**<0.001**	2.13 (0.68)	1	55.91	**<0.001**
V_{PSG}[‡]	2.22 (0.71)	1	197.85	**<0.001**	2.16 (0.69)	1	80.31	**<0.001**
$COV_{MSG,PSG}$	*-2.04 (0.66)*	*1*	*37.27*	**<0.001**	*-1.96 (0.64)*	*1*	*35.58*	**<0.001**
V_R	0.005 (0.0008)	-	-	-	0.004 (0.0006)	-	-	-

1190

1191 [†]not significantly different from each other (logLRT, PD_i: $\chi^2_1 = 0.038$, p=0.85; EGP_i: $\chi^2_1 = 0.002$, p=0.96)

1192 [‡] not significantly different from each other (logLRT, PD_i: $\chi^2_1 = 0$, p= 1; EGP_i: $\chi^2_1 = 0.006$, p=0.94)

1193

1194

1195

1196 Table S8. Repeatabilities (R) of variance components from reanalyses of EGP_i and log-transformed PD_i,

1197 reanalysed using 95% confidence pedigree. R calculated as variance component/sum of all variance

1198 components. Values for *$COV_{MSG,PSG}$* are correlation coefficients. M and P denote maternal and paternal

1199 individuals, while MSG and PSG denote the corresponding maternal and paternal social groups.

1200

	log(PD_i) R (SE)	EGP_i R (SE)
R_{year}	0.005(0.002)	0.004 (0.002)
R_M[a]	0.23 (0.04)	0.24 (0.04)
R_P[a]	0.22 (0.04)	0.24 (0.04)
R_{MSG}[b]	0.27 (0.05)	0.26 (0.05)
R_{PSG}[b]	0.27 (0.05)	0.26 (0.05)
$COR_{MSG,PSG}$	*-0.92 (0.05)*	*-0.92 (0.05)*
R_R	0.0007 (0.0002)	0.0005 (0.05)

1201

1202

1203

1204

1205 Table S9. Posterior mean (credible intervals) estimates of unsampled males and females per cohort
1206 estimated in MasterBayes simultaneously with parentage and paternity distance. Values for unsampled
1207 males represent population-level estimates, while number of unsampled females was estimated per social
1208 group.

| Year	Unsampled males	Unsampled females
1986	11.827 (0.407-43.445)	1.2651 (0.316-3.474)
1987	35.622 (6.093-89.752)	0.819 (0.1945-1.958)
1988	26.1864 (0.805-96.173)	0.975 (0.122-2.801)
1989	6.401 (0.250-21.541)	0.548 (0.044-1.626)
1990	10.764 (0.676-32.665)	0.803 (0.054-2.465)
1991	16.404 (1.693-44.149)	2.314 (0.960-4.439)
1992	37.147 (6.335-90.084)	0.380 (0.011-1.269)
1993	20.403 (0.6008-68.087)	0.843 (0.123-2.136)
1994	12.696 (0.680-39.903)	0.359 (0.009-1.341)
1995	40.303 (5.744-102.097)	6.283 (2.062-12.629)
1996	239.383 (32.810-812.610)	2.370 (0.091-7.561)
1997	16.930 (2.090-43.140)	0.980 (0.100-2.640)
1998	47.200 (17.15-89.200	0.540 (0.110-1.250)
1999	35.000 (10.390-71.470)	0.650 (0.160-1.400)
2000	35.081 (12.290-68.310)	0.799 (0.239-1.76)
2001	28.868 (12.280-50.640)	0.604 (0.226-1.174)
2002	55.9474 (12.230-132.450)	0.428 (0.010-1.610)
2003	24.937 (5.672-55.067)	0.517 (0.066-1.312)
2004	20.150 (2.961-49.365)	0.705 (0.142-1.695)
2005	10.192 (1.035-26.744)	0.919 (0.313-1.850)
2006	5.129 (0.208-17.236)	2.653 (1.141-4.951)
2007	6.859 (0.389-20.929)	1.363 (0.403 -2.915)
2008	18.010 (4.195-39.820)	1.896 (0.661-3.836)
2009	16.416 (5.726-31.224)	0.702 (0.210-1.462)
2010	18.812 (0.970-57.216)	2.353 (0.482-6.379)
2011	6.698 (0.168-23.674)	1.3234 (0.308-2.703)
2012	49.145 (13.680-105.520)	0.739 (0.153-1.795)
2013	50.206 (21.790-88.780)	0.614 (0.167-1.338)
2014	111.922 (49.660-217.92)	2.225 (1.000-4.137)

1209

1210

1211

1212

1213

1214

1215

1216

1217

1218   Figure S1. Inferred pedigree structure for 29 cohorts showing maternal assignments in red, paternal in blue
1219   and individuals as dots. Reconstructed pedigree has a maximum depth of six generations and contains 579
1220   maternal-cub and 596 paternal-cub links, 186 full sibships, 452 maternal half-sibs, and 927 paternal half
1221   sibs.
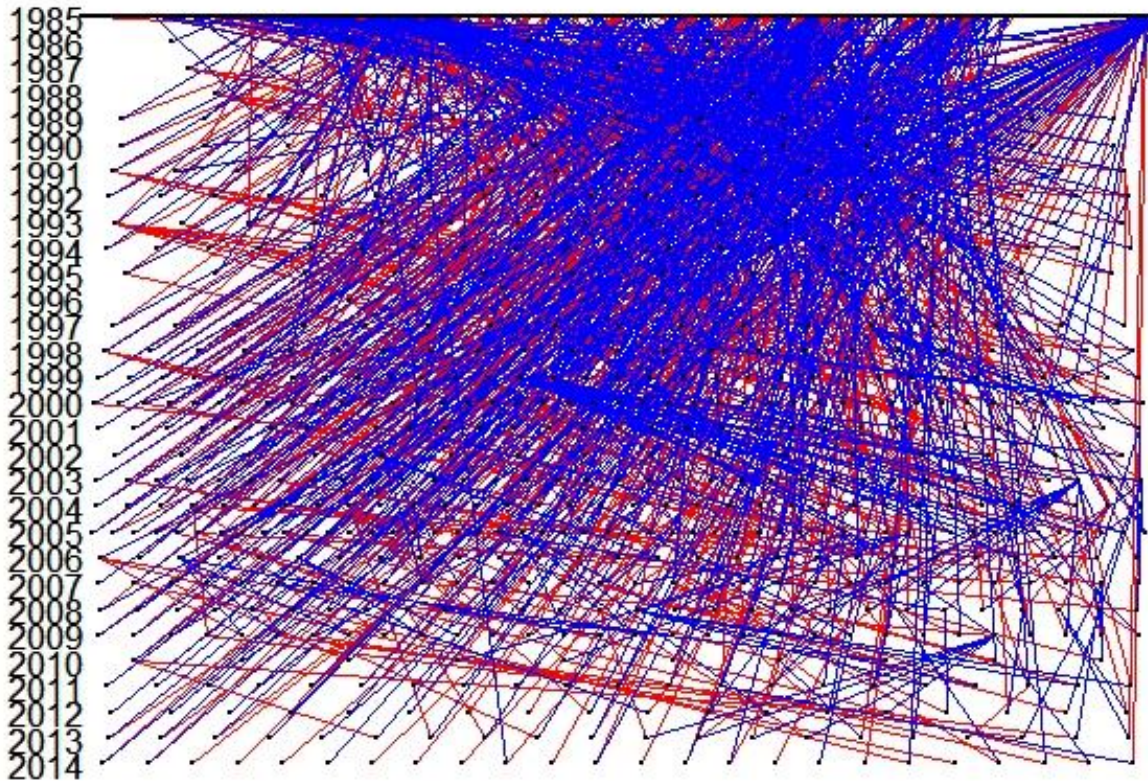
1222

1223

1224

1225

1226

1227

1228

1229

1230

1231

1232

1233

1234

1235

1236

1237

1238   Figure S2. Histogram of model residuals for binary EGP_i (0/1) run in ASReml 3.0 with a Gaussian error
1239   structure.
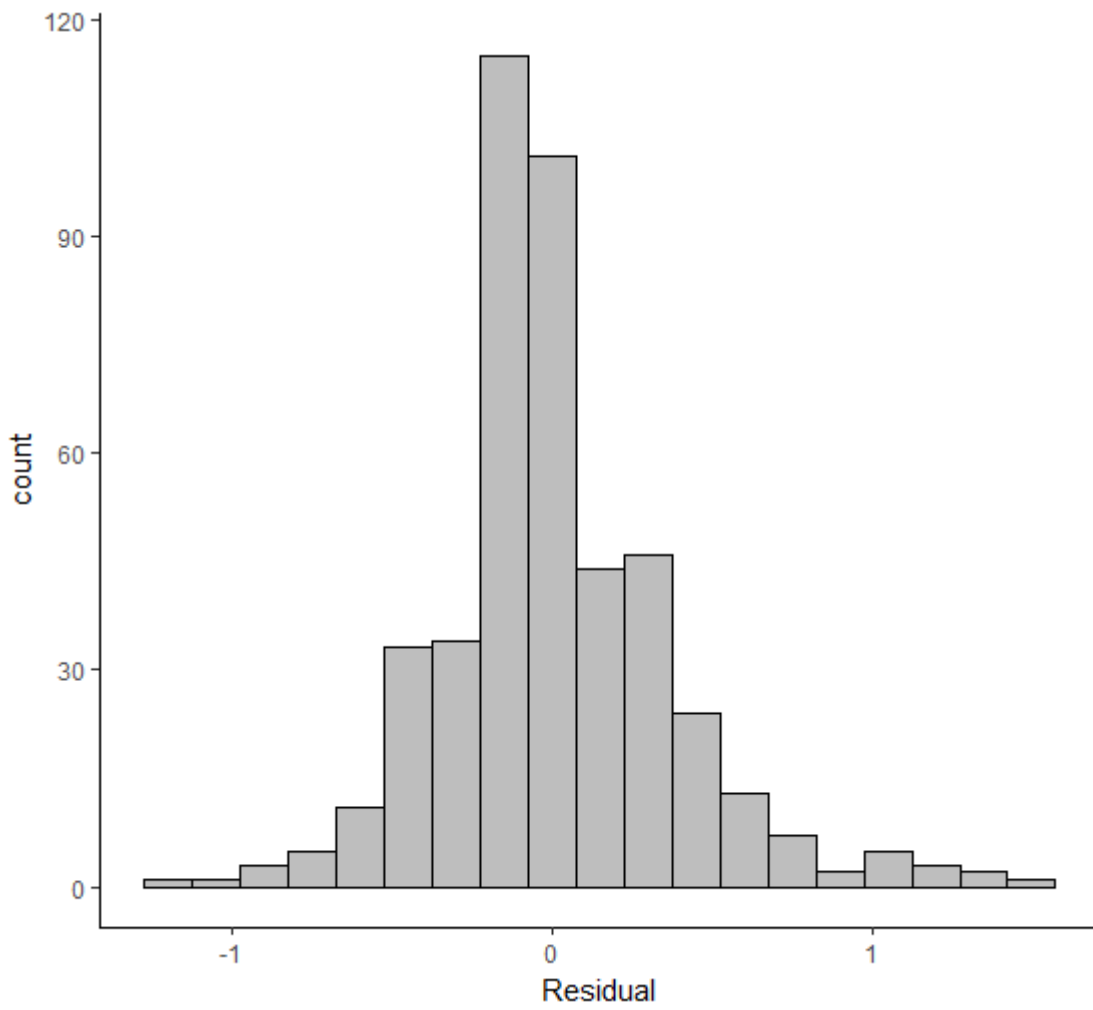
1240

1241

1242

1243

1244

1245

1246

1247

1248