

# ESTIMATING THE ONTOGENETIC AGE AND SEX COMPOSITION OF FAUNAL ASSEMBLAGES WITH BAYESIAN MULTILEVEL MIXTURE MODELS

Jesse Wolfhagen<sup>1</sup>

<sup>1</sup>Department of Anthropology, Purdue University, West Lafayette, Indiana, USA

email: jl.wolfhagen@gmail.com

*R Markdown version last compiled on 2023-03-03*

## ABSTRACT

Understanding the ontogenetic age and sex composition of zooarchaeological assemblages can reveal details about past human hunting and herding strategies as well as past animal morphology and behavior. As such, the accuracy of our estimates underlies our ability to ascertain details about site formation and gain insights into how people interacted with different animals in the past. Unfortunately, our estimates typically rely on only a small number of bones, limiting our ability to fruitfully use these estimates to make meaningful comparisons to theoretical expectations or even between multiple assemblages. This paper describes a method to use zooarchaeological remains with standard biometric measurements to estimate the ontogenetic age and sex composition of the assemblage, focused on immature, adult-sized female, and adult-sized male specimens. The model uses a Bayesian framework to ensure that the parameter estimates are biologically meaningful. Simulated assemblages show that the model can accurately estimate the biometry and composition of zooarchaeological assemblages. Two archaeological case studies also show how the model can be applied to produce tangible insights. The first, focused on sheep from Neolithic Pinarbaşı B, highlights the model's ability to elucidate site formation and function. The second, focused on cattle remains from four assemblages from 7th-6th Millennium BCE northwestern Anatolia, showcases how to use the mixture modeling results to compare assemblages to one another and to specific hypotheses. This modeling framework provides a new avenue for investigating long-term trajectories in animal biometry alongside contextual analyses of past human choices in butchery and consumption.

**Keywords:** *Zooarchaeology, Biometry, Logarithmic size index (LSI), Domestication, Bayesian statistics.*

## <sup>1</sup> 1. INTRODUCTION

<sup>2</sup> Different hunting and herding strategies target specific classes of animals among a herd that are determined  
<sup>3</sup> by the animal's ontogenetic age and sex (Dahl and Hjort 1976; Stiner 1990). In addition to human-driven  
<sup>4</sup> goals, sex differences in habitat use, diet quality, and reproductive capabilities among ungulate prey species  
<sup>5</sup> contribute to the susceptibility and desirability of males and females at different ages to human exploitation  
<sup>6</sup> (Corti and Shackleton 2002; Post et al. 2001; Ruckstuhl and Neuhaus 2002; Ruckstuhl 2007; Saïd et al.  
<sup>7</sup> 2011). These factors impact the formation of bone assemblages by affecting the probabilities that bones  
<sup>8</sup> from different classes of animals (e.g., immature, adult female, or adult male animals) are deposited before  
<sup>9</sup> being mediated by other taphonomic processes (Lyman 2008). The ontogenetic age and sex composition  
<sup>10</sup> of zooarchaeological assemblages can therefore reflect anthropologically-relevant aspects of past hunting  
<sup>11</sup> strategies—like seasonal site use and scale of exploitation (Speth 2013)—or general management goals of  
<sup>12</sup> past herding strategies (e.g., Payne 1973; Redding 1984).

<sup>13</sup> Reconstructing the ontogenetic age and sex composition of a zooarchaeological assemblage can enrich our  
<sup>14</sup> understanding of past human-animal interactions by complementing mortality profiles and inter-assemblage  
<sup>15</sup> comparisons. However, the disaggregated nature of faunal assemblages complicates efforts to conclusively  
<sup>16</sup> identify the ontogenetic age and sex of a specimen. Because articulated remains are rare, zooarchaeologists  
<sup>17</sup> typically cannot relate elements that are morphologically distinct between the sexes (e.g., the pelvis) to other  
<sup>18</sup> elements that can provide information about the animal's age-at-death (e.g., limb bones or mandibles). We  
<sup>19</sup> can, though, take advantage of the general pattern of sexual dimorphism among ungulate taxa by using size  
<sup>20</sup> differences in limb bones to distinguish between males and females.

### <sup>21</sup> 1.1 *Morphometric Sex Determination in Zooarchaeology*

<sup>22</sup> Some biometric methods to determine the sex of an animal bone are multivariate—using combinations of  
<sup>23</sup> measurements with bivariate plots or discriminant functions to predict the sex of archaeological specimens  
<sup>24</sup> based on distributions of known-sex specimens (e.g., Munro, Bar-Oz, and Hill 2011; Speth 1983; Widga  
<sup>25</sup> 2006). These methods typically combine dimensions from different planes of an element (e.g., the breadth  
<sup>26</sup> and depth of a distal articular end) to produce patterns that can be separated by a ‘cut point’ between  
<sup>27</sup> males and females, either visually in the case of bivariate plots or algorithmically in the case of discriminant  
<sup>28</sup> functions. The need to preserve multiple measured dimensions, even on the same end of an element, may  
<sup>29</sup> make it difficult to apply these methods to more heavily processed assemblages. Further, specimens from  
<sup>30</sup> animals that died before reaching adult body size may be misclassified as females, particular for dimensions  
<sup>31</sup> affected by post-fusion growth (Popkin et al. 2012).

32 Other sex determination methods are univariate—they use a single measurement from a specimen and  
33 typically use size index methods to associate those measurements from different elements together (e.g.,  
34 Weinstock 2006; Zeder and Lemoine 2020). This approach allows general descriptions of the sex ratio in  
35 an assemblage that can be used to identify changes in these sex ratios or overall biometry over time (e.g.,  
36 Arbuckle and Atici 2013; Grigson 1989). Zeder and Lemoine (2020) go further by using inter-quartile ranges  
37 of log size index (LSI) values from their reference population to create ‘cut-off’ values between immature,  
38 female, and male specimens to calculate specific ontogenetic age and sex ratios for elements and assemblages.

39 Regardless of whether the method uses multivariate or univariate data, these sex determination methods  
40 tend to have the same weaknesses. Practically, these methods rely on direct comparisons with reference  
41 populations (typically, but not always, modern populations of known sex). Thus, the analysis implies that  
42 the biometry of the archaeological population is the same as the reference population. However, this impli-  
43 cation is an untenable one in most cases, as animal biometry typically varies spatially and temporally due  
44 to climatic and ecological variation (e.g., Koch 1986; Davis 1982; Wright and Viner-Daniels 2015; Hill, Hill,  
45 and Widga 2008; Lebennzon and Munro 2022). Biometric variation has also been attributed to anthropogenic  
46 pressures, particularly among domesticated taxa (e.g., Arbuckle and Kassebaum 2021; Manning et al. 2015;  
47 Trentacoste et al. 2021; Grau-Sologestoa and Albarella 2019), though harvest pressure has also been at-  
48 tributed to biometric changes in wild taxa (e.g., Wolverton 2008; Munro, Lebennzon, and Sapir-Hen 2022).  
49 These environmental and anthropogenic pressures may affect males and females differently (e.g., Tchernov  
50 and Horwitz 1991; Zohary, Tchernov, and Horwitz 1998); pressures that reduce sexual dimorphism could  
51 interfere with analyses, as more specimens may be indeterminate or misclassified. Biometric variation be-  
52 tween populations can complicate efforts to estimate changes in the demographic (ontogenetic age and sex)  
53 composition of assemblages over time. Further, efforts to control for ontogenetic age (e.g., removing unfused  
54 specimens or those from early-fusing elements) distorts the relationship between the analyzed specimens and  
55 the rest of the assemblage, decreasing our ability to make reliable inferences about the entire assemblage  
56 (Zeder and Hesse 2000).

57 Philosophically, sex determinations made by these methods tend to be absolutist: specimens are identified  
58 as male or female (or immature) or are marked as indeterminate. As in taxonomic identifications, the use  
59 of absolutist determinations masks any underlying uncertainty in the determination (Wolfhagen and Price  
60 2017). Removing indeterminate specimens from consideration artificially reduces sample sizes and inflates  
61 reported accuracy rates. This produces a false sense of confidence in the sex determination results, especially  
62 when those results are then used to summarize the entire assemblage. More critically, any nuances or caveats  
63 in the sex determinations of an assemblage are lost when the results are used in synthetic analyses at larger  
64 spatial and temporal scales. What is necessary is a way to estimate the ontogenetic age and sex composition

65 of a faunal assemblage that preserves the uncertainty inherent in the process.

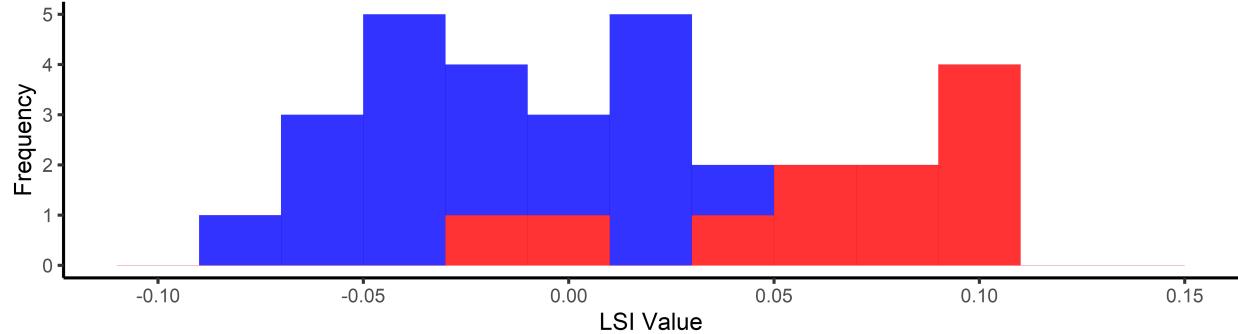
66 **1.2 Mixture Modeling in Zooarchaeology**

67 Mixture modeling provides just such a method, producing probabilistic sex identifications rather than ab-  
68 solute ones. By describing an assemblage of faunal measurements as a mixture of specimens from different  
69 animal groups (generally termed “mixture components”) like male and female specimens—described by pa-  
70 rameters for the proportion of the overall assemblage ( $\pi$ ), average size ( $\mu$ ), and size variability ( $\sigma$ ) of each  
71 animal group. A mixture model allows researchers to not only describe the overall composition of the as-  
72 semblage but to also estimate the probabilities that a specific specimen belongs to a particular animal group  
73 (Dong 1997; Monchot and Léchelle 2002). Additionally, mixture modeling does not rely on a reference popu-  
74 lation, allowing biometric variation between populations and even changes in the extent of sexual dimorphism  
75 (e.g., Helmer et al. 2005). These features allow mixture models the flexibility to track both biometric and  
76 demographic variation across assemblages over time and space (e.g., Arbuckle et al. 2016; Arbuckle and  
77 Kassebaum 2021).

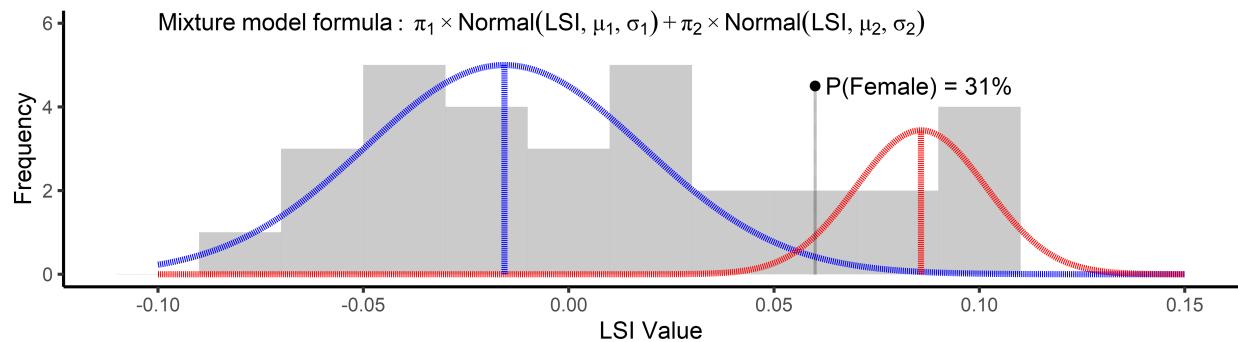
78 Conceptually, a mixture model can be thought of as a “latent state” or “missing data” problem: we  
79 know that measured specimens come from particular animal groups, but that information has been lost (Marin,  
80 Mengersen, and Robert 2005). If we knew every specimen’s group identity, then the calculation of the mixture  
81 model parameters (mixture proportion, average body size, and size variability for each animal group) would  
82 be trivial. In archaeological contexts, however, we cannot directly observe those group identities; we must  
83 therefore use probabilities of group membership and calculate group-specific parameters from those resulting  
84 probabilities (Monchot and Léchelle 2002). Figure 1 describes a schematic example of a mixture model:  
85 Figure 1A shows the distribution of LSI values from a reference population of 31 adult pig (*Sus domesticus*)  
86 distal tibia distal breadths (Tibia Bd: Driesch 1976) described in Zeder and Lemoine (2020), with specimens  
87 colored by their known identity (females = blue, males = red). Figure 1B shows the results of fitting a  
88 two-component (females and males) mixture model to the data using standard approaches (e.g., Monchot  
89 and Léchelle 2002; Arbuckle and Kassebaum 2021), ignoring those true identities.

90 The mixture model describes the assemblage as a mixture of the two ‘mixture components’ (males and  
91 females): each component is described with three parameters: a proportion ( $\pi$ ), an average size ( $\mu$ ), and a  
92 standard deviation ( $\sigma$ ). Taken together, these parameters determine a specimen’s probability of being in one  
93 of the groups, as shown in Panel B; a specimen with an  $LSI_e$  value of 0.06 has a 31% probability of being  
94 female based on the model. This results in 31 sets of probabilities, one for each specimen; Figure 1C shows  
95 four plausible simulated assemblages that result from the mixture model; every specimen’s membership  
96 probability is used to simulate a ‘true’ identity. Importantly, by leaving the mixture model results as

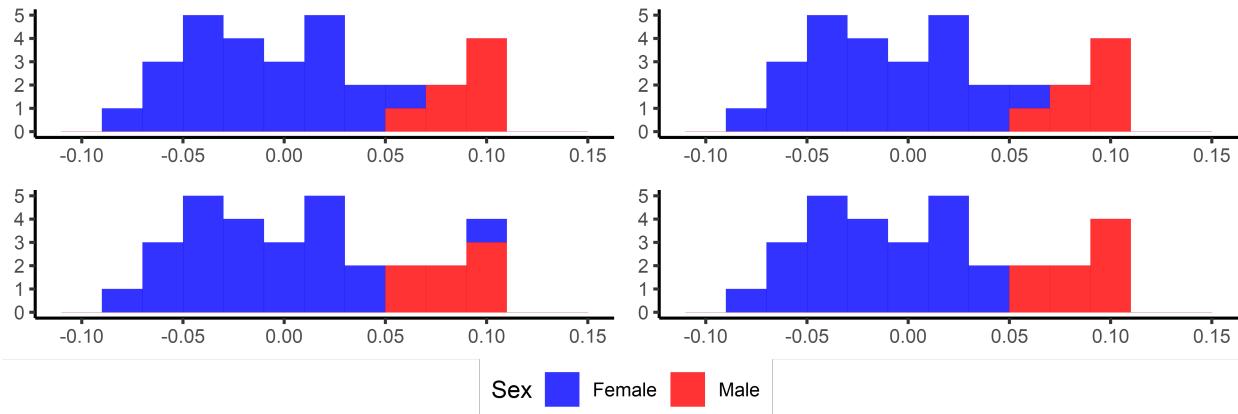
### A Adult Pig Tibia Bd (Known Identities)



### B Mixture Model Fit



### C Simulated Histograms



**Figure 1:** Walkthrough of the mixture modeling procedure using pig Tibia Bd measurements from Zeder and Lemoine (2022). A: the distribution of Tibia Bd LSI values (standard value: 33.5 mm; Hongo and Meadow 2000) of adult specimens, shaded by known sex (females = blue, males = red). B: The result of a standard (non-Bayesian) mixture model analysis on the LSI values, ignoring sex. Vertical lines show the estimated means for females and males; curves show the relative probability densities for the two distributions. C: Four simulations using the membership probabilities of each specimen based on the mixture model results.

97 specimen-specific probabilities of being female or male, mixture model results retain the uncertainty of the  
98 sex determination process: that is, a specimen with a 51% probability of being female is not treated as  
99 equivalent to one with a 95% probability of being female.

100 The previous example showcases the benefits of mixture modeling as a flexible probabilistic sex de-  
101 termination method. First, the model does not require comparison to a reference population to estimate  
102 differences between female and male specimens. Second, the model produces parametric estimates of body  
103 size and size variability that can be used for inter-site comparisons, rather than just determinations for the  
104 included specimens. Finally, the model produces probabilistic estimates for every specimen, rather than  
105 leaving some specimens indeterminate and obscuring variation in the confidence of the sex assignments.  
106 These theoretical and practical advantages of mixture modeling and its potentials for zooarchaeology have  
107 been apparent since its introduction to the field (e.g., Dong 1997; Monchot and Léchelle 2002; Monchot,  
108 Mashkour, and Vigne 2005), though its application has been piecemeal over the past two decades despite  
109 the existence of free scientific software that can perform the analysis (e.g., PAST: Hammer 2013; R packages  
110 “mixtools” and “mclust”: Benaglia et al. 2009; Scrucca et al. 2016).

111 The reasons for the patchy application of mixture modeling in zooarchaeology are less straightforward.  
112 High-profile early case studies of mixture modeling report size variability parameters (standard deviation  
113  $\sigma$ ) that vary widely and include very small values for some groups (e.g., De Cupere et al. 2005; Monchot,  
114 Mashkour, and Vigne 2005; Vigne 2011). These results suggest that the very flexibility that is a great  
115 strength of mixture modeling is actually identifying ‘groups’ that are not necessarily consistent with biological  
116 expectations (e.g., that the results are “overfitted” to the observed data). Such extreme differences in  
117 the standard deviation of different groups can result in counterintuitive implications: specimens may be  
118 considered more likely to come from the broad distribution (the one with the larger  $\sigma$  parameter) than the  
119 narrow distribution even when the value is more extreme than the narrow distribution’s mean (e.g., is larger  
120 than a larger mean or smaller than a smaller mean). Returning to the mixture model example in Figure 1  
121 can explain this issue more clearly. Table 1 shows the mixture model parameters for the two components:  
122 the standard deviation ( $\sigma$ ) for females is more than twice the standard deviation for males ( $\sigma_1 = 0.034$ ,  $\sigma_2 =$   
123 0.016). As such, higher numbers beyond the observed range will be considered likely females: an  $LSI_e$  value  
124 of 0.173 (Tibia Bd value: 39.83 mm) is more likely to be a female than a male using the mixture model’s  
125 results (probability of being female: 51%).

**Table 1:** Mixture model parameter estimates for the pig distal tibia Bd example. Estimates calculated using maximum-likelihood approaches.

Group	Mixture Component $\pi$	Average Size $\mu$	Size Variability $\sigma$
Female	0.76	-0.016	0.034
Male	0.24	0.086	0.016

126 Published mixture model examples show this issue, as well. De Cupere et al. (2005, Table 2) report  
 127 three groups of chicken carpometacarpus lengths from bones with medullary bone, providing the full set of  
 128 mixture model parameters (Group 1: proportion  $\pi = 0.285$ , mean  $\mu = 33.337$ , standard deviation  $\sigma = 0.3$ ;  
 129 Group 2:  $\pi = 0.608$ ,  $\mu = 35.416$ ,  $\sigma = 0.433$ ; Group 3:  $\pi = 0.107$ ,  $\mu = 37.866$ ,  $\sigma = 0.094$ ). According to De  
 130 Cupere et al. (2005, fig. 3), there is one carpometacarpus with a medullary bone whose greatest length is  
 131 roughly 41.5 mm. Counterintuitively, the analysis would suggest that this specimen is most likely to be a  
 132 member of Group 2; it even determines that the specimen is more likely to be a member of Group 1 than  
 133 Group 3. Vigne (2011, Table 3A) reports mixture modeling results of  $LSI_{10}$  values from cattle recovered  
 134 from Neolithic Shillourokambos, Cyprus to estimate females and males, using PAST (Hammer 2013). The  
 135 reported values for the Recentes phase (Female  $\pi = 0.75$ ,  $\mu = 0.120$ ,  $\sigma = 0.042$ ; Male  $\pi = 0.25$ ,  $\mu = 0.163$ ,  
 136  $\sigma = 0.007$ ) produce counterintuitive results: a specimen with an  $LSI$  value of 0.176-within the range of  $LSI$   
 137 values from this phase (Vigne 2011: Figure 2)-would be considered more likely to be female than male.  
 138 These issues extend to more recent publications. Arbuckle et al. (2016, fig. 5) report sex-specific  $LSI_{10}$   
 139 average sizes for cattle in the Eastern Fertile Crescent during the early-mid Holocene; because they report  
 140 their  $LSI_{10}$  data in a supplement, it can be shown that the smallest measurement from Ganj Dareh ( $LSI_{10}$   
 141 = -0.044, modeled female mean = -0.019, modeled male mean = 0.024) is considered more likely to be male  
 142 than female due to the extreme differences in standard deviations.

143 These examples highlight the difficulties researchers face when interpreting the results of mixture analyses  
 144 of zooarchaeological data. While mixture modeling provides the flexibility to model data from a pre-specified  
 145 or unknown number of groups, there is no guarantee that the identified ‘groups’ are biologically meaningful.  
 146 Analysts may identify inconsistent results from mixture analyses and exclude the analysis from reports,  
 147 leaving only mixture analyses that appear to have interpretable results (the “file drawer problem”: Rosenthal  
 148 1979). As these examples show, however, mixture analyses applied to more abstract quantities, like  $LSI$   
 149 values, or interpreted in light of less easily interpreted biological groups, like breeds, can have counterintuitive  
 150 implications. These examples are not meant to highlight the errors; on the contrary, the fact that the  
 151 authors report their full model results and/or data mean that such errors could be identified, highlighting

152 the importance of open scientific reporting and publishing (Marwick 2017; Ram and Marwick 2018).

153 Zooarchaeologists have a wealth of reference information that can inform them about the impacts of diet,  
154 sex, castration, and other factors on the size and variability of animal bones. These reference populations  
155 provide raw measurements from several taxa and generally include specimens of known age-at-death and  
156 sex, though sometimes these include archaeological data of (relatively) complete individuals that can be  
157 assigned to sex (e.g., sheep: Popkin et al. 2012; Davis 1996, 2000; pigs: Zeder and Lemoine 2020; Payne  
158 and Bull 1988; aurochs/cattle: Degerbøl 1970; bison: Speth 1983; Todd 1983). These data can provide  
159 useful information that could be relevant for interpreting a mixture model analysis; ideally, one could take  
160 advantage of relevant information from reference populations while still maintaining some aspects of a mixture  
161 model's flexibility. Unfortunately, standard mixture modeling algorithms do not provide a straightforward  
162 way to ensure that the model parameters ( $\mu$  and  $\sigma$ ) for the groups accord with our understanding of these  
163 parameters from reference populations. Bayesian inference, however, does provide a way to do this very  
164 thing by using data from reference populations to create prior distributions for mixture model parameters.  
165 Using prior distributions improves overall model performance because the analyst can use these sources of  
166 'prior knowledge' to inform them about the data that they have on-hand (Otárola-Castillo et al. 2022).

167 This paper describes a Bayesian approach to the mixture model analysis of faunal measurements that  
168 addresses these weaknesses of mixture modeling as currently applied. The model uses informative priors de-  
169 rived from a 'prior assemblage' of known age-at-death and sex individuals to constrain population parameter  
170 estimates to be biologically interpretable (Popkin et al. 2012). It also uses multilevel modeling to take ad-  
171 vantage of partial pooling and address aggregation issues to directly estimate parameters for each measured  
172 dimension in the analysis (Gelman 2006a; Wolfhagen 2020). In addition to modeling females and males,  
173 the model includes a third group consisting of "immature" specimens that died before reaching adult body  
174 size. The model also emphasizes inference of the entire assemblage rather than just the measured specimens  
175 by incorporating observations of the sex ratio (from morphological data) and the proportion of immature  
176 specimens (from fusion data) to inform population parameters of the proportions of these different groups.  
177 The model is used on sixteen simulated assemblages derived from the Popkin et al. (2012) Shetland sheep  
178 (*Ovis aries*) population to test its ability to accurately estimate the age and sex composition of assemblages.  
179 Two archaeological case studies then show the applicability of the model to archaeological assemblages for  
180 reconstructing the age and sex composition of assemblages and to highlight the importance of incorporating  
181 immature specimens into mixture modeling analyses.

182 **2. A BAYESIAN MULTILEVEL MIXTURE MODEL FOR ZOOARCHAEO-**  
183 **LOGICAL MEASUREMENTS**

184 The Bayesian model developed for this paper improves on standard mixture modeling for zooarchaeological  
185 measurements in four distinct ways. First, it addresses complications caused by measurements from unfused  
186 specimens and post-fusion growth by modeling three groups within the mixture: immature animals, (adult-  
187 sized) females, and (adult-sized) males, each with distinct size parameters. Second, the model's multilevel  
188 structure allows the model to balance bias due to aggregation and overfitting from small sample sizes.  
189 Second, the Bayesian foundation of the model provides an avenue for synthesizing information about the  
190 ontogenetic age and sex composition of the assemblage from non-metrical data (e.g., fusion rates, sex ratios  
191 based on morphological data) to inform the results of the mixture model. Finally, researchers can create  
192 prior distributions for mixture model parameters from prior assemblages or other sources, ensuring that the  
193 mixture model results are biologically interpretable. This section outlines these benefits; specific details of  
194 the model are described in a Model Supplement and in the analytical code (available at the project's GitHub  
195 page).

196 Observed measurements from different dimensions (e.g., humerus distal breadth “Humerus Bd,” radius  
197 proximal breadth “Radius Bp,” abbreviations following Driesch 1976) are first converted to logarithmic size  
198 index (LSI) values using a natural logarithm base to take advantage of the normalization LSI standardization  
199 provides (Meadow 1999; Wolfhagen 2020). The model uses a single LSI value per specimen, so any specimens  
200 with multiple observed dimensions are first summarized by estimating the mean LSI value from the observed  
201 dimensions; these specimen-level LSI values are the basis for the mixture model. These specimen-level LSI  
202 values can be clustered into different element portions—partial or complete elements that are the basic  
203 categorical unit of an analysis [e.g., “distal humerus” or “first phalanx”; compare to “skeletal part type” in  
204 Breslawski (2022)]. Table 2 provides a glossary of the key terms used in this text.

**Table 2:** Definitions of key terms used in this paper

Term	Definition
<b>Element Portion</b>	A complete or partial skeletal element defined by the zooarchaeologist, used as the foundation of the multilevel model (e.g., "distal humerus"). Model produces parameter estimates for all defined element portions, so element portions must be non-overlapping. Analogous to "skeletal part type" in Breslawski (2022).
<b>Dimension</b>	Specific type of observed measurement (e.g., "humerus distal breadth") on a specimen. Dimension definitions typically follow von den Driesch (1976).
<b>Measured Assemblage</b>	Assemblage of measured specimens from a defined number of element portions of a specific taxon.
<b>Modeled Assemblage</b>	Assemblage of specimens from a defined number of element portions of a specific taxon. Includes measured and non-measured specimens, though all element portions must have some number of measured specimens. Measurability is assumed to be effectively random (i.e., unrelated to whether the specimen came from an immature, female, or male individual).
<b>Full Assemblage</b>	Assemblage of specimens from a defined number of element portions of a specific taxon. Includes element portions that do not have any observed measurements. Measurability is assumed to be effectively random (i.e., unrelated to whether a specimen came from an immature, female, or male individual).

**2.1 Benefits of the Bayesian Multilevel Mixture Model**

Body size is affected by both ontogenetic age and sex; animals killed before reaching adult body size pose a complication for most sex determination models, which exclusively focus on distinguishing between (adult-size) female and male animals (but see Zeder and Lemoine 2020). Measurements from known age-at-death Shetland sheep show that specimens killed younger than one year of age are significantly smaller than those killed at older ages, regardless of fusion status and sex; after one year of age, size is no longer significantly impacted by age (Popkin et al. 2012). Thus, any measurement from an unfused epiphysis or from an element portion that does not fuse or exhibits significant post-fusion growth should be considered potentially immature and needs to be modeled with a three-member mixture model (Group 1 = immature, Group 2 = adult female, and Group 3 = adult males). On the other hand, the model excludes the possibility that

215 measurements from specimens that are conclusively not immature due to their fusion status could be from  
216 the immature group ( $\pi_1 = 0$ ), effectively fitting a two-member mixture model (adult females and adult  
217 males).

218 Typically, biometric analyses using LSI values aggregated across different element portions (e.g., Vigne  
219 2011; Sasson and Arter 2020; Arbuckle and Kassebaum 2021); this aggregation produces bias because  
220 it assumes that every element portion has the same mixture model parameter values (Wolfhagen 2020).  
221 Multilevel modeling uses partial pooling to allow the cluster-specific parameters to vary between clusters  
222 while reducing overfitting caused by small sample sizes (McElreath 2020; Fernée and Trimmis 2021). In the  
223 case of this mixture models, element portions are the relevant clusters—the multilevel model produces a set  
224 of mixture model parameters for each element portion (a set of  $\pi$ ,  $\mu$ , and  $\sigma$  parameters for each of the three  
225 animal groups). These cluster-specific parameters are related to each other through “hyper-parameters” that  
226 describe the average value of the mixture model parameters and the variability of model parameters across  
227 element portions (Wolfhagen 2020). This structure reduces overfitting caused by small sample sizes among  
228 some clusters while also avoiding the bias caused by aggregating all clusters together.

229 Of course, biometric data are not the only source of information on an assemblage’s ontogenetic age and  
230 sex composition. Fusion rates of elements that fuse around the age that animals reach adult body size can  
231 provide relevant information on the proportion of immature specimens in the assemblage (e.g., first and  
232 second phalanges in sheep: Popkin et al. 2012), just as sex ratios derived from morphologically distinct  
233 adult elements provide information about the adult sex ratio in an assemblage (e.g., fully fused pelvises:  
234 Stiner et al. 2022; horn cores: Twiss and Russell 2009). These estimates of assemblage composition do not  
235 supersede those produced by a mixture model, but they are also not irrelevant to the composition from a  
236 mixture model. Unlike other sex determination methods, the multilevel structure of the Bayesian multilevel  
237 mixture model allows the analyst to inform their model results with relevant fusion and morphological sex  
238 data from the assemblage. These data do not determine the proportion of immature animals and the adult  
239 sex ratio of the mixture model, but they do help the model make more precise estimates of the ontogenetic  
240 age and sex composition of the assemblage than possible with the measurement data alone.

241 Relevant information from a prior distribution can inform an analyst about reasonable values for model’s  
242 hyper-parameters, which can be summarized as prior distributions. Creating informed prior distributions  
243 allows the model to ensure that the hyper-parameters have biologically interpretable results. The multilevel  
244 structure of the model then ensures that mixture model parameters can vary between different element  
245 portions while still being informed by these hyper-parameters to maintain biological interpretability, even  
246 with small numbers of observations. Prior distributions draw explicit links between our sources of prior  
247 knowledge (e.g., reference populations, ethnographic data, ecological data, etc.) and our archaeological data.

248 Unlike absolutist models, we can define the prior distributions used in the Bayesian multilevel mixture model  
249 to be less specifically focused on the parameter values of the prior assemblage. Increasing the uncertainty of  
250 assemblage-derived prior distributions allows the mixture model to adjust to biometric differences between  
251 the prior assemblage and the assemblage being fit by the model. Care must still be taken to ensure that  
252 prior distribution definitions are at appropriate scales for the observations and not so broad as to include  
253 values that are known to be physically impossible (e.g., Gabry et al. 2019).

254 ***2.2 Developing Prior Distributions from a Prior Assemblage***

255 Prior distributions are central to Bayesian inference and describe one's prior beliefs in potential values of  
256 a model parameter. Prior distributions can be likened to a 'filter' from which parameter values are drawn  
257 to evaluate their fit with the data (Smith and Gelfand 1992). Several approaches exist for deciding how to  
258 describe this prior belief, ranging from 'objective' priors that provide equal weight to all possible values of  
259 a parameter to distinct distributions defined by a synthesis of previous or related research (Gelman 2006b).  
260 Objective priors poorly reflect our intuition about phenomena we are modeling, are inefficient, and can  
261 introduce errors into our analyses (Gabry et al. 2019); instead, 'weakly informative priors' or 'reference priors'  
262 use transformations of parameter values—like centering and scaling element portion-specific parameters—to  
263 describe variation in parameter values within reasonable values, with small deviations being more likely than  
264 large deviations (Gelman et al. 2008). Informative priors are derived from relevant knowledge, be it the  
265 results of earlier studies on the same subject, the quantification of expert opinion, or parameter values for  
266 related subjects (McCarthy and Masters 2005; Otárola-Castillo et al. 2022). Regardless of the distribution's  
267 source, it is important to evaluate how well the distribution reflects your prior knowledge about the system  
268 under study because the prior distributions influence the results of the analysis.

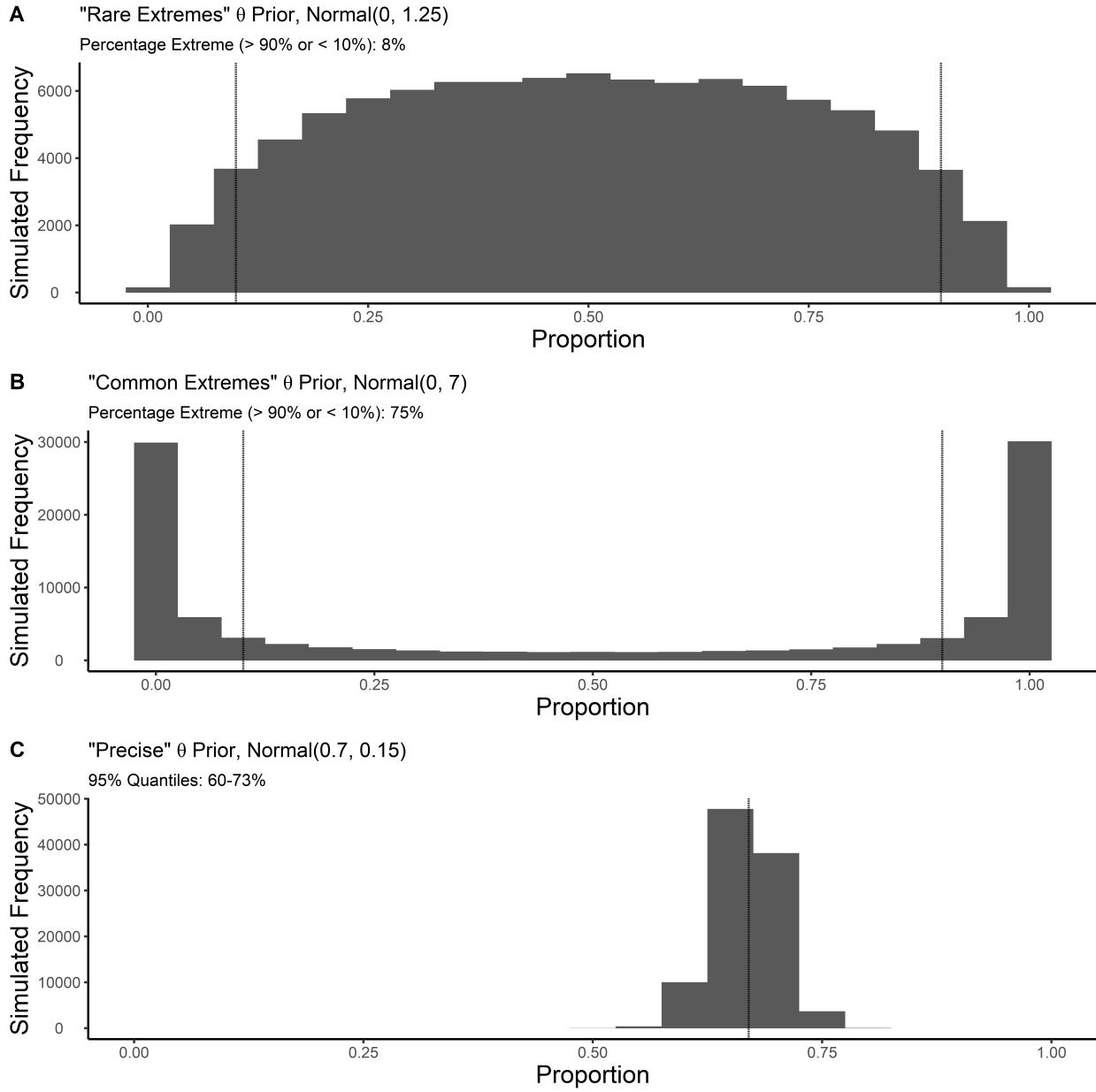
269 The mixture proportions summarize the composition of the assemblage and mediate the relative likeli-  
270 hoods of the different animal groups, adjusting a specimen's membership probabilities. Prior distributions  
271 for the mixture components reflect our prior beliefs about the relative proportions of immature, adult-sized  
272 females, and adult-sized males in the assemblage. Instead of estimating the prior belief for each of these three  
273 related categories, the model uses two prior distributions to estimate independent variables: the proportion  
274 of immature animals ( $\pi_1$ ) and the adult sex ratio estimated through the relative proportion of adult females  
275 ( $\frac{\pi_2}{\pi_2 + \pi_3}$ ) (see Model Supplement for more details). The following examples show some of the flexibility re-  
276 searchers have when describing their prior belief about the proportion of immature animals or the adult sex  
277 ratio in an assemblage; the Model Supplement shows the mathematical details necessary to create relevant  
278 prior distributions for a model.

279 Figure 2 shows three examples of prior distributions for one of these mixture proportion concepts (pro-

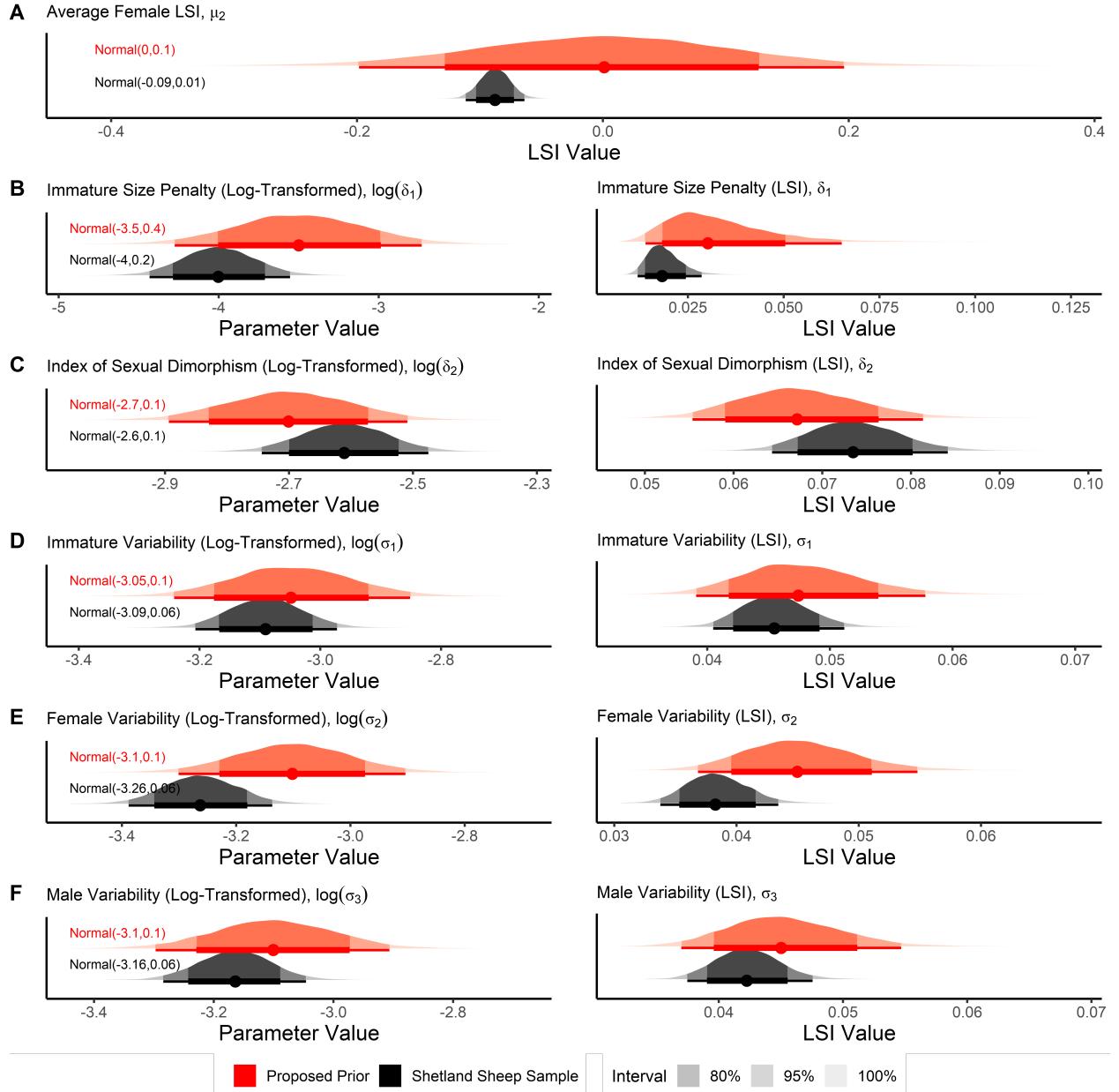
portion of immature or adult sex ratio) that reflect different expectations based on prior knowledge. Note that the Bayesian mixture model uses the observations of fusion rates and morphological sex ratios as observed data, so these prior distributions reflect knowledge prior to even those observations. Panel 2A shows a relatively broad, or uncertain, prior distribution where a researcher doesn't believe that the proportion is extreme (i.e., that the assemblage is either dominated by or bereft of immature animals or that the sex ratio is dominated by females or males) but has little opinion otherwise. Panel 2B shows a somewhat inverse situation, where the researcher is confident that the proportion (either the proportion of immature animals or the adult sex ratio) is at either one extreme or the other but isn't sure which extreme it is. Panel 2C displays a scenario where the researcher is confident that the proportion is centered around 67% before looking at the faunal data, presumably based on prior research or other contextual information.

Prior distributions for parameters governing average body size ( $\mu$ ) and size variability ( $\sigma$ ) are the keys to ensuring that the mixture model produces biologically feasible and interpretable results. These prior distributions are based on analysis of a 'prior assemblage' created by sampling immature, adult female, and adult male/castrate Shetland sheep from the Popkin et al. (2012) population (150 specimens for each animal group; see Model Supplement for more details). Castrates were considered males for the purposes of the model, as proximal and distal bone width measurements as a whole did not vary significantly between intact males and castrates (Popkin et al. 2012: 1783-1784). Modeling the average body size and size variability of these animal groups using a multilevel model created the starting point for prior distributions that could be used in the model. To generalize the prior distributions so that they are applicable to a variety of zooarchaeological scenarios, the results from the prior assemblage were given larger standard deviations to increase the uncertainty, which allows the model to better fit the data (see Figure 3 and Model Supplement for more details). The prior distributions act to prevent the model from accepting parameter values that are implausibly large or small given our prior knowledge about size variability and size differences between animal groups.

This section describes one approach for defining prior distributions of mixture models that are relevant for a wide range of zooarchaeological cases, particularly if researchers do not have strong preconceptions about the relevant parameters from prior research. It is important to remember that a model's prior distributions are choices made by the researcher to fit particular research questions, regardless of whether the distributions are informed by advice on reference priors, prior assemblages, or mathematical summaries of existing research. Zooarchaeological assemblages resulting from catastrophic kills would be expected to have a different ontogenetic age and sex composition compared to assemblages derived from sustained hunting or herding take-off (e.g., Stiner 1990). Similarly, other research contexts may provide an analyst with different prior expectations about animal size variability and overall biometry. Other reference populations,



**Figure 2:** Examples of prior distributions for demographic parameters (proportion of immature animals or adult sex ratio) with different expectations. A: A distribution where extreme values are considered unlikely but otherwise most values are about equally as likely. B: A distribution where extremely high or extremely low values are likely but indeterminate values are much less likely. C: A distribution where it is extremely likely that values are centered around 67%. The last scenario would only be appropriate if there is sufficient prior knowledge about the context.



**Figure 3:** Posterior distributions of model hyper-parameters from a sample of known-identity sheep specimens (red) and proposed prior distributions for mixture model applications (black). Prior distributions describe (A) average female body size, in LSIE units; (B) the size difference between average immature and female animals, in log-transformed (left) and LSIE (right) units; (C) the index of sexual dimorphism—the size difference between average female and male animals, in log-transformed (left) and LSIE (right) units; size variability for (D) immature, (E) female, and (F) male animals, in log-transformed (left) and LSIE (right) units. Proposed prior distributions provide useful baseline in the absence of relevant biometric information regarding sexual dimorphism and size variability.

313 particularly those from other taxa, could also be used to create prior assemblages and help determine the  
314 limits on biological feasibility. Researchers could and should adapt their prior distributions to best reflect  
315 their intuition about likely parameter values for their research context. Regardless of the prior distributions  
316 a researcher uses, it is crucial to formally describe the prior distributions that is used in a Bayesian analysis  
317 to ensure replicability. Further, researchers should examine the implications of different candidate prior  
318 distributions while developing a Bayesian model to test a research question; prior distributions should be  
319 regularly tested even before models are fit to datasets (Gelman, Carlin, et al. 2020).

320 **2.3 Extending the Multilevel Analysis to Multiple Sites**

321 The multilevel structure of the model that allows parameters to vary across element portions can also be used  
322 to extend the modeling approach to examine multiple assemblages at once. Combining multiple assemblages  
323 into a single model allows researchers to investigate regional variation in herd management strategies or  
324 outline diachronic trends in body size that may relate to population turnover (e.g., Arbuckle and Atici 2013;  
325 Arbuckle et al. 2016). It also allows researchers to model diachronic changes over the course of a multi-  
326 period site's occupation, as each occupation layer can be defined as a separate assemblage (e.g., Hongo et al.  
327 2009; Wolfhagen et al. 2021). By including the assemblages in the same model, estimates share the same  
328 hyper-parameters, which improves the precision of these estimates and allows researchers to directly compare  
329 assemblage-specific parameters by using contrasts. Further, adopting this structure provides the foundation  
330 for more sophisticated analyses that test specific hypotheses about variation in biometric or compositional  
331 parameters, such as spatiotemporal autocorrelation in body size.

332 An important consequence of extending the model to evaluate multiple sites at once is that the interpreta-  
333 tion of the overall hyper-parameters that the researcher inputs into the model changes. Instead of describing  
334 the overall estimates for a specific assemblage, these hyper-parameters now describe a 'grand mean' of the  
335 parameter value for all the included assemblages. These overall summaries could be interpretively useful if,  
336 for instance, all the assemblages come from a discrete archaeological culture or region. In other scenarios,  
337 however, the interpretation of these overall hyperparameters may be less meaningful than comparisons of  
338 assemblage-specific estimates that still account for anatomical variation within each assemblage (see Model  
339 Supplement for more details).

340 **2.4 Interpreting Model Results: Measured, Modeled, and Full Assemblages**

341 The results of the Bayesian multilevel mixture model include specimen-specific membership probabilities  
342 ( $\pi_{\text{Specimen}}$ ) based on the mixture model parameters. While these membership probabilities can be used to  
343 calculate "critical size limits" where the largest membership probability shifts from one group to another (e.g.,

344 Monchot and Léchelle 2002), they can also be used to simulate assemblages of known-group specimens to  
345 examine age/sex-stratified estimates of body part representation and sex-stratified fusion rates. Membership  
346 probabilities ( $\pi_{\text{Specimen}}$ ) are used to simulate the specimen's identity by sampling from the probabilities using  
347 a multinomial distribution; in each posterior sample, a single simulated assemblage is created, resulting in a  
348 distribution of simulated assemblages with known age/sex assignments (Crema 2011). The characteristics of  
349 these assemblages can then be used to summarize the overall assemblage or identify differences in composition  
350 based on element types, fusion states, sub-assemblage features, or other pertinent factors that a researcher  
351 is interested in examining in relation to the composition of the assemblage.

352 The usual goal of a mixture model analysis—like any sex determination analysis—is to estimate the  
353 composition of the *entire* (or *modeled*) faunal assemblage, rather than just the *measured* assemblage used  
354 by the analyst. Typical analyses elide these differences, smoothly translating the results of an analysis on a  
355 measured assemblage (i.e., the sex ratio) to describe the entire assemblage. Sometimes disparate results from  
356 different element portions require explanation, such as different butchery strategies for males and females  
357 (e.g., Speth 1983), but even in these cases the results from measured specimens are used to describe the  
358 entire set of bones from the same element portion. This elision creates a bias by ignoring the existence of  
359 unmeasured specimens in the assemblage and presents an interpretive dilemma for researchers, whose only  
360 recourse if they are unwilling to make this elision is to discount the model results as unrepresentative.

361 We can avoid this bias by formalizing the relationship between the *measured* and *modeled* assemblages  
362 by stating that the measured assemblage is a sample of the modeled assemblage, wherein inclusion is gov-  
363 erned by a specimen's *measurability*—the preservation of specific bony portions that allow for a biometric  
364 measurement. If we assume that measurability is unrelated to a specimen's ontogenetic age or sex, then  
365 we can assume that the measured assemblage is a random sample of the modeled assemblage. Thus, an  
366 unmeasured specimen will have the same model parameters (mixture proportions  $\pi$ , average size  $\mu$ , and  
367 size variability  $\sigma$ ) of the measured specimens from the same element portion. Crucially, this means that  
368 we can include unmeasured specimens in our simulated assemblages by using the relevant mixture propor-  
369 tions  $\pi$  (adjusted for the specimen's fusion data as necessary) as that specimen's membership probabilities  
370 ( $\pi_{\text{Specimen}}$ ). Leveraging the multilevel structure of the model further, we can assert that the overall mixture  
371 model hyper-parameters for *modeled* element portions are equally valid for *unmodeled* element portions. The  
372 Bayesian multilevel mixture model estimates hyper-parameters that describe the average value ( $\mu_{\text{Element}}$ )  
373 and expected variability ( $\sigma_{\text{Element}}$ ) of mixture model parameters element portions; these hyper-parameters  
374 can be used to estimate the relevant mixture model parameters of an unobserved element portion (Gelman,  
375 Carlin, et al. 2020; McElreath 2020). The resulting parameters, then, could be used to estimate  $\pi_{\text{Specimen}}$   
376 membership probabilities for the unmodeled (and unmeasured) specimens, as in the first extension, creating

377 an estimate of the composition of the *full* assemblage.

378 At first blush, these extensions may seem like a departure from concrete results of a mixture analysis  
379 into proxy-upon-proxy esoterica. However, by formalizing the relationship between what data are in the  
380 mixture analysis (the measured assemblage) and what data we are interested in describing (the modeled  
381 or full assemblage), these extensions are critical for creating a principled interpretation of an assemblage  
382 based on the analysis' results. Mixture analyses are based on the measurable sample of specimens from  
383 the modeled subset of all element portions; this does not mean that these results cannot produce useful  
384 information, but it does mean that we must contextualize those results by understanding how small the  
385 measured assemblage is in comparison with the modeled (or full) assemblage we are interested in describing.  
386 These extensions provide a way to do this—measured specimens will have much more certain membership  
387 probabilities than unmeasured or unmodeled specimens, owing to the information gained from its size. Thus,  
388 including unmeasured and unmodeled specimens will produce less ‘extreme’ results (e.g., a lower probability  
389 that a majority of the assemblage is from a single group). This will be especially clear when the measurement  
390 assemblage is much smaller than the modeled or full assemblage.

### 391 ***2.5 Computational Details of the Bayesian Analysis***

392 The Bayesian multilevel mixture model is written in Stan, version 2.30.1 (Team 2022). All analyses in this  
393 paper use R version 4.1.3 (2022-03-10), in Rstudio 2022.12.0.353 (Elisabeth Geranium) (R Core Team 2022;  
394 RStudio Team 2022); Table 3 lists the packages, versions, and citations for the packages used in the analytical  
395 scripts. The model Stan code and analytical R code necessary to replicate and apply the analyses in this  
396 paper are freely available in a GitHub page and Open Science Framework page. The files include a copy of  
397 the Shetland sheep data file from the supplemental files published in Popkin et al. (2012) and archaeological  
398 datasets for the case studies downloaded from OpenContext (Buitenhuis 2013; Carruthers 2006; Galik 2013;  
399 Gourichon and Helmer 2013). The analytical code includes two script files—a script for replication and one  
400 for application. The R markdown file (“ZooarchMixMod.Rmd”) file replicates the entire analytical workflow  
401 of the paper, with a specific seed set to ensure exact replicability of the submitted manuscript. Another set  
402 of scripts to standardize the analytical workflow for faunal datasets structured like the OpenContext faunal  
403 datasets used in these case studies, see the GitHub for more details. All scripts (R and Stan) are released  
404 under the MIT license and figures are released as CC-BY to encourage reuse and reproducibility (Marwick  
405 2017; Marwick and Pilaar Birch 2018).

**Table 3:** Software packages used in the analytical script for this paper.

Use	Package	Version	Citation
Data aggregation, analysis, and multi-core processing	boot	1.3.28	Canty & Ripley, 2021; Davison & Hinkley, 1997
Data aggregation, analysis, and multi-core processing	data.table	1.14.2	Dowle & Srinivasan, 2019
Data aggregation, analysis, and multi-core processing	readxl	1.4.1	Wickham & Bryan, 2022
Data aggregation, analysis, and multi-core processing	parallel	4.1.3	Microsoft Corporation & Weston, 2022
Data aggregation, analysis, and multi-core processing	doParallel	1.0.17	Microsoft Corporation & Weston, 2022
Creation of RMarkdown files	kableExtra	1.3.4	Zhu, 2021
Creation of RMarkdown files	knitr	1.39	Xie, 2022
Bayesian analysis and summarization	cmdstanr	0.5.3	Gabry & Češnovar, 2022
Bayesian analysis and summarization	rstan	2.21.5	Stan Development Team, 2022
Mixture model analysis and standard animal measurements	mixtools	1.2.0	Benaglia, et al., 2009
Mixture model analysis and standard animal measurements	zoolog	1.0.1	Pozo, et al., 2021
Visualization	Cairo	1.6.0	Urbanek & Horner, 2022
Visualization	ggplot2	3.3.6	Wickham, 2016
Visualization	ggdist	3.2.0	Kay, 2022
Visualization	ggpubr	0.4.0	Kassambara, 2020
Visualization	ggrepel	0.9.1	Slowikowski, 2021
Visualization	rnaturrearth	0.1.0	South, 2017a
Visualization	rnaturrearthdata	0.1.0	South, 2017b
Visualization	sf	1.0.8	Pebesma, 2018

### 406 3. TESTING THE BAYESIAN MULTILEVEL MIXTURE MODEL

407 Two sets of tests are used to evaluate different aspects of the Bayesian multilevel mixture model. First,  
408 the accuracy of the model’s ability to reconstruct the age and sex composition of assemblages is tested by  
409 using simulated faunal assemblages of known age and sex from the Shetland sheep population. This test  
410 evaluates both the single-assemblage model and the multi-assemblage model. Second, two archaeological  
411 case studies showcase the applicability of the model to archaeological data and the added insights gained  
412 from adopting Bayesian multilevel mixture models. The simulated assemblage case study and the single  
413 assemblage archaeological case study use sheep (*Ovis aries*) measurements, with standard measurements  
414 coming from a female wild sheep (*Ovis orientalis* FMC 57951: Uerpmann and Uerpmann 1994: Table 12).  
415 The multiple assemblage case study uses cattle (*Bos taurus*) measurements, with standard measurements  
416 coming from a wild female aurochs (*Bos primigenius* “Ullerslev”: Degerbøl 1970). Two dimensions of the  
417 standard cow (Scapula GLP: 89 mm; and Calcaneus GB: 46 mm) were not included in the ‘zoolog’ output  
418 and were included manually, drawn from the referenced source.

**Table 4:** Measurements included in the simulation analyses. Dimension definitions follow von den Driesch (1976)

Element Portion	Dimension
Scapula	GLP
Humerus	Bd
Humerus	BT
Radius	Bp
Radius	Bd
Metacarpus	Bp
Metacarpus	Bd
Femur	Bd
Tibia	Bd
Astragalus	Bd
Metatarsus	Bp
Metatarsus	Bd

*419* **3.1 Simulated Assemblages**

*420* A series of simulated assemblages of known age and sex composition are created from the Shetland sheep  
*421* population by randomly drawing element portions (and all associated measurements) from the total assem-  
*422* blage without replacement. Table 4 describes the measured dimensions included in the simulation analyses  
*423* from the 10 element portions. The first test, using a single-assemblage model, uses 150 element portions from  
*424* the Shetland sheep population where every element portion has an equal probability of being selected. There  
*425* is no guarantee, however, that the element portions have equal representation or even that all element por-  
*426* tions are present in the simulated assemblage, which better approximates archaeological assemblages. The  
*427* result of this first simulation produces an assemblage of 231 measurements from 125 individual animals. The  
*428* second test creates 15 simulated assemblages using the same procedure that are analyzed in a single multi-  
*429* assemblage model. To further test the model's flexibility, these assemblages vary in sample size and some  
*430* are manipulated to vary in average body size and expected composition from the original Shetland sheep  
*431* population. Demographic observations for phalanx fusion rates and pelvis sex ratios were also simulated  
*432* from the Shetland sheep population using the same underlying probabilities as the measurement assem-  
*433* blages. Table 5 describes the sample sizes of the measurement assemblages, including any manipulations to  
*434* the measurement values. The specific elemental composition and measurements of the assemblages, along  
*435* with the simulated demographic observations, used in both simulations can be recovered from the replication  
*436* script with the recorded random seed (see also Supplemental Tables 1-3); using another random seed would  
*437* provide a conceptual replication of new assemblages drawn from the same underlying populations.

**Table 5:** Group composition of the simulated measurement assemblages (element portions)

Assemblage	Demographics	Size	Immature	Female	Male	Total
Single Assemblage	13% Immature, 46% Female, 40% Male	1.00	23	80	47	<b>150</b>
Site 01	13% Immature, 46% Female, 40% Male	1.00	2	11	17	<b>30</b>
Site 02	13% Immature, 46% Female, 40% Male	1.00	3	3	4	<b>10</b>
Site 03	13% Immature, 46% Female, 40% Male	1.20	4	13	13	<b>30</b>
Site 04	13% Immature, 46% Female, 40% Male	0.80	3	11	16	<b>30</b>
Site 05	13% Immature, 46% Female, 40% Male	1.20*	4	13	13	<b>30</b>
Site 06	20% Immature, 70% Female, 10% Male	1.00	5	24	1	<b>30</b>
Site 07	20% Immature, 70% Female, 10% Male	1.00	0	8	2	<b>10</b>
Site 08	20% Immature, 70% Female, 10% Male	1.20	4	23	3	<b>30</b>
Site 09	20% Immature, 70% Female, 10% Male	0.80	4	25	1	<b>30</b>
Site 10	20% Immature, 70% Female, 10% Male	1.20*	11	17	2	<b>30</b>
Site 11	5% Immature, 35% Female, 60% Male	1.00	2	10	18	<b>30</b>
Site 12	5% Immature, 35% Female, 60% Male	1.00	0	3	7	<b>10</b>
Site 13	5% Immature, 35% Female, 60% Male	1.20	1	17	12	<b>30</b>
Site 14	5% Immature, 35% Female, 60% Male	0.80	3	11	16	<b>30</b>
Site 15	5% Immature, 35% Female, 60% Male	1.20*	2	12	16	<b>30</b>

*Note:*

Demographics in the Single Assemblage and Sites 01-05 reflect original Shetland sheep composition

\* Size increased for males only

Rather than trying to reconstruct the exact parameter values of the simulated assemblages, parametric accuracy is focused on relating the parameter distributions of the assemblage (the sample) to the respective values in the full Shetland sheep assemblage (the population from which the sample is derived), including any relevant demographic or size modifications. In this sense, the goal is not 100% accuracy: instead, the goal is being well-calibrated, wherein credible intervals about a sample parameter contain the true population values the specified percentage of the time (e.g., 95% of a model's 95% credible intervals contain the true population values). If too few population values are contained in the interval statements, then the model has overfit to the sample and the posterior distributions are too narrow: a researcher may falsely distinguish between two assemblages from the same underlying population (i.e., a false positive). If too many population values are contained in the interval statements, then the model has underfit to the sample and the posterior distributions are too wide: a researcher may be unable to distinguish between two assemblages that derive from different underlying populations (i.e., a false negative).

Compositional accuracy does not have the same structure as parametric accuracy because there is no underlying population value for composition: there is only the true number of immature, adult female, and adult male specimens in the measured and modeled assemblages. Again, though, it is important to understand accuracy in the context of overfitting and underfitting. Overfitted results, wherein credible intervals about the number of immature, adult female, and adult male specimens contain the true abundances

455 at a lower rate than designed (e.g., fewer than 95% of the 95% credible intervals), could lead to a researcher  
456 declaring an imbalance in the demographic composition of an element where one does not exist (or is even  
457 imbalanced in the opposite direction). Underfitted results, by contrast, would mean that a researcher is  
458 unable to identify an imbalance where one exists because the credible intervals are too wide. It is important,  
459 then, to use the simulations to understand the kinds of errors the model is prone to making so that researchers  
460 avoid overinterpretation.

461 Modeled assemblages were created for the single-assemblage simulation and the multi-assemblage sim-  
462 ulation by assuming that measured specimens represent 20% of the overall assemblage and sampling more  
463 specimens from the Shetland sheep population to create the remaining 80% of the assemblage. For example,  
464 in the single-assemblage simulation with 150 measured specimens, this means sampling 600 more specimens  
465 from the Shetland sheep population to create a total modeled assemblage of 750 specimens. Specimens could  
466 not be repeatedly sampled, though multiple specimens could be from the same individual. As described in  
467 Section 2.5, unmeasured specimens use the relevant  $\pi$  parameters for the element portion. For the multi-  
468 site simulation, this potentially includes element portions where there are no relevant measurements in the  
469 specific assemblage.

470 Because the “grand mean” parameters in the multisite simulation no longer represent the same thing  
471 as in the single assemblage model (see Model Supplement), the prior distributions must also be changed to  
472 reflect different expectations. Again, the goal of these prior distribution definitions is to prevent extreme  
473 overfitting so that parameter estimates are biologically feasible (Gelman et al. 2008). In general, the centers  
474 of the distributions stayed the same, but the uncertainty was increased to reflect the fact that there’s  
475 less certainty about biologically feasible values for multiple populations, especially if there is size variation  
476 expected between the assemblages. These prior distributions are listed in the Model Supplement.

477 ***3.2 Archaeological Case Studies***

478 The Bayesian multilevel mixture model is applied to two archaeological case studies to showcase the utility  
479 of the model for both interpreting a single assemblage and examining multiple assemblages. In both case  
480 studies, the sheep and cattle measurements have been previously published on OpenContext and the general  
481 zooarchaeological summaries of the assemblages have been published, as well (Buitenhuis 2008, 2013; Car-  
482 ruthers 2005, 2006; Galik 2013; Gerritsen and Özbal 2019; Gourichon and Helmer 2008, 2013). Again, LSI<sub>e</sub>  
483 values are calculated using the same standard animal as the simulation analysis for the single assemblage  
484 analysis, the *Ovis orientalis* female standard animal (FMC 57951) from Uerpmann and Uerpmann (1994,  
485 Table 12) and the *Bos primigenius* female standard animal (“Ullerslev”: Degerbøl 1970; Grigson 1989),  
486 operationalized through ‘zoolog’ functions (Pozo et al. 2021). Alongside metric data, the OpenContext

487 faunal tables provide demographic data that can be used to observe relevant estimates of the age and sex  
488 composition of the assemblages. The goal of applying the mixture model to these assemblages, then, is to  
489 use the metric data to improve these estimates of the age and sex composition of the assemblage, biometric  
490 estimates, and sex-specific fusion rates.

491 **3.2.1 Single Assemblage: Biometric Analysis of Sheep from 7th Millennium BCE Central**  
492 **Anatolia (Pinarbaşı B)**

493 The site of Pinarbaşı, located in the Konya Plain of central Turkey, consists of a series of rock shelter and  
494 open-air sites at the foothills of the Karadağ volcanic region and Lake Hotamış and its associated wetlands  
495 (Baird et al. 2011; Kabukcu 2017). This case study examines the Pinarbaşı B late Neolithic occupation,  
496 which is dated to the second half of the 7th millennium BCE and includes a large number of domesticated  
497 sheep and goat remains (Baird et al. 2011; Carruthers 2005). Carruthers (2005) analyzed fauna from the  
498 1994-1995 excavations by Trevor Watkins (Watkins 1996), interpreting the presence of fetal sheep remains  
499 and other juvenile remains in the assemblage as evidence for herders penning sheep on-site. The Neolithic  
500 assemblage was thus described as the result of seasonal occupation by sheep and goat herders during the  
501 lambing season and the fall, with culling in the spring possibly focused on young males (Carruthers 2005).  
502 This analysis makes several claims that can be evaluated with the Bayesian multilevel mixture model: the  
503 dominance of immature remains and a female-dominated adult sex ratio.

504 The Bayesian multilevel mixture model for the late Neolithic Pinarbaşı B assemblage uses 44 sheep  
505 measurements from 44 specimens (see Table 6). In addition to these measurements, the observed proportion  
506 of immature animals from unfused first and second phalanges is 59 / 62 (95%), including specimens identified  
507 to sheep and to sheep/goat. There are 0 observed sheep (or sheep/goat) pelvis bones with sex identifications;  
508 this is entered into the model by having an observed adult sex ratio for the assemblage of 0 / 0 (females /  
509 females + males). All data come from the Pinarbaşı faunal assemblage uploaded to OpenContext, focusing  
510 only on specimens in the Site B Neolithic contexts (Carruthers 2006). The Pinarbaşı B sheep model uses  
511 the same prior distribution definitions for the model hyper-parameters as the single assemblage simulation  
512 since both models, even though the sheep body sizes likely differ between the two populations, showcasing  
513 the flexibility of the standard prior distribution definitions.

514 **3.2.2 Multiple Assemblages: Biometric Analysis of Cattle from 7th-6th Millennium BCE**  
515 **Northwest Anatolia (Barçın Höyük, İlipınar Höyük, Menteşe Höyük)**

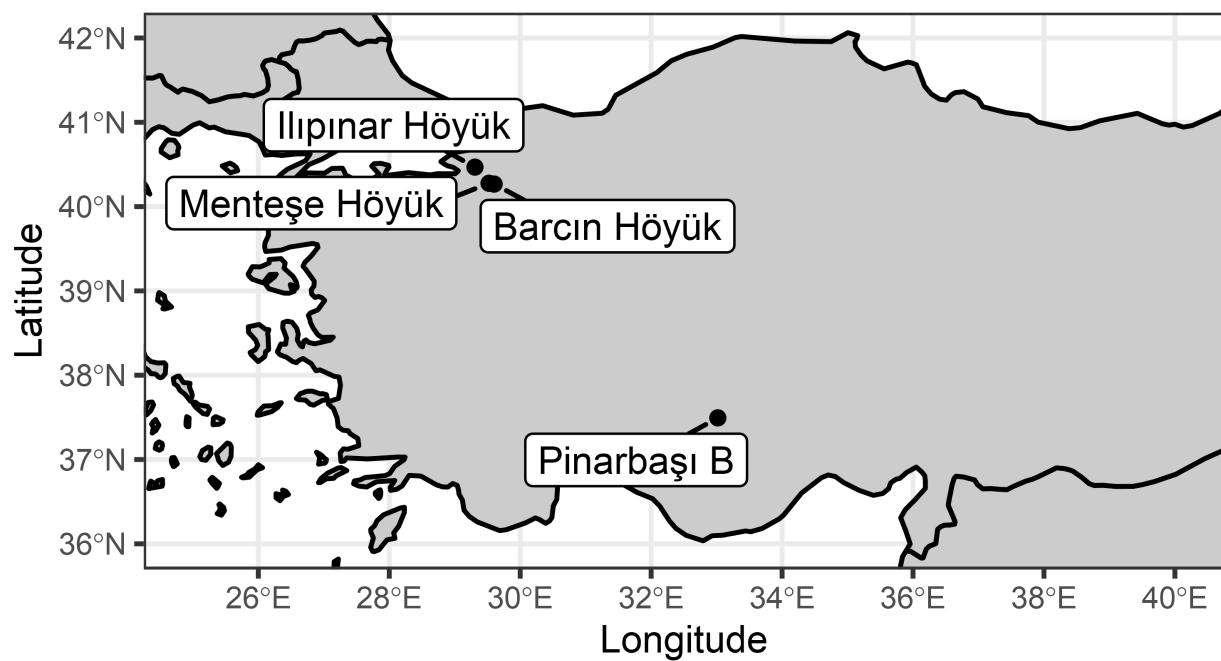
516 Understanding the development of Neolithic communities in northwestern Anatolia has long been of interest  
517 for researchers interested in studying the spread of agricultural lifeways from southwest Asia into Europe (e.g.,

**Table 6:** Elemental composition of the Pinarbaşı B assemblage. Dimensions definitions follow von den Driesch (1976).

Element Portion	Measurement	N
Astragalus	Bd	10
Calcaneus	GB	9
Humerus	Bd	1
Metacarpal (Distal)	Bd	2
Metatarsal (Distal)	Bd	4
Metatarsal (Proximal)	Bp	2
First Phalanx	Bp	9
Radius (Distal)	Bd	2
Tibia (Distal)	Bd	3
Tibia (Proximal)	Bp	2

Çakırlar 2013; Karul 2019; Özdoğan 2011, 2019). Agricultural communities first appear in the Marmara region in the mid-seventh millennium BCE in sites like Barçın Höyük (Gerritsen and Özbal 2019; Karul 2019). The domestic animal economies of these Late Neolithic and Early Chalcolithic communities appears to be focused on cattle and caprine (sheep and goat) herding, rather than pig husbandry (Buitenhuis 2008; Çakırlar 2013; Gourichon and Helmer 2008). Milk residues on pottery recovered from these sites suggest that these communities regularly consumed milk, potentially orienting herd management strategies of sheep, goats, and particularly cattle to specialize in milk production (Evershed et al. 2008; Thissen et al. 2010).

Four archaeological components from three sites are used in this case study, located near Lake İznik and on the Yenişehir Plain in the Bursa province of Turkey (Figure 4). The Neolithic layers from Barçın Höyük (Phase VI) is the earliest of these assemblages, with occupation roughly from 6500-6000 cal BCE; excavations revealed a subsistence economy focused on cereal agriculture and the herding of cattle, sheep, and goat (Galik 2013; Gerritsen and Özbal 2019). Menteşe Höyük is located approximately five km west of Barçın Höyük on the Yenişehir Plain; the three Neolithic layers at the site date to 5800-5600 cal BCE (Gourichon and Helmer 2013; Roodenberg et al. 2003). Previous faunal analysis of the Neolithic assemblage identified animal economies that shifted from predominantly cattle to sheep herding over the course of the occupation (Gourichon and Helmer 2008). İlipınar Höyük is located near Lake İznik, separated from the Yenişehir Plain by a mountain ridge (Roodenberg 2012a). The Neolithic/Early Chalcolithic occupation of the site spanned 6200-5400 cal BCE (Buitenhuis 2013); the assemblage is into two sub-assemblages (Neolithic İlipınar = Phases X-VII, 6000-5700 cal BCE; Chalcolithic İlipınar = Phases VI-V, 5600-5400 cal BCE), marked by the introduction of mudbrick architecture and expanded storage (Roodenberg 2012a, 2012b). Sheep and goat are common in the earlier assemblages of the site, with cattle becoming predominate in later phases of the site (Buitenhuis 2008; Roodenberg 2012a). Notably for this biometric analysis, Buitenhuis (2008) notes that cattle body sizes are stable throughout site's occupation.



**Figure 4:** Map of archaeological sites included in this analysis

541 The northwest Anatolian cattle bone assemblages consist of 614 measured specimens spread unevenly  
 542 across the four components (Barçın Höyük N = 67, Menteşe Höyük N = 45, Neolithic İlipinar N = 249, Chal-  
 543 colithic İlipinar N = 253). All measured Bos remains were included in the analysis, rather than separating out  
 544 those identified as aurochs (*Bos primigenius*, N = 3) or identified only to *Bos* spp. (N = 134) in the İlipinar  
 545 Höyük dataset; all specimens were only labeled as “Bos” in the Menteşe Höyük dataset. Table 7 shows the  
 546 composition of the four northwest Anatolian measurement assemblages. Demographic observations of the  
 547 proportion of immature animals and the adult sex ratio for each assemblage describe these assemblage-level  
 548 parameters. For the four northwest Anatolian assemblages, estimates of the assemblage-level proportion of  
 549 immature specimens based on the fusion rates of proximal and middle phalanges for cattle specimens are 28  
 550 / 87 (32%) for Barçın Höyük, 28 / 184 (15%) for Neolithic İlipinar, 8 / 25 (32%) for Menteşe Höyük, and 9 /  
 551 89 (10%) for Chalcolithic İlipinar. The observed adult sex ratios (females / females + males) based on cattle  
 552 pelvis morphology are 3 / 4 (75%) for Barçın Höyük, 0 / 0 for Neolithic İlipinar, 0 / 0 for Menteşe Höyük,  
 553 and 3 / 5 (60%) for Chalcolithic İlipinar. As in the Pinarbaşı B example, observations of 0 / 0 impart no  
 554 information onto the prior distribution of the adult sex ratio. All demographic and measurement data come  
 555 from the OpenContext datasets (Buitenhuis 2013; Galik 2013; Gourichon and Helmer 2013); the “R script  
 556 for replication” includes the steps for data processing and analysis.

**Table 7:** Elemental composition of the Northwest Anatolian cattle measurement assemblages.

Element Portion	Barçın Höyük	İlipinar Höyük (Late Neolithic/Transitional)	İlipinar Höyük (Early Chalcolithic)	Menteşe Höyük
Astragalus	5	19	14	15
Calcaneus	4	13	13	1
Femur (Distal)	0	5	3	0
Femur (Proximal)	0	4	4	0
Humerus	0	17	39	2
Metacarpal (Distal)	2	6	28	0
Metacarpal (Proximal)	3	19	9	3
Metatarsal (Distal)	5	6	16	2
Metatarsal (Proximal)	4	10	10	1
First Phalanx	16	76	35	8
Second Phalanx	21	49	29	14
Radius (Distal)	1	8	22	0
Radius (Proximal)	2	10	35	0
Scapula	0	9	14	0
Tibia (Distal)	4	9	21	0
Tibia (Proximal)	0	2	4	0
<b>Total</b>	<b>67</b>	<b>262</b>	<b>296</b>	<b>46</b>

557 Previous syntheses of the Late Neolithic and Early Chalcolithic animal economies in northwest Anatolia  
 558 provide several prior inferences about the age and sex structure of cattle bone assemblages that can be  
 559 evaluated with the results of the Bayesian multilevel mixture model. First, the general cultural continuity  
 560 of the assemblages suggests that the biometry and composition of cattle bone assemblages may be similar  
 561 at the sites, having been produced by similar processes (Çakırlar 2013; Özdoğan 2019); Buitenhuis (2008,  
 562 312) explicitly states that there is no size change among cattle bones across the İlipinar assemblage. Second,

563 the widespread evidence of milk consumption from pottery residue analyses from these sites and others in  
564 the region (Evershed et al. 2008; Thissen et al. 2010) has led some researchers to argue that cattle were  
565 managed for milk production (Gourichon and Helmer 2008; Roodenberg 2012a). Gourichon and Helmer  
566 (2008, 440) argue that the cattle tooth eruption and wear data at Menteşe indicate exploitation focused on  
567 milk consumption; one consequence of this pattern should be female-dominated adult sex ratios, including  
568 higher fusion rates for later-fusing elements among females than males (Zeder and Hesse 2000). The multilevel  
569 modeling results can be used to evaluate the feasibility of these inferences by examining posterior distributions  
570 of relevant parameters and simulations of sex-specific fusion rates.

571 Because this application is a multisite model and deals with a different taxon than the original simulations,  
572 the prior distributions for the model hyper-parameters are redefined to reflect different expectations of  
573 biological feasibility. While the multisite simulation provides useful prior distribution definitions for most  
574 of the parameters, two other parameters (average body size of females  $\mu_2$  and index of sexual dimorphism  
575  $\log_e(\delta_2)$ ) should be further changed because of different expectations modeling cattle rather than sheep.  
576 The change in the prior distribution definition of  $\mu_2$  reflects the fact that the standard measurements for  
577 cattle come from an aurochs female (Degerbøl 1970), which is expected to be larger than the domestic cattle  
578 females in the assemblages. Cattle are expected to be more sexually dimorphic than sheep, which is reflected  
579 in increasing the average expected value of  $\log_e(\delta_2)$ , resulting in an expectation of 0.14 LSI<sub>e</sub> units between  
580 males and females on average. This is slightly lower than the index of sexual dimorphism seen in the Degerbøl  
581 (1970) aurochs specimens [Grigson (1989): Figure 2, which uses LSI<sub>10</sub>; the equivalent size difference is 0.06  
582 on the LSI<sub>10</sub> scale], though domestic cattle may be expected to be less sexually dimorphic than their wild  
583 counterparts (e.g., Tchernov and Horwitz 1991); these prior distribution definitions are listed in the Model  
584 Supplement.

## 585 4. RESULTS

### 586 4.1 Simulated Assemblages: testing model accuracy

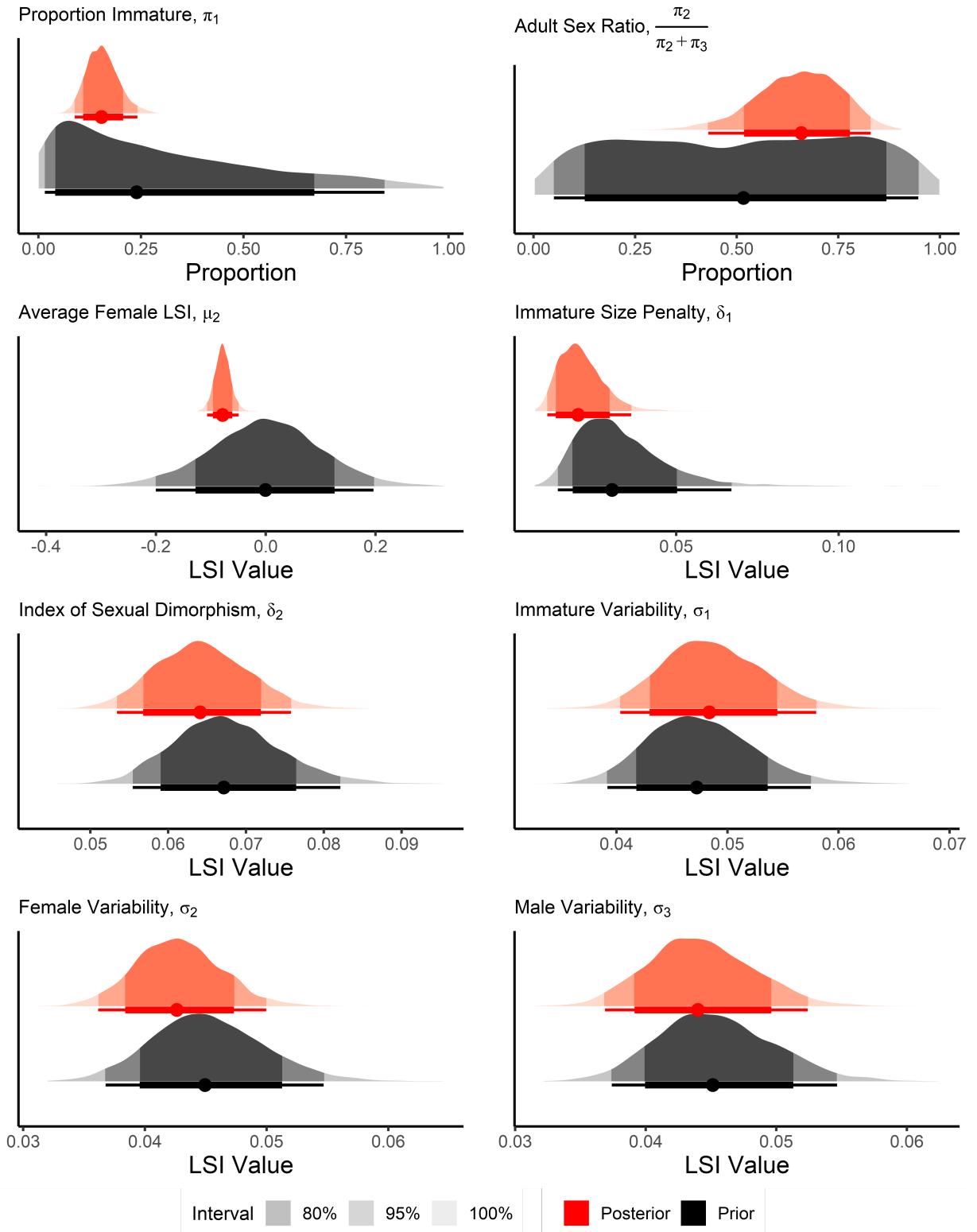
587 Bayesian models work by updating prior information with new data to produce posterior distributions of  
588 parameters of interest (Otárola-Castillo et al. 2022). Thus, the difference between a model parameter's  
589 prior and posterior distribution shows the amount that the model "learns" from the data. If the data  
590 do not provide relevant information on a parameter's potential values, then the posterior distribution will  
591 resemble the prior distribution. Figure 5 compares the prior and posterior distributions of the main model  
592 hyper-parameters for the single assemblage simulation. The results show that the data provides much more  
593 information about the likely values of the two demographic parameters (the proportion of immature animals,

594  $\pi_1$ , and the adult sex ratio,  $\frac{\pi_2}{\pi_2 + \pi_3}$ ) and the average body size for females ( $\mu_2$ ). This is largely to be expected,  
595 as the prior distribution definitions were weakly-informative priors (Gelman et al. 2008), but also shows how  
596 these choices did not appear to severely influence the resulting posterior distributions.

597 The prior distribution definitions for the size offsets ( $\delta_1$  and  $\delta_2$ ) and the size variability estimates ( $\sigma_1$ ,  $\sigma_2$ ,  
598 and  $\sigma_3$ ) have a lot more overlap between the prior distributions and their respective posterior distributions.  
599 This overlap stresses the importance of using a Bayesian framework, particularly one relying on informative  
600 prior distributions, to produce meaningful parameter estimates from zooarchaeological data. But it also  
601 highlights the interpretive weight given to the reference population. However, the overlap is not necessarily  
602 a drawback of the model, as again the prior distribution definitions were designed as informative priors,  
603 specifically to ensure that the resulting parameter estimates would be biologically feasible. Further, the  
604 simulated population also has the same underlying biological population (the Shetland sheep population)  
605 that was used to develop the prior distributions, so it is possible that this overlap reflects that fact.

606 The parametric and compositional accuracy of both simulation tests are summarized in Table 8. The  
607 single assemblage model is well conditioned when examining parametric accuracy, though the multisite  
608 model overfits in this respect; this is driven by poor performance on size variability ( $\sigma$ ) parameters—the  
609 model estimates average body size ( $\mu$ ) parameters well. The multisite model also has a tendency to underfit  
610 when examining site-specific compositional accuracy. In both models, though, the compositional accuracy  
611 improves (in the sense of no longer underfitting) by using the modeled assemblages rather than the measured  
612 assemblages. This makes intuitive sense, as the measured assemblage is itself theoretically a sample from  
613 the modeled assemblage (based on the assumption that “measurability” is random). In these simulations, of  
614 course, this theory is held to be explicitly true, though the relationship between the measured and modeled  
615 assemblage is generally held to be true implicitly in zooarchaeology and can be explicitly tested (see Section  
616 2.4).

617 Figures 6 and 7 show the posterior distributions of the site-level parameters for the single-assemblage and  
618 multisite simulations. In each figure, Panel A shows the posterior distributions for the mixing proportions ( $\pi$   
619 parameters), Panel B shows the posterior distributions for the average body size ( $\mu$  parameters), and Panel  
620 C shows the posterior distributions for the size variability ( $\sigma$  parameters) for immature (black), adult female  
621 (blue), and adult male (red) specimens in the assemblage. Vertical bars summarize the 80% and 95% credible  
622 intervals for the parameter, while the solid horizontal lines denote the “true” parameter value for the Shetland  
623 sheep population from which the assemblage was sampled, including any deviations made for the multisite  
624 simulation (see Table 5). The model excels at estimating the average body sizes of the animal groups, even  
625 when those body sizes are manipulated (Figures 6B and 7B). The multisite model’s estimates of the overall  
626 proportions tend to be conservative (Figure 7A); that is, underestimating the proportions for animal groups



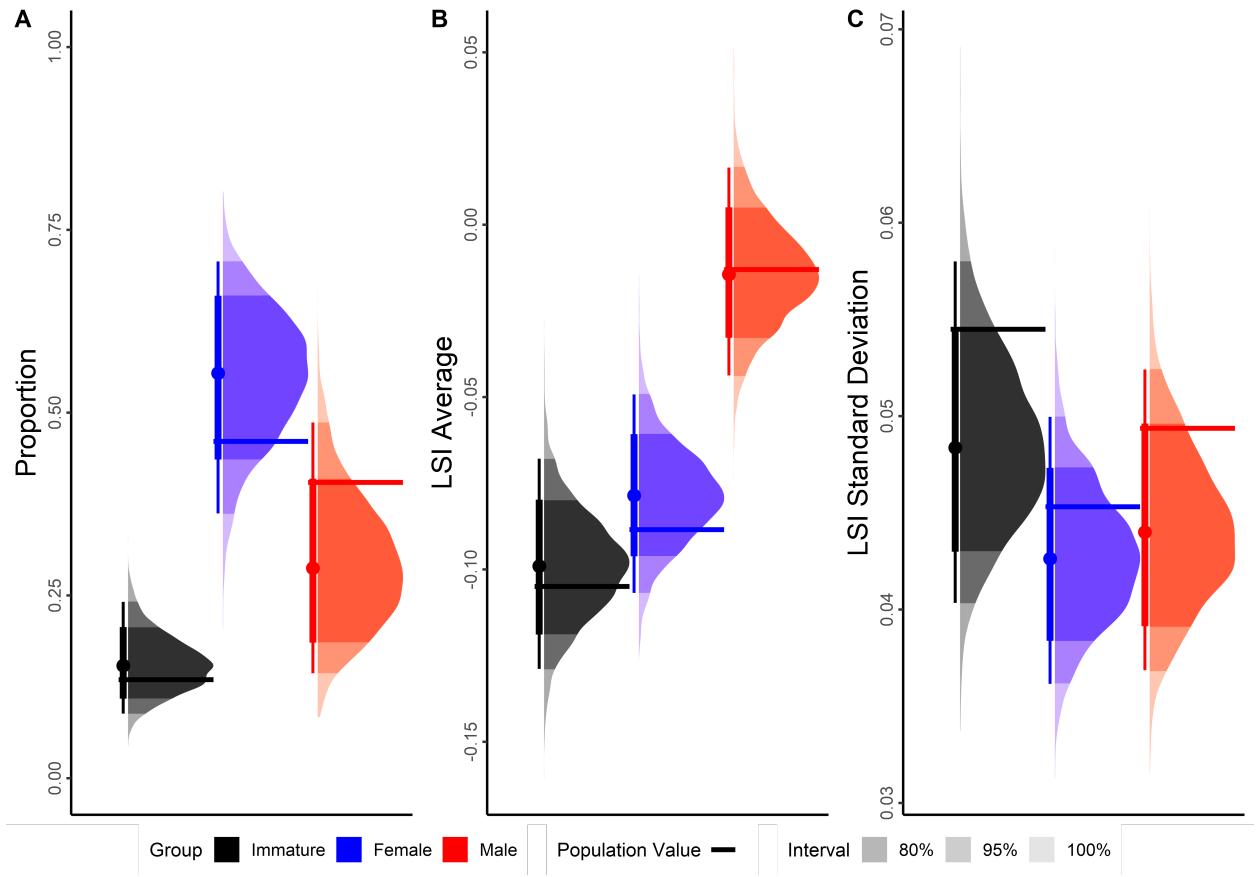
**Figure 5:** Comparison of prior and posterior distributions for mixture model hyper-parameters of the simulated single assemblage. The model hyper-parameters serve as assemblage-wide estimates accounting for size and composition variation across element portions.

**Table 8:** Parametric accuracy rates and expected accuracy ranges for the simulation models compared to the population values from which the samples were drawn. Expected accuracy ranges are based on the number of parameters being evaluated. If the modeled accuracy rate is above the range, then the model has underfit to the data. If the modeled accuracy rate is below the range, then the model has overfit to the data.

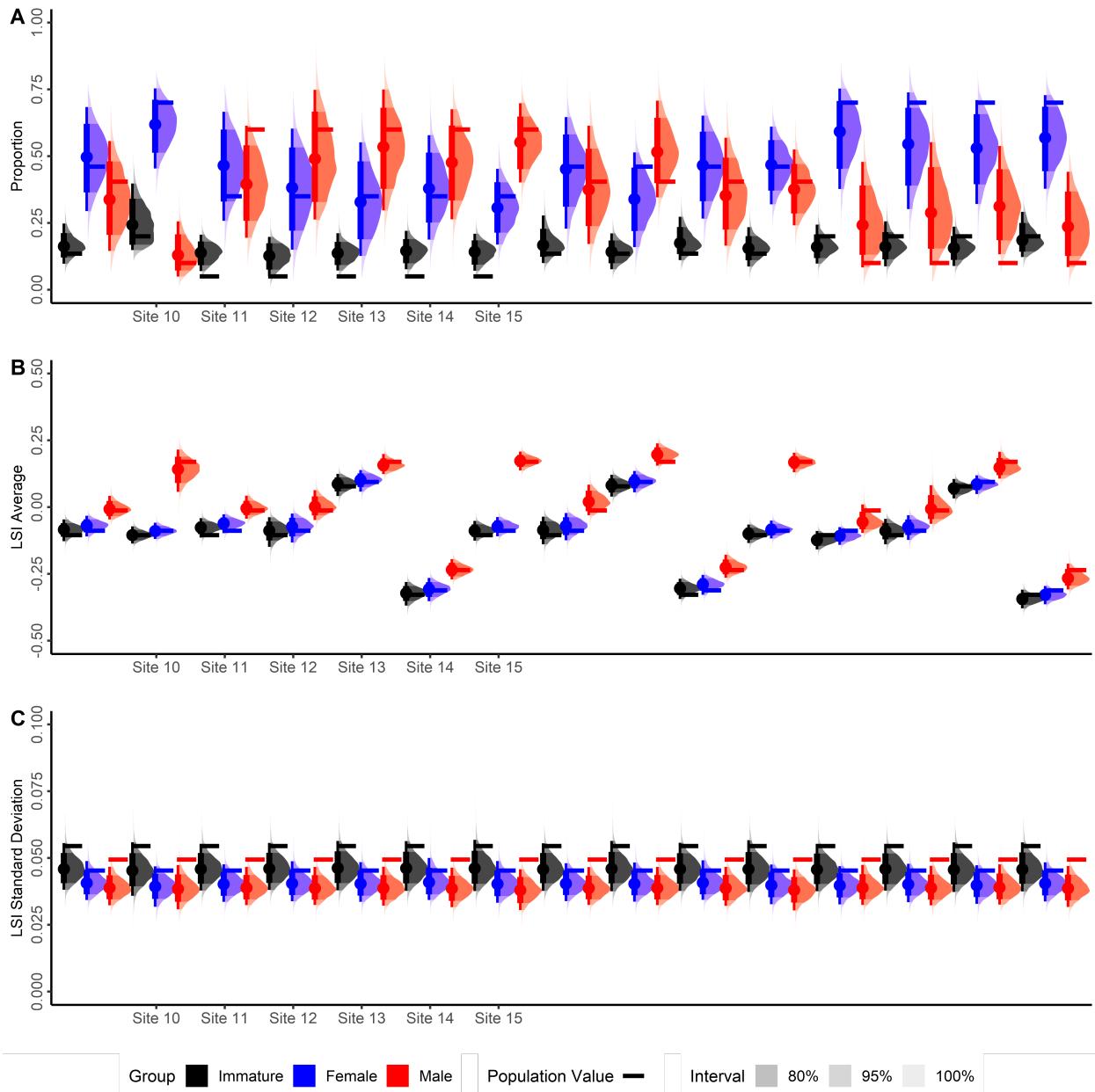
Quantity	Assemblage	Accuracy	Expected Accuracy	Interpretation
Parameters (80% CI)	Single Assemblage	9 / 9 (100%)	5-9	
Parameters (80% CI)	Multisite (Overall)	79 / 135 (59%)	99-117	Overfit
Parameters (80% CI)	Multisite (Average Size)	41 / 45 (91%)	31-41	
Parameters (95% CI)	Single Assemblage	9 / 9 (100%)	7-9	
Parameters (95% CI)	Multisite (Overall)	114 / 135 (84%)	123-133	Overfit
Parameters (95% CI)	Multisite (Average Size)	45 / 45 (100%)	40-45	
Measurement Composition (80% CI)	Single Assemblage	24 / 24 (100%)	15-23	Underfit
Measurement Composition (80% CI)	Multisite (Overall)	21 / 24 (88%)	15-23	
Measurement Composition (80% CI)	Multisite (By Site)	331 / 339 (98%)	256-285	Underfit
Measurement Composition (95% CI)	Single Assemblage	24 / 24 (100%)	20-24	
Measurement Composition (95% CI)	Multisite (Overall)	24 / 24 (100%)	20-24	
Measurement Composition (95% CI)	Multisite (By Site)	338 / 339 (100%)	314-329	Underfit
Modeled Composition (80% CI)	Single Assemblage	18 / 24 (75%)	15-23	
Modeled Composition (80% CI)	Multisite (Overall)	17 / 24 (71%)	15-23	
Modeled Composition (80% CI)	Multisite (By Site)	302 / 360 (84%)	273-303	
Modeled Composition (80% CI)	Multisite (New Elements)	18 / 21 (86%)	13-20	
Modeled Composition (95% CI)	Single Assemblage	21 / 24 (88%)	20-24	
Modeled Composition (95% CI)	Multisite (Overall)	22 / 24 (92%)	20-24	
Modeled Composition (95% CI)	Multisite (By Site)	349 / 360 (97%)	334-350	
Modeled Composition (95% CI)	Multisite (New Elements)	20 / 21 (95%)	18-21	

627 with extremely low “true” proportions (e.g., adult males in Sites 6-10, immature animals in Sites 11-15)  
 628 while overestimating the proportions for animal groups with extremely high “true” proportions (e.g., adult  
 629 females in Sites 6-10). This is likely due to the multilevel modeling structure “shrinking” estimates towards  
 630 a global mean, though may also reflect overfitting from small sample sizes. This is particularly the case when  
 631 examining variability parameters ( $\sigma$ : Figure 7C), which the multilevel model tends to underestimate.

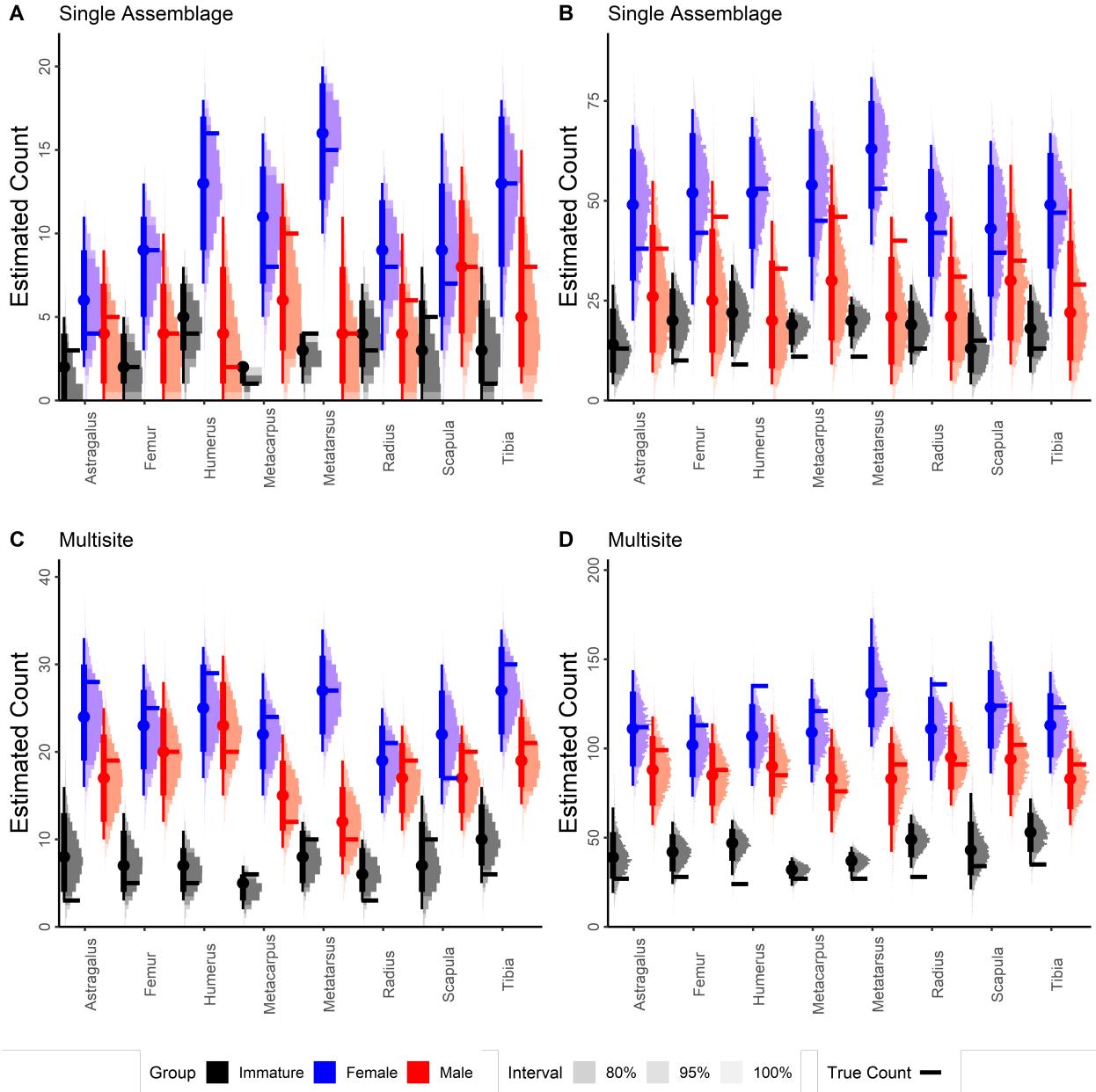
632 Figure 8 shows posterior distributions of simulated group-specific compositions for both the single-  
 633 assemblage and total composition of the multisite models (i.e., all sites combined) alongside true counts  
 634 for each group. The underfitting performance of the multisite simulation, particularly at the level of indi-  
 635 vidual sites, may be a consequence of low statistical power due to small sample sizes. The median number  
 636 of specimens per element portion is 4 for the measured assemblage and 17 for the modeled assemblage. By  
 637 contrast, the single-assemblage simulation has a median number of specimens per element portion of 19.5  
 638 for the measured assemblage and 92 for the modeled assemblage. More important than overall sample sizes,  
 639 though, is the group-specific element counts, where the contrast between the median number of specimens in  
 640 the measured and modeled assemblages from the multisite simulation (measurement assemblage: 1, modeled  
 641 assemblage: 4) and the measured and modeled assemblages from the single-assemblage simulation (measure-  
 642 ment assemblage: 5, modeled assemblage: 26) is starker. The larger sample sizes in the modeled assemblages  
 643 also partially explains the reduced underfitting relative to the measured assemblages.



**Figure 6:** Posterior distributions of the mixture model hyper-parameters (assemblage-level estimates) for the simulated single assemblage. Horizontal lines denote the relevant parameter values from the Shetland sheep population from which the assemblage was sampled.



**Figure 7:** Posterior distributions of assemblage-level mixture model parameters for the simulated multisite assemblages. Horizontal lines denote the relevant parameter values from the Shetland sheep population from which the assemblage was sampled, including modifications to the composition and body size for some assemblages (see Table 4).



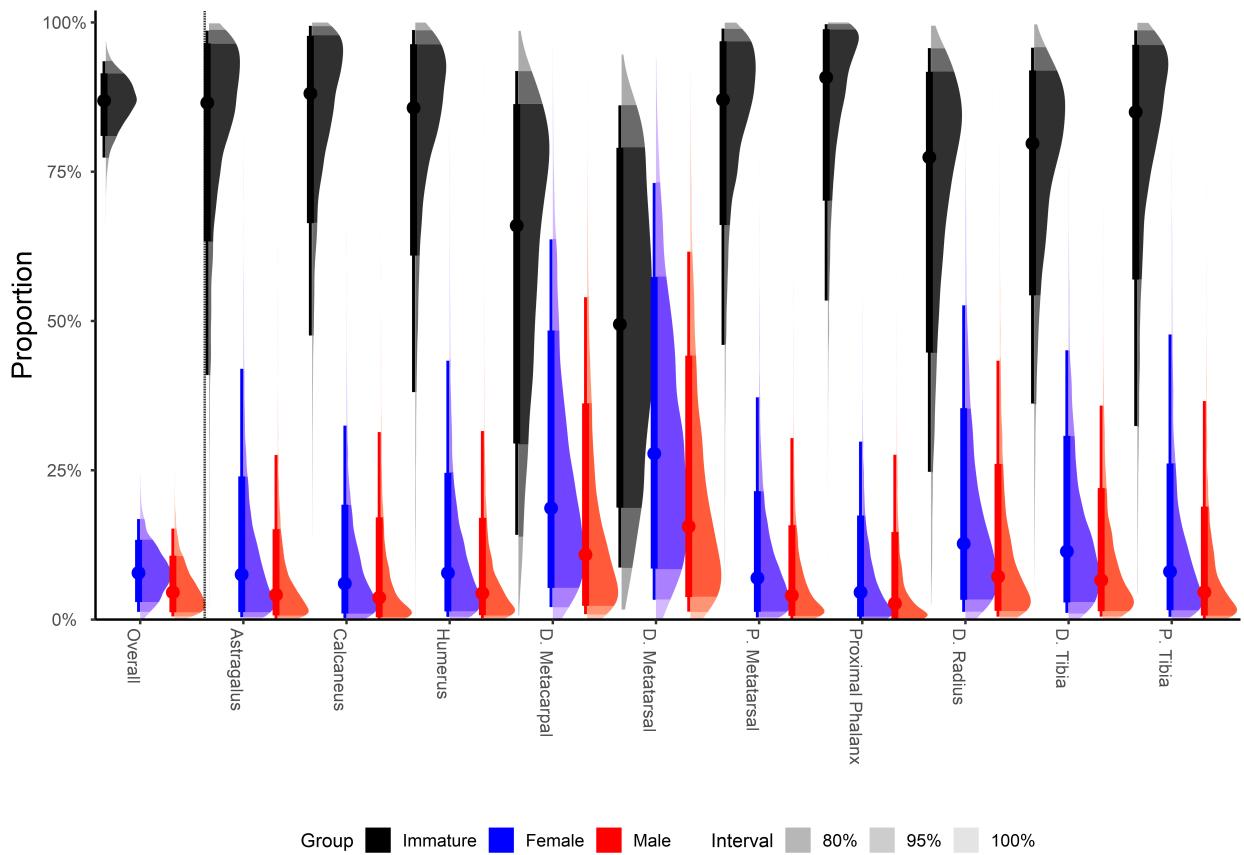
**Figure 8:** Posterior distributions of simulated group-specific composition for the simulated assemblages. Top row: single assemblage model (A) measured assemblage, (B) modeled assemblage. Bottom row: multisite assemblage model (combined counts) (C) measured assemblages, (D) modeled assemblages. Horizontal lines denote the true group-specific composition of the relevant measured or modeled assemblage.

644 It is also noteworthy that the additional sampling to create the multisite simulation's modeled assemblage  
645 resulted in seven newly observed element portions at certain sites. The model can estimate the composition of  
646 these element portions due to its multilevel structure, which estimates element-specific offsets and interaction  
647 terms ( $\nu_{\text{Element}}$  and  $\nu_{\text{Element}}$ ) for elements that are present in at least one site. The compositional accuracy of  
648 these newly observed element portions is well-calibrated (see Table 8), despite having no observed measured  
649 specimens from the element portion for those sites. That these element portions can be accurately modeled  
650 despite lacking observed measurements for the site lends support to the idea that researchers could extend the  
651 same multilevel model structure (element-specific offsets and interaction terms) to estimate the composition  
652 of unobserved (i.e., completely unmeasured) element portions in an assemblage.

653 **4.2 Pinarbaşı B Sheep: The impact of immature specimens**

654 The first three distributions in Figure 9 show the posterior estimates of the assemblage-level proportion  
655 of immature, female, and male sheep at Pinarbaşı B. In general, the Pinarbaşı B sheep assemblage is  
656 overwhelmingly composed of immature animals (posterior  $\mu_{\pi_1}$  median = 87%; 95% posterior credible interval  
657 for  $\mu_{\pi_1} = 77\text{-}94\%$ ), somewhat lower than the observed fusion rate of proximal and middle phalanges (59 / 62  
658 = 95%). Even though the overall proportions of female and male animals in the assemblage is low relative  
659 to immature specimens, the Bayesian multilevel mixture model can produce a posterior distribution of the  
660 overall adult sex ratio that suggests that adult females are more common than adult males (median  $\theta_{\text{Female}}$   
661 = 64%; 95% posterior credible interval for  $\theta_{\text{Female}} = 10\text{-}96\%$ ); the wide credible interval of this distribution  
662 reflects the small proportions of mature specimens in the assemblage overall. Despite this uncertainty, we  
663 can use this distribution to estimate the probability that adult females are more common than adult males  
664 (i.e., that  $\theta_{\text{Female}}$  is greater than 0.5); 68% of the posterior  $\theta_{\text{Female}}$  samples are above 0.5, providing some  
665 confidence in the interpretation that the adult sex ratio is skewed towards females. Note that any threshold  
666 value could be chosen to test a hypothesis about the adult sex ratio—for instance, only 45% of the posterior  
667  $\theta_{\text{Female}}$  samples are above 0.67 (i.e., a 2:1 female:male ratio), providing poor support that the mature portion  
668 of the assemblage is heavily skewed towards females.

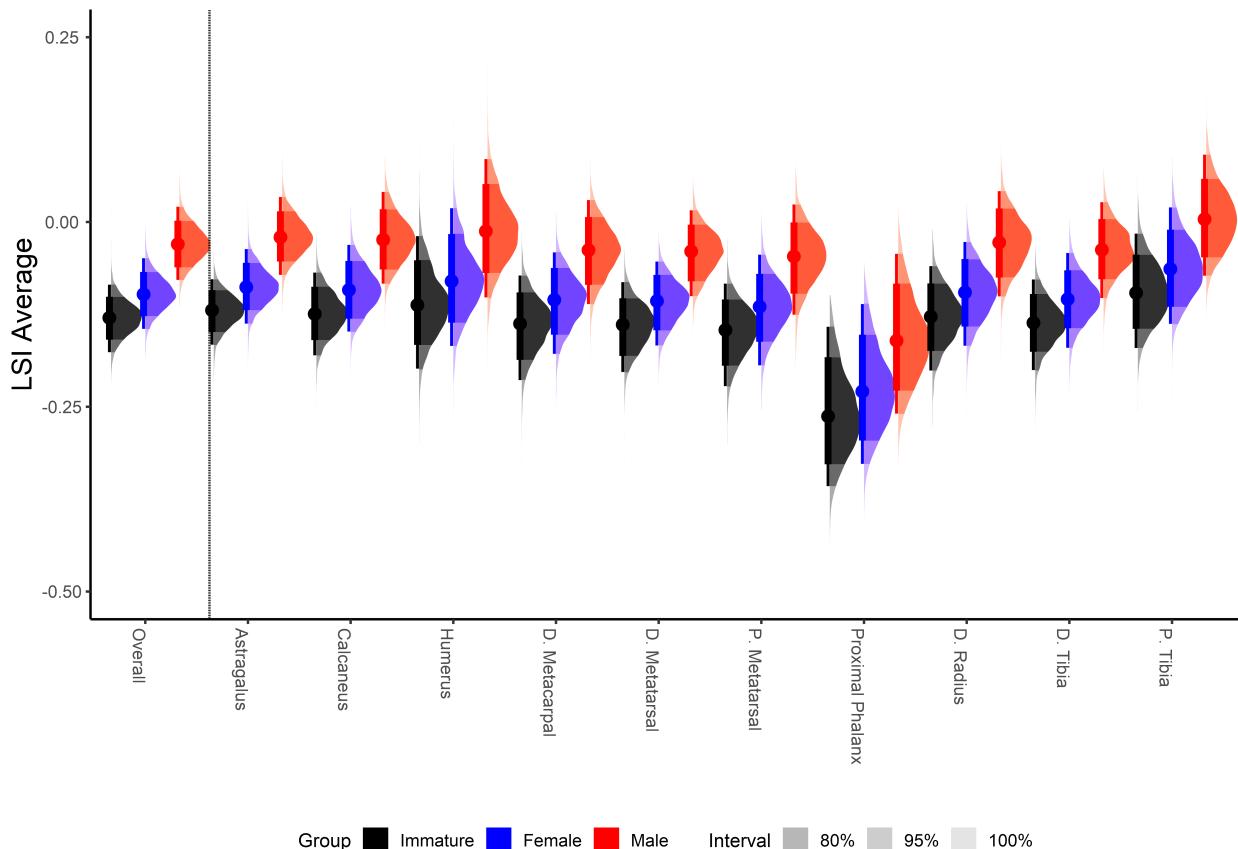
669 The other distributions show element-specific composition estimates for the Pinarbaşı B assemblage.  
670 Most of the element-specific proportions of immature animals ( $\pi_1$ ) broadly match the overall esitmates, with  
671 posterior medians over 75%. However, these element-specific distributions also have long tails extending into  
672 lower  $\pi_1$  values, conveying less certainty about element-specific  $\pi_1$  estimates relative to the assemblage-wide  
673 estimate. This likely owes to small element-specific sample sizes (the astragalus, calcaneus, and proximal  
674 phalanx have 9-10 specimens, all other element-specific samples sizes are 1-4, see Table 6) and to the presence  
675 of some element portions with lower modeled proportions of immature animals. Two element portions—the



**Figure 9:** Posterior distributions of composition estimates ( $\pi$ ) for the Pinarbaşı B sheep assemblage. The three distributions to the left of the vertical line are the assemblage-level estimates, while the distributions to the right of the line show element-specific composition estimates.

676 distal metacarpal and distal metatarsal—have posterior  $\pi_1$  median values below 75%, though again have long  
 677 tails that extend in both direction (Distal metacarpal:  $\pi_1$  posterior median = 66%, 95% posterior credible  
 678 interval: 46-92%; distal metatarsal:  $\pi_1$  posterior median = 49%, 95% posterior credible interval: 32-86%).  
 679 Notably, all measured specimens from these two element portionss have fused distal ephysises, meaning that  
 680 the model considers it impossible for the specimens to be immature.

681 Figure 10 shows the same comparison (overall and element-specific distributions) for the average size of  
 682 female animals ( $\mu_2$ ) in the Pinarbaşı B assemblage. Average body sizes vary across elements, highlighting  
 683 some allometric variation between Pinarbaşı B sheep and the standard sheep and the importance of using  
 684 a multilevel model to account for allometric variation. The multilevel structure of the model provides a  
 685 parameter ( $\sigma_{\text{Element}}[3]$ ) that directly estimates this variation: the posterior mean of this distribution is 0.05,  
 686 with an upper 95% quantile of 0.08 on the  $LSI_e$  scale. Combining all the measurements into a single LSI  
 687 analysis would confound this element-level variation with differences in composition, muddling the ability to  
 688 compare the biometry of Pinarbaşı B sheep with sheep from contemporary assemblages.



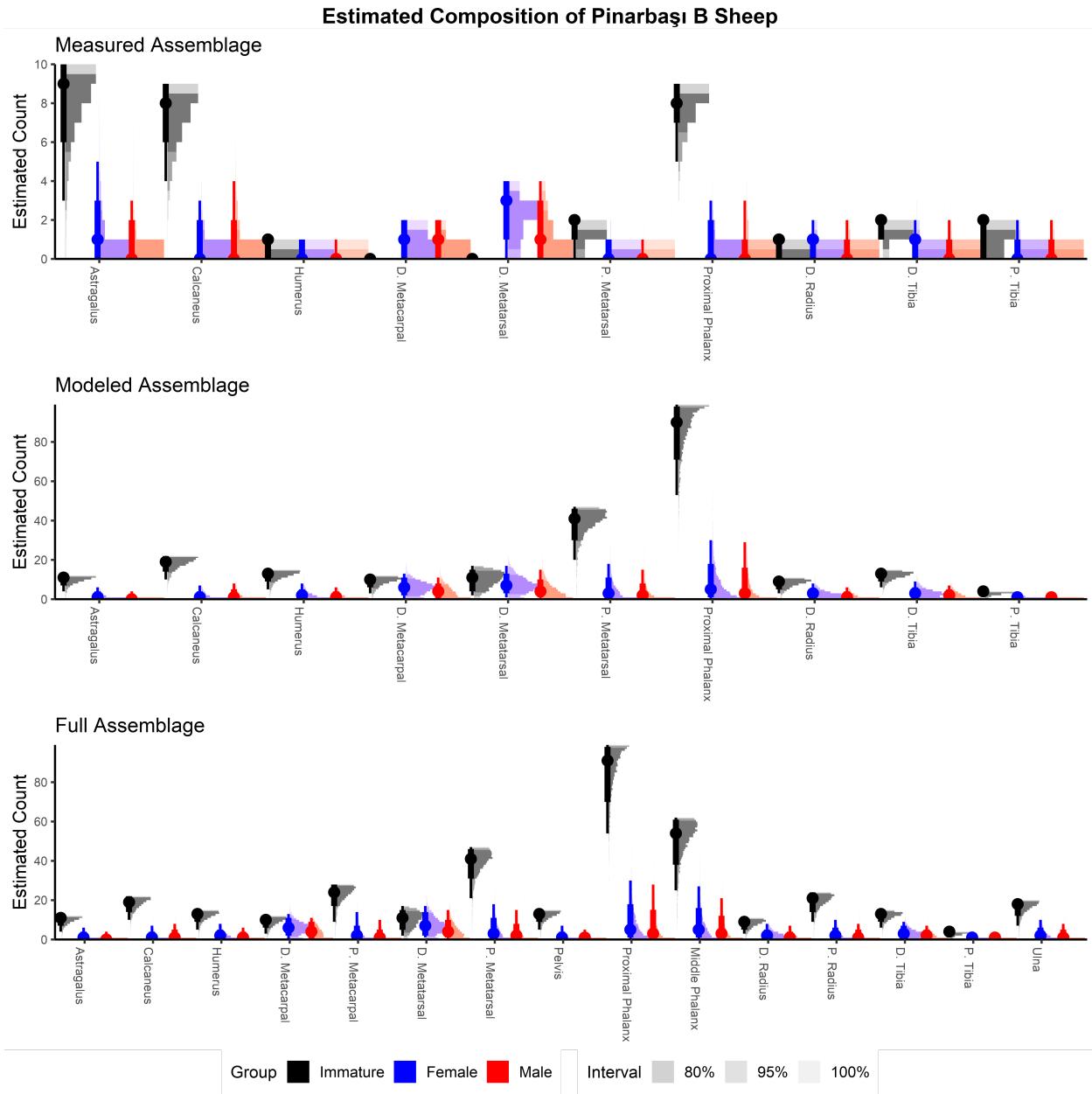
**Figure 10:** Posterior distributions of average LSI value for female animals ( $\mu_2$ ) for the Pinarbaşı B sheep assemblage. The distribution to the left of the vertical line is the assemblage-level estimate, while the distributions to the right of the line show element-specific size estimates.

689 Figure 11 shows the distribution of simulated compositions for immature, female, and male specimens  
690 in three Pinarbaşı B assemblages: the measured assemblage ( $N = 44$ ), the assemblage of modeled element  
691 portions ( $N = 277$ ), and the full sheep assemblage including five element portions that were not modeled  
692 due to lack of measurements (additional elements: proximal radius, ulna, proximal metacarpal, pelvis, and  
693 middle phalanx; total  $N = 428$ ). Beyond visualizing the group-specific composition of the assemblage, the  
694 simulated compositions provide more insight into the assemblage's formation. Five element portions in the  
695 full assemblage (astragalus, calcaneus, proximal radius, proximal fused metacarpal bones 3 and 4, pelvis)  
696 contain no mature specimens in 18% of the simulations; by contrast, no element portion has more than  
697 2% of the simulations without immature specimens (see Table 9). Additionally, element portions vary in  
698 the probability that they contain male specimens: distal metapodials (no males in 8% of simulations) and  
699 phalanges (no males in 16% of simulations) contain males much more frequently than other element portions  
700 (no males in 39% of simulations; Table 9). While this analysis focuses on appendicular elements, these results  
701 suggest that mature animals—particularly mature males—did not enter the assemblage as complete animals  
702 but potentially as raw materials for bone tools (e.g., distal metapodials, distal tibia: Russell and Griffitts  
703 2013) or on skins (e.g., proximal and middle phalanges). Immature animals, by contrast, appear to have  
704 entered the assemblage as complete animals: the element portions with the highest probability of missing  
705 immature specimens may be affected more by density-mediated attrition (e.g., proximal tibia: Symmons  
706 2005) and identifiability of neonatal specimens (e.g., distal metapodials; see Discussion).

707 These results strongly reinforce the interpretation that Pinarbaşı B was used by herders as a camp where  
708 sheep gave birth, with on-site culling largely reflecting either the first seasonal cull of animals before winter or  
709 animals that died naturally in their first year of life (Carruthers 2005; Martín et al. 2015). That this method  
710 reinforces previous analyses is not surprising; however, the mixture modeling results go beyond these earlier  
711 interpretations by creating group-specific biometric estimates and providing a probabilistic framework to  
712 estimate a specimen's membership into the three groups. The biometric estimates not only account for the  
713 presence of immature specimens in the assemblage but also for allometric variation across element portions.  
714 The probabilistic identifications allow analysts to simulate assemblage compositions, highlighting potential  
715 differences in the ways that bones from immature, adult female, and adult male entered the Pinarbaşı B  
716 assemblage.

717 **4.3 7th-6th Millennium BCE Northwest Anatolian Cattle: Examining differences between  
718 assemblages**

719 Figure 12 shows the posterior distributions of average body sizes for female cattle ( $\mu_2$ ) from the four analyzed  
720 assemblages. These distributions are produced from posterior samples; assemblage-specific estimates from a



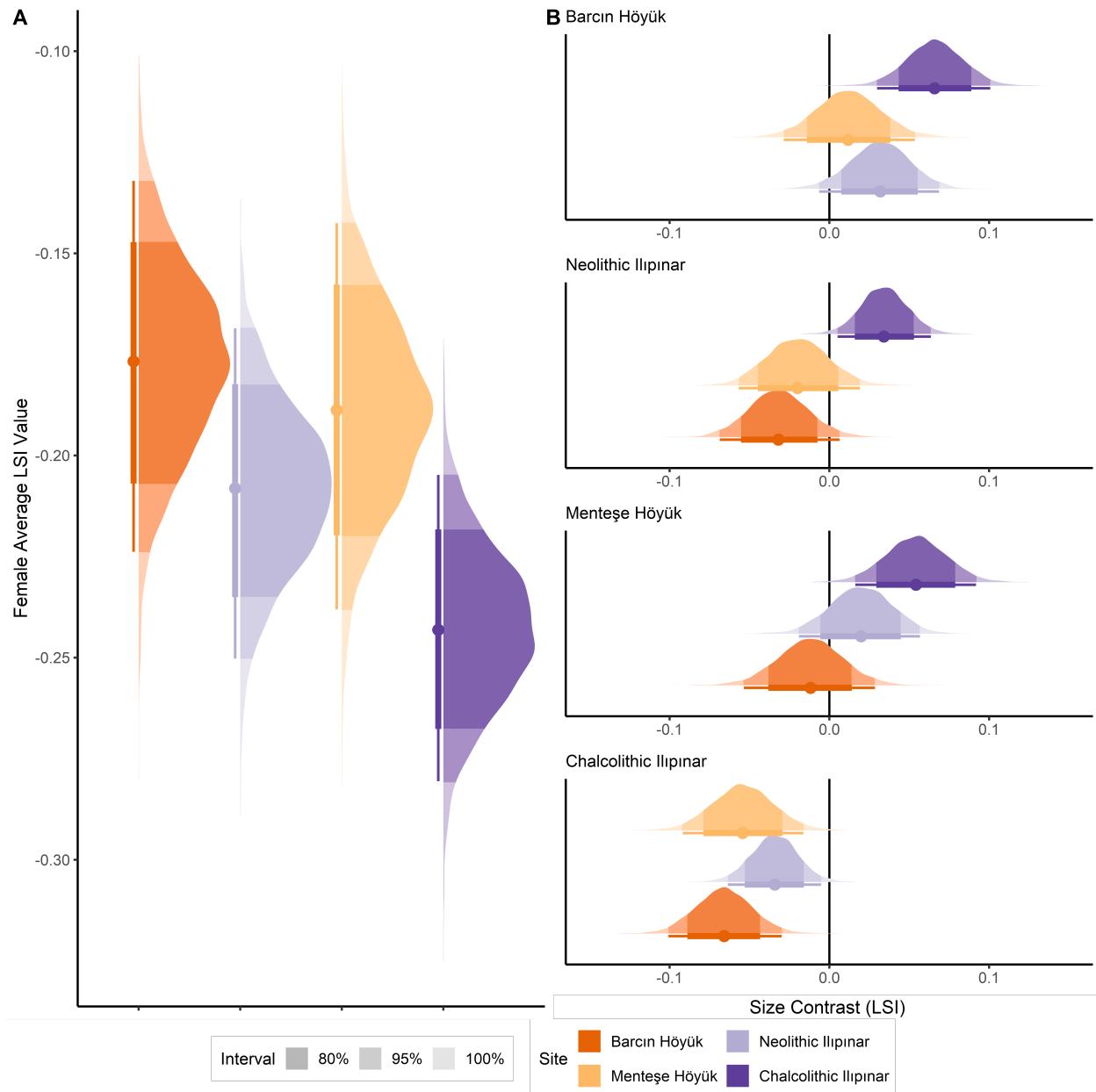
**Figure 11:** Posterior distributions of simulated group-specific composition for the Pınarbaşı B sheep (A) measured, (B) modeled, and (C) full assemblages. The full assemblage includes five additional element portions compared to the measured and modeled assemblages (proximal metacarpal, pelvis, middle phalanx, proximal radius, ulna).

**Table 9:** Percentage of simulated group-specific compositions of the full Pinarbaşı B sheep assemblage that exclude a different population category by element portion. Unmodeled element portions included in the composition analysis are proximal metacarpal, pelvis, middle phalanx, proximal radius, and ulna.

Element Portion	p(No Immature)	p(No Adults)	p(No Females)	p(No Males)
D. Metatarsal	1%	0%	2%	6%
D. Metacarpal	0%	0%	3%	10%
Proximal Phalanx	0%	0%	7%	16%
D. Radius	0%	0%	7%	33%
Middle Phalanx	0%	0%	8%	17%
D. Tibia	0%	0%	9%	22%
P. Metatarsal	0%	7%	16%	27%
Ulna	0%	0%	16%	32%
Humerus	0%	0%	18%	36%
P. Metacarpal	0%	12%	24%	36%
P. Radius	0%	12%	26%	39%
P. Tibia	2%	0%	30%	49%
Calcaneus	0%	18%	35%	41%
Pelvis	0%	22%	38%	50%
Astragalus	0%	28%	41%	64%

721 single posterior sample share the same relevant hyper-parameters ( $\mu_{\mu_2}$  and  $\sigma_{\text{Site}}[4]$ ), meaning that they covary  
 722 with one another to an extent. To compare these distributions, then, a contrast is necessary to account for this  
 723 potential covariation. This is done by simply evaluating the difference between two parameters (e.g., between  
 724 the average female LSI<sub>e</sub> value  $\mu_2$  for Barçın Höyük and  $\mu_2$  for Neolithic İlipinar) in each posterior sample,  
 725 shown in the right-hand panel of Figure 12. These contrasts show that the female cattle from Chalcolithic  
 726 İlipinar are likely smaller, on average, than female cattle from the other sites. These cattle measurements are  
 727 3-7% smaller, on average, than those from the other northwestern Anatolian sites relative to the standard  
 728 animal's measurements. Thus, the mixture modeling approach not only identifies a size difference that was  
 729 unrecognized when using standard LSI analysis (e.g., Buitenhuis 2008), it also reveals a diachronic trend  
 730 in cattle body size that can be studied to evaluate factors like climate or herding practices that affected  
 731 animal body size over time (e.g., Munro, Lebzon, and Sapir-Hen 2022; Wright and Viner-Daniels 2015).  
 732 Accounting for differences in the elemental and demographic composition of different assemblages provides  
 733 more accurate reconstructions of body size that allow us to critically examine patterns of animal size change  
 734 in the past (e.g., Arbuckle and Kassebaum 2021; Manning et al. 2015).

735 In addition to identifying size differences among the assemblages, the modeling results also reveals varia-  
 736 tion in the age and sex composition of the four assemblages. Figure 13 shows the distributions of assemblage-  
 737 level demographic variables—the proportion of immature animals and the adult sex ratio (the proportion of



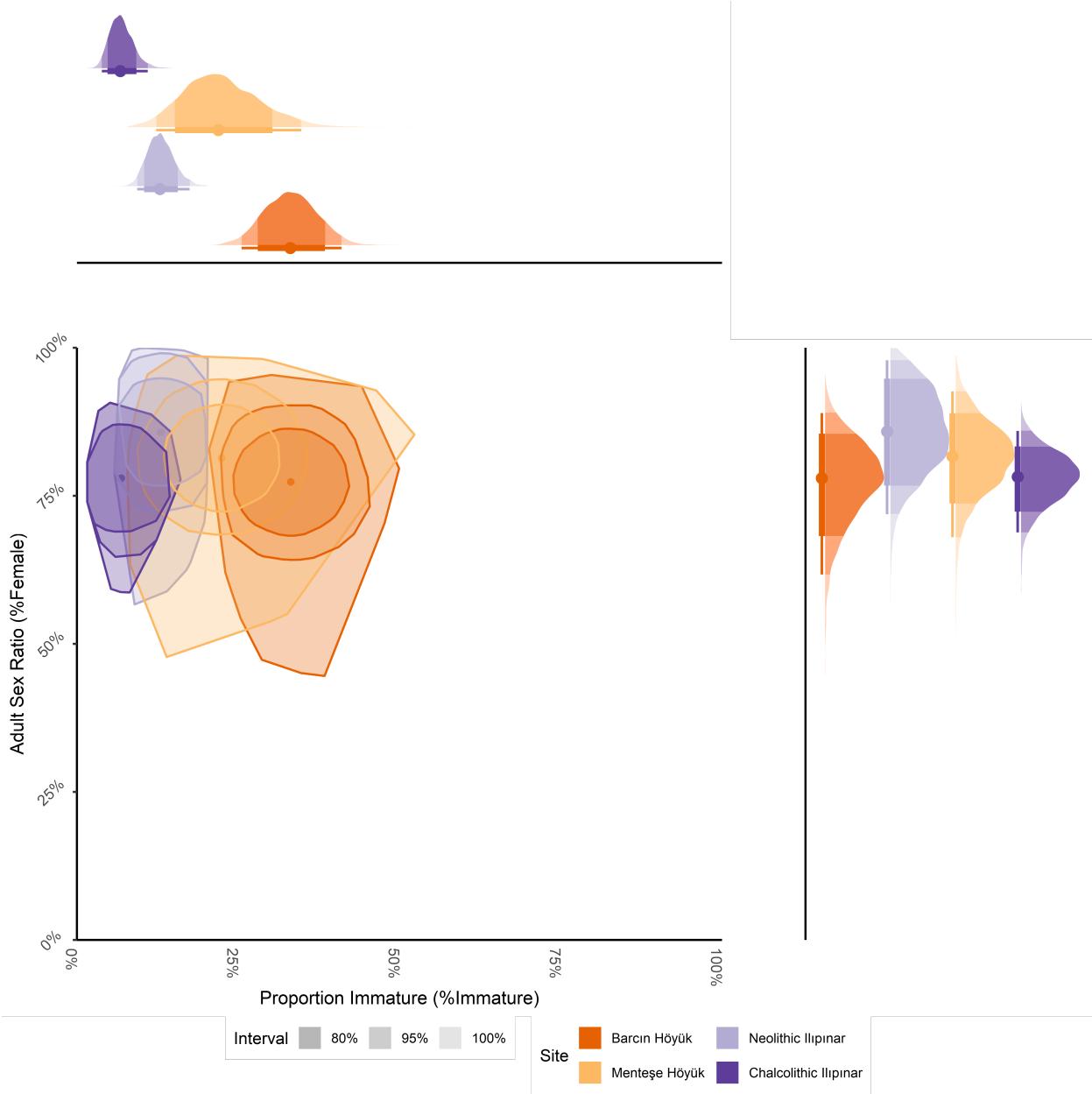
**Figure 12:** Posterior distributions of site-specific average LSI value for female animals ( $\mu_2$ ) for the Northwest Anatolian cattle assemblages. The left-side panel (A) shows the distributions in broadly temporal order from left to right, while the right-side panel (B) shows site-specific contrasts for average female body size, indicating specific size differences between pairs of sites. The title describes the focal assemblage, while the color scheme for the distributions is consistent across both panels. If the contrast distribution is greater than zero, then the focal assemblage is larger than the displayed assemblage (e.g., the top-most distribution in the top panel shows that Barçın Höyük is significantly larger than Chalcolithic İlipinar Höyük).

738 adults that are female)—for the four northwest Anatolian assemblages. The assemblages have broadly similar  
739 estimates for the adult sex ratio (right-hand panel of Figure 13), with strong evidence that they contain more  
740 females than males. However, modeling shows significant variation in the proportion of immature animals  
741 in the assemblages (top panel of Figure 13): the proportion of immature cattle in the Menteşe Höyük and  
742 Barçın Höyük assemblages is significantly higher than the proportions of immature cattle in either İlipinar  
743 Höyük assemblage. This difference between the sites could be investigated further to understand whether  
744 there are taphonomic factors that have selectively removed immature specimens from the İlipinar Höyük  
745 assemblage post-depositionally or whether herd management strategies differed among communities at these  
746 sites (e.g., Gillis et al. 2014, 2015).

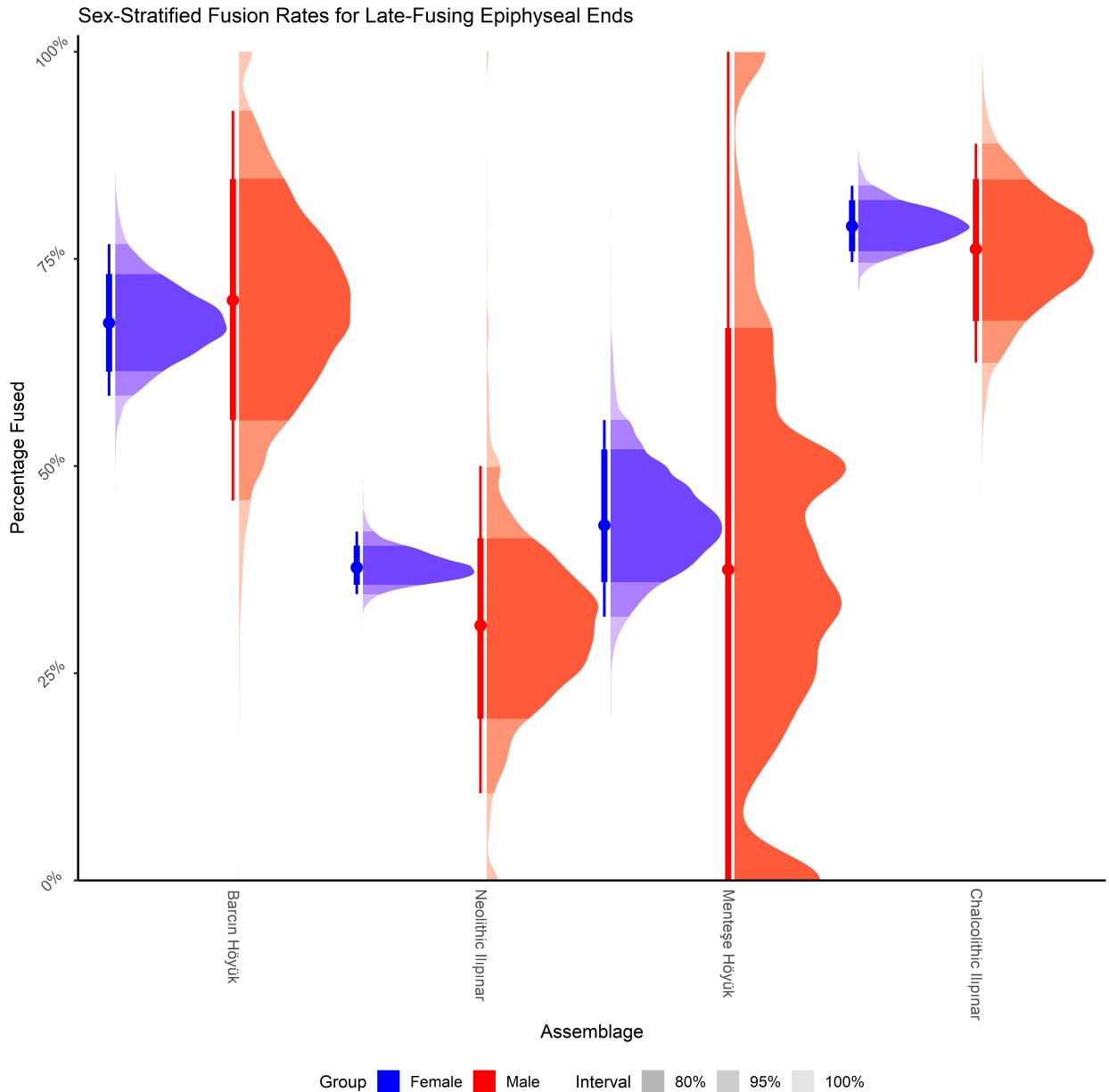
747 Simulating sex-specific fusion rates for late-fusing elements (proximal femur, distal femur, proximal  
748 humerus, distal radius, proximal tibia, proximal ulna: Grigson 1982) from the full northwest Anatolian  
749 assemblages highlights the complexities of examining sex-specific fusion rates in zooarchaeological assem-  
750 blages. In each assemblage, estimates of male fusion rates are extremely uncertain, owing to the small  
751 number of estimated males in each iteration and thus large potential shifts in the denominator for fusion  
752 rates (Figure 14). This uncertainty makes it difficult to clearly establish whether fusion rates differed be-  
753 tween males and females; regardless, in 68% of the posterior samples female fusion rates were higher than  
754 male fusion rates for Chalcolithic İlipinar. These results complicate regional syntheses that tie the presence  
755 of milk residues to milk-oriented cattle management (e.g., Evershed et al. 2008); the ability to directly  
756 estimate sex-specific fusion rates allows researchers to test the validity of these exploitation models for past  
757 assemblages (e.g., Arbuckle and Atici 2013).

## 758 5. DISCUSSION

759 The simulation analyses show that the Bayesian multilevel mixture model presented here can accurately  
760 reconstruct age- and sex-specific biometry of a faunal population represented in a measured assemblage, while  
761 also producing relatively accurate estimates of the “demographic” (ontogenetic age and sex) composition of  
762 the assemblage. The archaeological applications of the mixture model highlight how the model can highlight  
763 meaningful variation the composition and relative size of specimens across element portions. The results  
764 can point to potential differences in the how animals entered an assemblage, as between immature and  
765 mature sheep at Pinarbaşı B, or reveal variation in demographic proportions that could highlight taphonomic  
766 differences or variation in management strategies across the sites, as in the northwest Anatolian cattle. While  
767 broadly supporting the earlier analyses of these assemblages, applying mixture models to the measurement  
768 assemblages also opened new lines of inquiry based on exploring the drivers of variation in body size and



**Figure 13:** Comparison of posterior distributions of site-specific demographic parameters (proportion of immature  $\pi_{\text{Immature}}$  and adult sex ratio  $\theta_{\text{Female}}$ ) for the Northwest Anatolian cattle assemblages. Top panel shows the marginal plot for the proportion of immature specimens, while the right-hand panel shows the marginal plot for the adult sex ratio (in proportion of females among mature animals).



**Figure 14:** Posterior distributions of simulated sex-specific fusion rates for late-fusing elements among Northwest Anatolian cattle full assemblages. Included element portions are distal tibia, distal metapodials, calcaneus, femur, proximal ulna, distal radius, proximal tibia, and proximal humerus.

769 assemblage composition.

770 The performance of the Bayesian multilevel mixture models relies on the prior distributions, which  
771 provide constraints against overfitting and ensure that the model produces biologically reasonable parameter  
772 estimates. The prior distribution definitions in this paper were derived largely from the measurements of a  
773 herd of known-age, known-sex population of Shetland sheep (Popkin et al. 2012), though for the multisite  
774 cattle model some of the definitions were changed based on data on European aurochsen (Degerbøl 1970). It  
775 is important to note that prior distribution definitions can be derived from many different sources—including  
776 quantification based on one's judgment (e.g., Gelman et al. 2008; McCarthy and Masters 2005). This could  
777 allow researchers to use different kinds of demographic priors to adjust for their expectations about the ways  
778 that assemblages are created, for instance if one may expect catastrophic profiles (e.g., Stiner 1990). More  
779 important than the source of one's prior distribution definitions, is investigating the expectations of those  
780 prior distribution definitions by performing prior predictive checks as in Section 2.3 (Gabry et al. 2019;  
781 Gelman, Carlin, et al. 2020). Further, emphasis should be paid to increasing the diversity of known-age,  
782 known-sex animal populations with individual measurement data (e.g., Lebzon and Munro 2022; Zeder  
783 and Lemoine 2020), which could help develop prior distributions relevant to different taxa and to understand  
784 how variable different parameters, especially size variability ( $\sigma$ ) parameters, are across populations.

785 One of the central tenets of the mixture model's extension to modeled assemblages is the idea that  
786 "measurability" (adequate preservation to maintain a biometric dimension) is unrelated to a specimen's  
787 status as immature, female, or male. Variation in the mixture proportions among elements, especially  
788 the proportion of immature specimens  $\pi_1$  may highlight group-specific biases in the deposition of speci-  
789 mens but could also indicate issues with the assumption that "measurability" is random. The Pinarbaşı B  
790 sheep assemblage potentially demonstrates this issue, as the distal metapodials have much lower element-  
791 specific  $\pi_1$  estimates than other element portions. While metapodial bones from mature sheep could have  
792 been selectively over-represented in the assemblage, it is also likely that distal metapodials from immature  
793 animals—particularly very young animals—are less likely to be measurable compared to adult animals. Because  
794 the distal breadth measurement requires both distal condyles to be present, distal metapodial specimens from  
795 neonatal or extremely juvenile individuals may be missed while those from other element portions (e.g., prox-  
796 imal metapodial, distal humerus) would still be theoretically measurable (Martín and García-González 2015).  
797 The inclusion of condyle-specific measurements could address this issue, though would require identifying  
798 whether the isolated condyle is medial or lateral (e.g., width of condyle: Payne 1969).

799 The ability to create accurate simulated estimates of age and sex composition provides many opportunities  
800 for further analyses. For instance, comparison of the composition of animals in different depositional contexts  
801 could support contextual taphonomic analyses (e.g., Meier 2020). Access to certain kinds of animals could

802 highlight systems of provisioning or status-related restrictions (Arbuckle 2012; Twiss 2019: pp. 73-97).  
803 Differences in the ontogenetic age and sex composition of different body parts could also highlight ritual  
804 behaviors reflected in the use of certain contexts or sites (e.g., Madgwick and Mulville 2015). In a similar  
805 vein, tracking adult sex ratios could identify the use of castrates to take advantage of secondary products like  
806 wool or labor; the models used here would identify castrates as males, though alternative prior distributions  
807 and measured dimensions could distinguish these groups (e.g., Popkin et al. 2012: Figure 7). On a more  
808 practical level, providing specimen-specific probabilities of being immature, female, or male can provide a  
809 useful baseline for sampling strategies focused on ancient DNA or stable isotopes, allowing researchers to  
810 explore potential sex differences in diets (e.g., Post et al. 2001) or more easily identify male specimens to  
811 isolate Y-chromosomal DNA to explore sex-specific selection (e.g., Daly et al. 2021; McGrory et al. 2012).

812 The archaeological case studies highlight the importance of considering the presence of immature speci-  
813 mens and elemental variation in body size when summarizing the biometry and composition of an assemblage.  
814 Variation in the proportion of immature animals in the assemblage, as in the multisite case study for North-  
815 west Anatolia, may point to differences in culling strategies or even the seasonality of animal presence at  
816 the sites. Most of the Pinarbaşı B material derives from immature specimens, which could complicate inter-  
817 assemblage biometric analyses that do not use sex-specific size estimates (e.g., Arbuckle et al. 2014; Helmer  
818 et al. 2005). Restricting measurements only to fused specimens removes useful information, particularly  
819 when fusion rates may differ between male and female animals (Zeder and Hesse 2000); further, it does  
820 not resolve the problem of immature animals in the measurement assemblage if early-fusing elements like  
821 the distal humerus are still included because of post-fusion growth (Popkin et al. 2012). The ability to  
822 create sex-specific biometric estimates is important to document large-scale spatial and temporal dynamics  
823 in animal body size (e.g., Arbuckle et al. 2016; Arbuckle and Kassebaum 2021; Wolfhagen et al. 2021).

824 Examining  $LSI_e$  (or  $LSI_{10}$ ) values of Pinarbaşı B sheep without accounting for the impact of immature  
825 animals would mislead a researcher into believing that sheep were smaller, on average, than sheep from  
826 contemporaneous sites (e.g., late Çatalhöyük: Baird et al. 2011). While stable isotopic analyses have  
827 suggested that the Pinarbaşı B and Çatalhöyük sheep had similar diets (e.g., Baird et al. 2011, 2018;  
828 Middleton 2018), biometric analyses had not previously supported the idea that these assemblages derived  
829 from the same animal population (e.g., Arbuckle et al. 2014: Figure 4). Only eight measurements from the  
830 Pinarbaşı B sheep are excluded from being immature due to fusion status, four of which are distal metatarsals;  
831 the median  $LSI_e$  of these specimens is -0.12, compared to a median of -0.06 for the 362 Late Çatalhöyük  
832 sheep that must be mature based on fusion status (Wolfhagen et al. 2021). Even though the mixture  
833 model's estimate for the average  $LSI_e$  for female sheep ( $\mu_2$ ) at Pinarbaşı B is relatively uncertain (95%  
834 credible interval: -0.14 - -0.05  $LSI_e$ ), it still provides a useful reference point for comparison that supports

835 the idea that these assemblages have similar biometries (95% credible interval of  $\mu_2$  for late Çatalhöyük  
836 sheep: -0.11 - -0.06 LSI<sub>e</sub>: Wolfhagen et al. 2021).

837 The increased ability to specify the age and sex composition of faunal assemblages with Bayesian multi-  
838 level mixture models also highlights the limitations of our current language used to describe and interpret  
839 these compositions. Many discussions that examine changes in the composition of faunal assemblages to  
840 identify shifts in exploitation patterns use terms like “prime-dominated age structure” (e.g., Stiner 1990),  
841 “dominance of females,” (e.g., Peters, Driesch, and Helmer 2005) or “young male slaughter/kill-off” (e.g.,  
842 Zeder and Hesse 2000; Arbuckle and Atici 2013). These terms are deceptive in their utility—they describe  
843 some empirical pattern, but it is up to the individual researcher to define the cut-off between a “dominant”  
844 and “non-dominant” assemblage. In the case of the adult sex ratio for Barcın Höyük, 98% of the poste-  
845 rior samples are above 60% (1.5 females:1 male), but only 68% of the posterior samples are over 75% (3  
846 females:1 male). Meanwhile, 79% of the posterior samples for the adult sex ratio for Neolithic İlpınar are  
847 over 80% (4 females:1 male). Are both assemblages “dominated by females”? More formalized language  
848 in our hypotheses—or, rather, the adoption of statistical modeling frameworks (McElreath 2020: 4-17)—is  
849 necessary to clarify what changes in assemblage-level estimates of biometry and composition mean for past  
850 human-animal interactions.

## 851 6. CONCLUSIONS

852 This paper describes a new method for estimating the biometry and ontogenetic age and sex composition of  
853 faunal assemblages based on standard measurement data using Bayesian multilevel mixture modeling. The  
854 model produces accurate estimates of sex-specific biometry, which can provide a more useful framework for  
855 inter-assemblage analysis (e.g., Arbuckle et al. 2016; Helmer et al. 2005). Such a framework could better  
856 explore broad spatial and chronological patterns in animal biometry while accounting for differences in as-  
857 semblage composition across the assemblages, ensuring reliable comparisons of animal body size in relation  
858 to other variables. These analyses could investigate the processes behind size fluctuation in animals, partic-  
859 ularly in relation to changing human-animal interactions and adaptation to new lifeways and anthropogenic  
860 environments.

861 Furthermore, the estimates of the age and sex composition of the assemblage can be used to simulate  
862 assemblages of specimens with known group assignment (immature, female, and male). These simulations  
863 are the baseline for comparing differences in the composition of sub-assemblages. Using these simulations  
864 allows researchers to make testable statements about the composition of the assemblage and to directly  
865 test hypotheses about differences in the age and sex composition of animal bones from different parts of a

site, different fusion groups, or other categories. The Bayesian structure of the model allows researchers the flexibility to create bespoke hypotheses that can be tested directly, rather than relying on null hypothesis testing for inference (Otárola-Castillo and Torquato 2018; Otárola-Castillo et al. 2022). Thus, the mixture modeling framework described here provides a foundation for biometric and compositional analyses that operate at multiple scales and presents a new avenue for summarizing and comparing zooarchaeological assemblages.

## ##ACKNOWLEDGEMENTS

The initial modeling work was developed as part of my dissertation research, which was partially funded by a Wenner-Gren Foundation Dissertation Fieldwork Grant (Grant #9192). Thank you to R. Lee Lyman, Ryan Breslawski, and an anonymous reviewer for their dedicated and constructive feedback to improve this manuscript's accessibility and argumentation. Thank you as well to Natalie Munro and the University of Connecticut Zooarchaeology Lab (Roxanne Lebzon, Elic Weitzel, Nick Gonzalez, and Mario Mata-González—visiting from Universität Tübingen), whose members discussed an earlier version of this paper with me and provided insightful advice. I am grateful to the comments and insights on earlier versions of this manuscript from Chris W. Carleton, Zenobie Garrett, Ashley Lingle, Heather Lynch, Max D. Price, and Katheryn Twiss. Any errors are my own.

## References

- Arbuckle, Benjamin S. 2012. “Animals and Inequality in Chalcolithic Central Anatolia.” Journal Article. *Journal of Anthropological Archaeology* 31 (3): 302–13. <https://doi.org/10.1016/j.jaa.2012.01.008>.
- Arbuckle, Benjamin S., and Levent Atici. 2013. “Initial Diversity in Sheep and Goat Management in Neolithic South-Western Asia.” Journal Article. *Levant* 45 (2): 219–35.
- Arbuckle, Benjamin S., Sarah Whitcher Kansa, Eric C. Kansa, David C. Orton, Canan Cakirlar, Lionel Gourichon, A. Levent Atici, et al. 2014. “Data Sharing Reveals Complexity in the Westward Spread of Domestic Animals Across Neolithic Turkey.” Journal Article. *PLoS One* 9 (6): 1–11. <https://doi.org/10.1371/journal.pone.0099845>.
- Arbuckle, Benjamin S., and Theo M. Kassebaum. 2021. “Management and Domestication of Cattle (Bos Taurus) in Neolithic Southwest Asia.” Journal Article. *Animal Frontiers* 11 (3): 10–19. <https://doi.org/10.1093/af/vfab015>.
- Arbuckle, Benjamin S., and V. Özkaya. 2006. “Animal Exploitation at Körtik Tepe: An Early Aceramic Neolithic Site in Southeastern Turkey.” Journal Article. *Paléorient* 32 (2): 113–36.
- Arbuckle, Benjamin S., Max D. Price, Hitomi Hongo, and Banu Öksüz. 2016. “Documenting the Initial

- 897 Appearance of Domestic Cattle in the Eastern Fertile Crescent (Northern Iraq and Western Iran).”  
898 Journal Article. *Journal of Archaeological Science* 72: 1–9. <https://doi.org/10.1016/j.jas.2016.05.008>.
- 899 Baird, Douglas, Denise Carruthers, Andrew Fairbairn, and Jessica A. Pearson. 2011. “Ritual in the Land-  
900 scape: Evidence from Pınarbaşı in the Seventh-Millennium Cal BC Konya Plain.” Journal Article.  
901 *Antiquity* 85: 380–94.
- 902 Baird, Douglas, Andrew Fairbairn, Emma Jenkins, Louise Martin, Caroline Middleton, Jessica A. Pearson,  
903 Eleni Asouti, et al. 2018. “Agricultural Origins on the Anatolian Plateau.” Journal Article. *Proc Natl  
904 Acad Sci U S A*. <https://doi.org/10.1073/pnas.1800163115>.
- 905 Benaglia, Tatiana, Didier Chauveau, David R. Hunter, and Derek Young. 2009. “mixtools: An R Package  
906 for Analyzing Finite Mixture Models.” *Journal of Statistical Software* 32 (6): 1–29. <http://www.jstatsoft.org/v32/i06/>.
- 907 Betancourt, Michael. 2017. “A Conceptual Introduction to Hamiltonian Monte Carlo.” Journal Article.  
908 *arXiv* 1701.02434 [Preprint]: Available from <https://arxiv.org/abs/1701.02434v1>.
- 909 Breslawski, Ryan P. 2022. “Minimum Animal Units and the Standardized Count Problem.” Journal Article.  
910 *Journal of Archaeological Method and Theory*. <https://doi.org/10.1007/s10816-022-09563-9>.
- 911 Breslawski, Ryan P., and David A. Byers. 2015. “Assessing Measurement Error in Paleozoological Osteo-  
912 metrics with Bison Remains.” Journal Article. *Journal of Archaeological Science* 53: 235–42. <https://doi.org/10.1016/j.jas.2014.10.001>.
- 913 Buitenhuis, Hjalke. 2008. “İlipinar: The Faunal Remains from the Late Neolithic and Early Chalcolithic  
914 Levels.” Book Section. In *Archaeozoology of the Near East VIII*.
- 915 ———. 2013. “İlipinar Zooarchaeology.” Dataset. <http://opencontext.org/projects/D297CD29-50CA-4B2C-4A07-498ADF3AF487>. <https://doi.org/10.6078/M76H4FB>.
- 916 Çakırlar, Canan. 2013. “Rethinking Neolithic Subsistence at the Gateway to Europe With New Archaeo-  
917 zoological Evidence from Istanbul.” Book Section. In *Barely Surviving or More Than Enough?: The  
918 Environmental Archaeology of Subsistence, Specialisation and Surplus Food Production*, edited by Maaike  
919 Groot, Daphne Lentjes, and Jørn Zeiler, 59–80. Leiden: Sidestone Press.
- 920 Carty, Angelo, and B. D. Ripley. 2021. *Boot: Bootstrap r (s-Plus) Functions*.
- 921 Carruthers, Denise. 2005. “Hunting and Herding in Central Anatolian Prehistory: The Sites at Pinarbasi.”  
922 Book Section. In *Archaeozoology of the Near East VI*.
- 923 ———. 2006. “Pınarbaşı 1994: Animal Bones.” Dataset. <https://opencontext.org/projects/1677643a-15d0-1b2c-3a35-37f04c765387>.
- 924 Corporation, Microsoft, and Steve Weston. 2022. *doParallel: Foreach Parallel Adaptor for the 'Parallel'  
925 Package*. <https://CRAN.R-project.org/package=doParallel>.

- 930 Corti, Paulo, and David M. Shackleton. 2002. "Relationship Between Predation-Risk Factors and Sexual  
931 Segregation in Dall's Sheep (*Ovis Dalli Dalli*).” Journal Article. *Canadian Journal of Zoology* 80 (12):  
932 2108–17. <https://doi.org/10.1139/z02-207>.
- 933 Crabtree, Pam J. 1990. *West Stow: Early Anglo-Saxon Animal Husbandry*. Book. East Anglian Archaeology.  
934 Suffolk: Suffolk County Planning Dept.
- 935 Crema, Enrico R. 2011. "Modelling Temporal Uncertainty in Archaeological Analysis." Journal Article.  
936 *Journal of Archaeological Method and Theory* 19 (3): 440–61. <https://doi.org/10.1007/s10816-011-9122-3>.
- 937 Dahl, Gudrun, and Anders Hjort. 1976. *Having Herds: Pastoral Herd Growth and Household Economy*.  
938 Book. Stockholm Studies in Social Anthropology. Stockholm: University of Stokholm.
- 940 Daly, K. G., V. Mattiangeli, A. J. Hare, H. Davoudi, H. Fathi, S. B. Doost, S. Amiri, et al. 2021. "Hherded  
941 and Hunted Goat Genomes from the Dawn of Domestication in the Zagros Mountains." Journal Article.  
942 *Proc Natl Acad Sci U S A* 118 (25). <https://doi.org/10.1073/pnas.2100901118>.
- 943 Davis, Simon J. M. 1981. "The Effects of Temperature Change and Domestication on the Body Size of Late  
944 Pleistocene to Holocene Mammals of Israel." Journal Article. *Paleobiology* 7 (1): 101–14.  
945 ———. 1982. "Climatic Change and the Advent of Domestication: The Succession of Ruminant Artiodactyls  
946 in the Late Pleistocene-Holocene in the Israel Region." Journal Article. *Paléorient* 8 (2): 5–15.  
947 ———. 1996. "Measurements of a Group of Adult Female Shetland Sheep Skeletons from a Single Flock:  
948 A Baseline for Zooarchaeologists." Journal Article. *Journal of Archaeological Science* 23: 593–612.  
949 ———. 2000. "The Effect of Castration and Age on the Development of the Shetland Sheep Skeleton and  
950 a Metric Comparison Between Bones of Males, Females and Castrates." Journal Article. *Journal of  
951 Archaeological Science* 27 (5): 373–90. <https://doi.org/10.1006/jasc.1999.0452>.
- 952 Davison, A. C., and D. V. Hinkley. 1997. *Bootstrap Methods and Their Applications*. Cambridge: Cambridge  
953 University Press. <http://statwww.epfl.ch/davison/BMA/>.
- 954 De Cupere, Bea, Wim Van Neer, Hervé Monchot, Elina Rijmenants, Mircea Udrescu, and Marc Waelkens.  
955 2005. "Ancient Breeds of Domestic Fowl (*Gallus Gallus f. Domestica*) Distinguished on the Basis of  
956 Traditional Observations Combined with Mixture Analysis." Journal Article. *Journal of Archaeological  
957 Science* 32 (11): 1587–97. <https://doi.org/10.1016/j.jas.2005.04.015>.
- 958 Degerbøl, Magnus. 1970. "The Urus (*Bos Primigenius Bojanus*) and Neolithic Domesticated Cattle (*Bos  
959 Taurus Domesticus Linné*) in Denmark." Journal Article. *Det Kongelige Danske Videnskabernes Selskab  
960 Biologiske Skrifter* 17 (1): 5–177.
- 961 Dong, Zhuan. 1997. "Mixture Analysis and Its Preliminary Application in Archaeology." Journal Article.  
962 *Journal of Archaeological Science* 24: 141–61.

- 963 Dowle, Matt, and Arun Srinivasan. 2021. *Data.table: Extension of ‘Data.frame’*. <https://CRAN.R-project.org/package=data.table>.
- 964
- 965 Driesch, Angela von den. 1976. *A Guide to the Measurement of Animal Bones from Archaeological Sites*.  
966 Book. Peabody Museum Bulletins. Cambridge, MA: Harvard University.
- 967 Evershed, Richard P., Sebastian Payne, Andrew G. Sherratt, Mark S. Copley, Jennifer Coolidge, Duska  
968 Urem-Kotsu, Kostas Kotsakis, et al. 2008. “Earliest Date for Milk Use in the Near East and Southeastern  
969 Europe Linked to Cattle Herding.” Journal Article. *Nature* 455 (7212): 528–31. <https://doi.org/10.1038/nature07180>.
- 970
- 971 Fernández, Hélène, and Hervé Monchot. 2007. “Sexual Dimorphism in Limb Bones of Ibex (*Capra Ibex*  
972 l.): Mixture Analysis Applied to Modern and Fossil Data.” Journal Article. *International Journal of  
973 Osteoarchaeology* 17: 479–91. <https://doi.org/10.1002/oa.876>.
- 974 Fernée, Christianne L., and Konstantinos P. Trimmis. 2021. “Detecting Variability: A Study on the  
975 Application of Bayesian Multilevel Modelling to Archaeological Data. Evidence from the Neolithic Adri-  
976 atic and the Bronze Age Aegean.” Journal Article. *Journal of Archaeological Science* 128. <https://doi.org/10.1016/j.jas.2021.105346>.
- 977
- 978 Gabry, Jonah, and Rok Cešnovar. 2022. *Cmdstanr: R Interface to ‘CmdStan’*.
- 979 Gabry, Jonah, Daniel Simpson, Aki Vehtari, Michael Betancourt, and Andrew Gelman. 2019. “Visualization  
980 in Bayesian Workflow.” Journal Article. *Journal of the Royal Statistical Society, Series A* 182 (2): 389–  
981 402.
- 982 Galik, Alfred. 2013. “Barçın Höyük Zooarchaeology.” Dataset. <http://opencontext.org/projects/74749949-4FD4-4C3E-C830-5AA75703E08E>. <https://doi.org/10.6078/M78G8HM0>.
- 983
- 984 Gelman, Andrew. 2006a. “Multilevel (Hierarchical) Modeling: What It Can and Cannot Do.” Journal  
985 Article. *Technometrics* 48 (3): 432–35. <https://doi.org/10.1198/004017005000000661>.
- 986 ———. 2006b. “Prior Distributions for Variance Parameters in Hierarchical Models (Comment on Article  
987 by Browne and Draper).” Journal Article. *Bayesian Analysis* 3: 515–34.
- 988 Gelman, Andrew, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin. 2020.  
989 *Bayesian Data Analysis*. Book. Third Edition.
- 990 Gelman, Andrew, Aleks Jakulin, Maria Grazia Pittau, and Yu-Sung Su. 2008. “A Weakly Informative  
991 Default Prior Distribution for Logistic and Other Regression Models.” Journal Article. *The Annals of  
992 Applied Statistics* 2 (4): 1360–83. <https://doi.org/10.1214/08-aos191>.
- 993 Gelman, Andrew, Aki Vehtari, Daniel Simpson, Charles C. Margossian, Bob Carpenter, Yuling Yao, Lauren  
994 Kennedy, Jonah Gabry, Paul-Christian Bürkner, and Martin Modrák. 2020. “Bayesian Workflow.”  
995 Journal Article. *ArXiv*.

- 996 Gerritsen, Fokke, and Rana Özbal. 2019. “Barcin Höyük, a Seventh Millennium Settlement in the Eastern  
997 Marmara Region of Turkey.” Journal Article. *Documenta Praehistorica* 46: 58–67. <https://doi.org/10.4312/dp.46.4>.
- 999 Gillis, R., R. M. Arbogast, J. F. Piningre, K. Debue, and J. D. Vigne. 2015. “Prediction Models for Age-  
1000 at-Death Estimates for Calves, Using Unfused Epiphyses and Diaphyses.” Journal Article. *International  
1001 Journal of Osteoarchaeology* 25 (6): 912–22. <https://doi.org/10.1002/oa.2377>.
- 1002 Gillis, R., I. Carrère, M. Saña Seguí, G. Radi, and J. D. Vigne. 2014. “Neonatal Mortality, Young Calf  
1003 Slaughter and Milk Production During the Early Neolithic of North Western Mediterranean.” Journal  
1004 Article. *International Journal of Osteoarchaeology* 26 (2): 303–13. <https://doi.org/10.1002/oa.2422>.
- 1005 Gourichon, Lionel, and Daniel Helmer. 2008. “Etude de La Faune Neolithique de Mentesé.” Book Section. In  
1006 *Life and Death in a Prehistoric Settlement in Northwest Anatolia. The Ilipinar Excavations, Volume III.*,  
1007 edited by Jacob Roodenberg and Songül Alpaslan Roodenberg, 435–48. Leiden: Nederlands Instituut  
1008 voor he Nabije Oosten.
- 1009 ———. 2013. “Faunal Data from Neolithic Mentesé.” Dataset. <https://opencontext.org/projects/42ead4dbbaed-4a58-9b9b-7ec85266d2a9>. <https://doi.org/https://doi.org/10.6078/M7S46PVN>.
- 1011 Grau-Sologestoa, Idoia, and Umberto Albarella. 2019. “The ‘Long’ Sixteenth Century: A Key Period of  
1012 Animal Husbandry Change in England.” Journal Article. *Archaeological and Anthropological Sciences* 11  
1013 (6): 2781–2803. <https://doi.org/10.1007/s12520-018-0723-6>.
- 1014 Grigson, Caroline. 1982. “Sex and Age Determination of Some Bones and Teeth of Domestic Cattle: A  
1015 Review of the Literature.” Book Section. In *Ageing and Sexing Animal Bones from Archaeological Sites*,  
1016 edited by Bob Wilson, Caroline Grigson, and Sebastian Payne, 7–23. BAR British Series. Oxford: British  
1017 Archaeological Reports.
- 1018 ———. 1989. “Size and Sex: Evidence for the Domestication of Cattle in the Near East.” Book Section.  
1019 In *People and Culture in Change: Preceedings of the 2nd Symposium on Upper Paleolithic, Mesolithic,  
1020 and Mesolithic Populations of Europe and the Mediterranean Basin*, edited by Israel Hershkovitz, 77–109.  
1021 BAR International Series. Oxford: Archaeopress.
- 1022 Hammer, Oyvind. 2013. *PAST PAleontological STatistics Reference Manual*. Book.
- 1023 Helmer, Daniel, Lionel Gourichon, Hervé Monchot, Joris Peters, and Maria Saña Segui. 2005. “Identifying  
1024 Early Domestic Cattle from Pre-Pottery Neolithic Sites on the Middle Euphrates Using Sexual Dimor-  
1025 phism.” Book Section. In *First Steps of Animal Domestication: New Archaeozoological Approaches*,  
1026 edited by Jean-Denis Vigne, Joris Peters, and Daniel Helmer, 86–95. Proceedings of the 9th Conference  
1027 of the International Council of Archaeozoology, Durham, August 2002. Oxford: Oxbow Books.
- 1028 Hill, Matthew E., Matthew G. Hill, and Christopher C. Widga. 2008. “Late Quaternary Bison Diminution

- 1029       on the Great Plains of North America: Evaluating the Role of Human Hunting Versus Climate Change.”  
1030       Journal Article. *Quaternary Science Reviews* 27 (17-18): 1752–71. <https://doi.org/10.1016/j.quascirev>.  
1031       2008.07.002.
- 1032       Hongo, Hitomi, Jessica A. Pearson, Banu Öksüz, and Gülcin Ilgezdi. 2009. “The Process of Ungulate  
1033       Domestication at Çayönü, Southeastern Turkey: A Multidisciplinary Approach Focusing on Bos Sp.  
1034       And *Cervus Elaphus*.” Journal Article. *Anthropozoologica* 44 (1): 63–78.
- 1035       Jasra, A., C. C. Holmes, and D. A. Stephens. 2005. “Markov Chain Monte Carlo Methods and the Label  
1036       Switching Problem in Bayesian Mixture Modeling.” Journal Article. *Statistical Science* 20 (1): 50–67.  
1037       <https://doi.org/10.1214/088342305000000016>.
- 1038       Kabukcu, Ceren. 2017. “Woodland Vegetation History and Human Impacts in South-Central Anatolia  
1039       16,000–6500 Cal BP: Anthracological Results from Five Prehistoric Sites in the Konya Plain.” Journal  
1040       Article. *Quaternary Science Reviews* 176: 85–100. <https://doi.org/10.1016/j.quascirev.2017.10.001>.
- 1041       Karul, Necmi. 2019. “Early Farmers in Northwestern Anatolia in the Seventh Millennium.” Book Section.  
1042       In *Concluding the Neolithic: The Near East in the Second Half of the Seventh Millennium BC*, edited by  
1043       Arkadiusz Marciniak, 269–86. Atlanta, Georgia: Lockwood Press.
- 1044       Kassambara, Alboukadel. 2020. *Ggpubr: 'Ggplot2' Based Publication Ready Plots*. <https://CRAN.R-project.org/package=ggpubr>.
- 1045       Kay, Matthew. 2022. *ggdist: Visualizations of Distributions and Uncertainty*. <https://doi.org/10.5281/zenodo.3879620>.
- 1046       Koch, Paul L. 1986. “Clinal Geographic Variation in Mammals: Implications for the Study of Chronoclines.”  
1047       Journal Article. *Paleobiology* 12 (3): 269–81.
- 1048       Lebenzon, Roxanne, and Natalie D. Munro. 2022. “Body Size Variation in a Modern Collection of Mountain  
1049       Gazelle (*Gazella Gazella*) Skeletons.” Journal Article. *Journal of Archaeological Science: Reports* 41.  
1050       <https://doi.org/10.1016/j.jasrep.2021.103285>.
- 1051       Lyman, R. Lee. 2008. *Quantitative Paleozoology*. Book. Cambridge Manuals in Archaeology. Cambridge:  
1052       Cambridge University Press.
- 1053       Madgwick, Richard, and Jacqui Mulville. 2015. “Reconstructing Depositional Histories Through Bone  
1054       Taphonomy: Extending the Potential of Faunal Data.” Journal Article. *Journal of Archaeological Science*  
1055       53: 255–63. <https://doi.org/10.1016/j.jas.2014.10.015>.
- 1056       Manning, Katie, Adrian Timpson, Stephen Shennan, and Enrico Crema. 2015. “Size Reduction in Early  
1057       European Domestic Cattle Relates to Intensification of Neolithic Herding Strategies.” Journal Article.  
1058       *PLoS One* 10 (12): e0141873. <https://doi.org/10.1371/journal.pone.0141873>.
- 1059       Marin, Jean-Michel, Kerrie Mengerson, and Christian P. Robert. 2005. “Bayesian Modelling and Inference  
1060       53

- 1062       on Mixtures of Distributions.” Journal Article. *Handbook of Statistics* 25 (Bayesian Thinking, Modeling  
1063       and Computation): 459–509.
- 1064       Marom, Nimrod, and Guy Bar-Oz. 2013. “The Prey Pathway: A Regional History of Cattle (*Bos Taurus*)  
1065       and Pig (*Sus Scrofa*) Domestication in the Northern Jordan Valley, Israel.” Journal Article. *PLoS One*  
1066       8 (2): e55958. <https://doi.org/10.1371/journal.pone.0055958.g001>.
- 1067       Martín, Patricia, and Ricardo García-González. 2015. “Identifying Sheep (*Ovis Aries*) Fetal Remains in  
1068       Archaeological Contexts.” Journal Article. *Journal of Archaeological Science* 64: 77–87. <https://doi.org/10.1016/j.jas.2015.10.003>.
- 1069       Martín, Patricia, Ricardo García-González, Jordi Nadal, and Josep Maria Vergès. 2015. “Perinatal Ovi-  
1070       caprine Remains and Evidence of Shepherding Activities in Early Holocene Enclosure Caves: El Mirador  
1071       (Sierra de Atapuerca, Spain).” Journal Article. *Quaternary International*. <https://doi.org/10.1016/j.quaint.2015.08.024>.
- 1072       Marwick, Ben. 2017. “Computational Reproducibility in Archaeological Research: Basic Principles and a  
1073       Case Study of Their Implementation.” Journal Article. *Journal of Archaeological Method and Theory* 24  
1074       (2): 424–50. <https://doi.org/10.1007/s10816-015-9272-9>.
- 1075       Marwick, Ben, and Suzanne E. Pilaar Birch. 2018. “A Standard for the Scholarly Citation of Archaeological  
1076       Data as an Incentive to Data Sharing.” Journal Article. *Advances in Archaeological Practice* 6 (02):  
1077       125–43. <https://doi.org/10.1017/aap.2018.3>.
- 1078       McCarthy, Michael A., and P. I. P. Masters. 2005. “Profiting from Prior Information in Bayesian Analyses  
1079       of Ecological Data.” Journal Article. *Journal of Applied Ecology* 42 (6): 1012–19. <https://doi.org/10.1111/j.1365-2664.2005.01101.x>.
- 1080       McElreath, Richard. 2020. *Statistical Rethinking: A Bayesian Course with Examples in r and Stan*. Book.  
1081       Texts in Statistical Science. Boca Raton, FL: CRC Press.
- 1082       McGrory, S., E. M. Svensson, A. Götherström, J. Mulville, A. J. Powell, M. J. Collins, and T. P. O’Connor.  
1083       2012. “A Novel Method for Integrated Age and Sex Determination from Archaeological Cattle Mandibles.”  
1084       Journal Article. *Journal of Archaeological Science* 39 (10): 3324–30. <https://doi.org/10.1016/j.jas.2012.05.021>.
- 1085       Meadow, Richard. 1999. “The Use of Size Index Scaling Techniques for Research on Archaeozoological  
1086       Collections from the Middle East.” Book Section. In *Historia Animalium Ex Ossibus: Beiträge Zur  
1087       Paläoanatomie, Archäologie, Ägyptologie, Ethnologie Und Geschichte Der Tiermedizin*, edited by Cornelia  
1088       Becker, Henriette Manhart, Joris Peters, and Jörg Schibler, 285–300. Leidorf: Verlag Marie.
- 1089       Meier, Jacqueline. 2020. “The Contextual Taphonomy of Middens at Neolithic Kfar HaHoresh.” Journal  
1090       Article. *Journal of Archaeological Science: Reports* 33. <https://doi.org/10.1016/j.jasrep.2020.102531>.

- 1095 Middleton, Caroline. 2018. "The Beginning of Herding and Animal Management: The Early Development  
1096 of Caprine Herding on the Konya Plain, Central Anatolia." Journal Article. *Anatolian Studies* 68: 1–31.  
1097 <https://doi.org/10.1017/s0066154618000017>.
- 1098 Milner, J. M., E. B. Nilsen, and H. P. Andreassen. 2007. "Demographic Side Effects of Selective Hunting in  
1099 Ungulates and Carnivores." Journal Article. *Conserv Biol* 21 (1): 36–47. [https://doi.org/10.1111/j.1523-1739.2006.00591.x](https://doi.org/10.1111/j.1523-<br/>1100 1739.2006.00591.x).
- 1101 Monchot, Hervé, and Jacques Léchelle. 2002. "Statistical Nonparametric Methods for the Study of Fossil  
1102 Populations." Journal Article. *Paleobiology* 28 (1): 55–69. [https://doi.org/http://dx.doi.org/10.1666/0094-8373\(2002\)028%3C0055:SNMFTS%3E2.0.CO;2](https://doi.org/http://dx.doi.org/10.1666/0094-8373(2002)028%3C0055:SNMFTS%3E2.0.CO;2).
- 1103
- 1104 Monchot, Hervé, Marjan Mashkour, and Jean-Denis Vigne. 2005. "Kernel Smoothing and Mixture Analyses  
1105 for the Determination of the Sex Ratios at Death, at the Beginning of the Domestication of Ungulates." Book Section. In *First Steps of Animal Domestication: New Archaeozoological Approaches*, edited by  
1106 Jean-Denis Vigne, Joris Peters, and Daniel Helmer, 55–60. Oxford: Oxbow Books.
- 1107
- 1108 Munro, Natalie D., Guy Bar-Oz, and Austin C. Hill. 2011. "An Exploration of Character Traits and Linear  
1109 Measurements for Sexing Mountain Gazelle (*Gazella Gazella*) Skeletons." Journal Article. *Journal of  
1110 Archaeological Science* 38 (6): 1253–65. <https://doi.org/10.1016/j.jas.2011.01.001>.
- 1111 Munro, Natalie D., Roxanne Lebzon, and Lisa Sapir-Hen. 2022. "Revisiting Late Pleistocene-Early  
1112 Holocene Mountain Gazelle (*Gazella Gazella*) Body Size Change in the Southern Levant: A Case for  
1113 Anthropogenic Impact." Journal Article. *PLoS One* 17 (8): e0273024. <https://doi.org/10.1371/journal.pone.0273024>.
- 1114
- 1115 Otárola-Castillo, Erik, and Melissa G. Torquato. 2018. "Bayesian Statistics in Archaeology." Journal Article. *Annual Review of Anthropology* 47 (1). <https://doi.org/10.1146/annurev-anthro-102317-045834>.
- 1116
- 1117 Otárola-Castillo, Erik, Melissa G. Torquato, Jesse Wolfhagen, Matthew E. Hill, and Caitlin E. Buck. 2022.  
1118 "Beyond Chronology, Using Bayesian Inference to Evaluate Hypotheses in Archaeology." Journal Article. *Advances in Archaeological Practice*.
- 1119
- 1120 Özdogan, Mehmet. 2011. "Archaeological Evidence on the Westward Expansion of Farming Communities  
1121 from Eastern Anatolia to the Aegean and the Balkans." Journal Article. *Current Anthropology* 52 (S4):  
1122 S415–30. <https://doi.org/10.1086/658895>.
- 1123 ———. 2019. "Early Farmers in Northwestern Turkey: What Is New?" Book Section. In *Concluding  
1124 the Neolithic: The Near East in the Second Half of the Seventh Millennium BC*, edited by Arkadiusz  
1125 Marciniak, 307–28. Atlanta, Georgia: Lockwood Press.
- 1126 Payne, Sebastian. 1969. "A Metrical Distinction Between Sheep and Goat Metacarpals." Book Section. In  
1127 *The Domestication and Exploitation of Plants and Animals*, edited by Peter J. Ucko and Geoffrey W.

- 1128 Dimbleby, 295–305. London: Duckworth.
- 1129 ———. 1973. “Kill-Off Patterns in Sheep and Goats: The Mandibles from Aşvan Kale.” Journal Article.  
1130 *Anatolian Studies* 23: 281–303.
- 1131 Payne, Sebastian, and Gail Bull. 1988. “Components of Variation in Measurements of Pig Bones and Teeth,  
1132 and the Use of Measurements to Distinguish Wild from Domestic Pigs.” Journal Article. *Archaeozoologia*  
1133 II (1,2): 27–66.
- 1134 Pebesma, Edzer. 2018. “Simple Features for R: Standardized Support for Spatial Vector Data.” *The R  
1135 Journal* 10 (1): 439–46. <https://doi.org/10.32614/RJ-2018-009>.
- 1136 Pérez-Barbería, F. Javier, Iain J. Gordon, and M. Pagel. 2002. “The Origins of Sexual Dimorphism in Body  
1137 Size in Ungulates.” Journal Article. *Evolution* 56 (6): 1276–85.
- 1138 Peters, Joris, Angela von den Driesch, and Daniel Helmer. 2005. “The Upper Euphrates-Tigris Basin:  
1139 Cradle of Agro-Pastoralism?” Book Section. In *The First Steps of Animal Domestication*, edited by  
1140 Jean-Denis Vigne, Joris Peters, and Daniel Helmer, 96–124. Proceedings of the 9th Conference of the  
1141 International Council of Archaeozoology, Durham, August 2002. Oxford: Oxbow Books.
- 1142 Popkin, Peter R. W., Polydora Baker, Fay Worley, Sebastian Payne, and Andy Hammon. 2012. “The Sheep  
1143 Project (1): Determining Skeletal Growth, Timing of Epiphyseal Fusion and Morphometric Variation  
1144 in Unimproved Shetland Sheep of Known Age, Sex, Castration Status and Nutrition.” Journal Article.  
1145 *Journal of Archaeological Science* 39 (6): 1775–92. <https://doi.org/10.1016/j.jas.2012.01.018>.
- 1146 Post, Diane M., Trent S. Armbrust, Eva A. Horne, and Jacob R. Goheen. 2001. “Sexual Segregation  
1147 Results in Differences in Content and Quality of Bison (*Bos Bison*) Diets.” Journal Article. *Journal of  
1148 Mammalogy* 82 (2): 407–13.
- 1149 Pozo, Jose M, Valenzuela-Lamas, Silvia, Trentacoste, Angela, Nieto-Espinet, Ariadna, Guimarães Chiarelli,  
1150 and Silvia. 2021. “Zoolog: Zooarchaeological Analysis with Log-Ratios.” *Journal of Statistical Software*  
1151 In preparation. <https://josempozo.github.io/zoolog/>.
- 1152 Proaktor, G., T. Coulson, and E. J. Milner-Gulland. 2007. “Evolutionary Responses to Harvesting in  
1153 Ungulates.” Journal Article. *J Anim Ecol* 76 (4): 669–78. [https://doi.org/10.1111/j.1365-2656.2007.01244.x](https://doi.org/10.1111/j.1365-2656.2007.<br/>1154 01244.x).
- 1155 R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R  
1156 Foundation for Statistical Computing. <https://www.R-project.org/>.
- 1157 Ram, Karthik, and Ben Marwick. 2018. “Building Toward a Future Where Reproducible, Open Science Is  
1158 the Norm.” Book Section. In *The Practice of Reproducible Research: Case Studies and Lessons from the  
1159 Data-Intensive Sciences*, edited by Justin Kitzes, Daniel Turek, and Fatma Deniz, 69–78. Oakland, CA:  
1160 University of California Press.

- 1161 Redding, Richard W. 1984. "Theoretical Determinants of a Herder's Decisions: Modeling Variation in the  
1162 Sheep/Goat Ratio." Book Section. In *Animals and Archaeology*, edited by Juliet Clutton-Brock and  
1163 Caroline Grigson, 3. Early Herders and their Flocks:223–41. BAR International Series. Oxford: British  
1164 Archaeological Reports.
- 1165 Roodenberg, Jacob. 2012a. "Change in Food Production and Its Impact on an Early 6th Millennium  
1166 Community in Northwest Anatolia. The Example of Ilipinar." Journal Article. *Praehistorische Zeitschrift*  
1167 87 (2). <https://doi.org/10.1515/pz-2012-0015>.
- 1168 ———. 2012b. "Ilipinar: A Neolithic Settlement in the Eastern Marmara Region." Book Section. In  
1169 *The Oxford Handbook of Ancient Anatolia (10,000-323 BCE)*, edited by Gregory McMahon and Sharon  
1170 Steadman. <https://doi.org/10.1093/oxfordhb/9780195376142.013.0044>.
- 1171 Roodenberg, Jacob, A. van As, L. Jacobs, and M. H. Wijnen. 2003. "Early Settlement in the Plain of  
1172 Yenisehir (NW Anatolia): The Basal Occupation Layers at Mentese." Journal Article. *Anatolica* XXIX  
1173 (1): 18–58.
- 1174 Rosenthal, Robert. 1979. "The "File Drawer Problem" and Tolerance for Null Results." Journal Article.  
1175 *Psychological Bulletin* 86 (3): 638–41.
- 1176 RStudio Team. 2022. *RStudio: Integrated Development Environment for r*. Boston, MA: RStudio, PBC.  
1177 <http://www.rstudio.com/>.
- 1178 Ruckstuhl, K. E. 2007. "Sexual Segregation in Vertebrates: Proximate and Ultimate Causes." Journal  
1179 Article. *Integr Comp Biol* 47 (2): 245–57. <https://doi.org/10.1093/icb/icm030>.
- 1180 Ruckstuhl, K. E., and P. Neuhaus. 2002. "Sexual Segregation in Ungulates: A Comparative Test of  
1181 Three Hypotheses." Journal Article. *Biological Reviews* 77 (1): 77–96. <https://doi.org/10.1017/s1464793101005814>.
- 1183 Russell, Nerissa, and Janet L. Griffitts. 2013. "Catalhoyuk Worked Bones: South and 4040 Areas." Book  
1184 Section. In *Substantive Technologies at Catalhoyuk: Reports from the 2000-2008 Seasons*, edited by Ian  
1185 Hodder, 277–306. Los Angeles.
- 1186 Said, Sonia, Vincent Tolon, Serge Brandt, and Eric Baubet. 2011. "Sex Effect on Habitat Selection in  
1187 Response to Hunting Disturbance: The Study of Wild Boar." Journal Article. *European Journal of  
1188 Wildlife Research* 58 (1): 107–15. <https://doi.org/10.1007/s10344-011-0548-4>.
- 1189 Sasson, Aharon, and Susan Arter. 2020. "Earliest Utilization of Chicken in Upper California: The Zooar-  
1190 chaeology of Avian Remains from the San Diego Royal Presidio." Journal Article. *American Antiquity*  
1191 85 (3): 516–34. <https://doi.org/10.1017/aaq.2020.27>.
- 1192 Scrucca, Luca, Michael Fop, T. Brendan Murphy, and Adrian E. Raftery. 2016. "mclust 5: Clustering,  
1193 Classification and Density Estimation Using Gaussian Finite Mixture Models." *The R Journal* 8 (1):

- 1194 289–317. <https://doi.org/10.32614/RJ-2016-021>.
- 1195 Smith, A. F. M., and Alan E. Gelfand. 1992. “Bayesian Statistics Without Tears: A Sampling-Resampling  
1196 Perspective.” Journal Article. *The American Statistician* 46 (2): 84–88.
- 1197 South, Andy. 2017a. *Rnaturalearth: World Map Data from Natural Earth*. <https://CRAN.R-project.org/>  
1198 package=rnaturalearth.
- 1199 ———. 2017b. *Rnaturalearthdata: World Vector Map Data from Natural Earth Used in 'Rnaturalearth'*.  
1200 <https://CRAN.R-project.org/package=rnaturalearthdata>.
- 1201 Speth, John D. 1983. *Bison Kills and Bone Counts: Decision Making by Ancient Hunters*. Book. Chicago:  
1202 University of Chicago Press.
- 1203 ———. 2013. “Thoughts about Hunting: Some Things We Know and Some Things We Don’t Know.”  
1204 Journal Article. *Quaternary International* 297: 176–85. <https://doi.org/10.1016/j.quaint.2012.12.005>.
- 1205 Stan Development Team. 2021. “RStan: The R Interface to Stan.” <https://mc-stan.org/>.
- 1206 Stiner, Mary C. 1990. “The Use of Mortality Patterns in Archaeological Studies of Hominid Predatory  
1207 Adaptations.” Journal Article. *Journal of Anthropological Archaeology* 9: 305–51.
- 1208 Stiner, Mary C., Natalie D. Munro, Hiljke Buitenhuis, G. Duru, and M. Ozbasaran. 2022. “An Endemic  
1209 Pathway to Sheep and Goat Domestication at Asikli Hoyuk (Central Anatolia, Turkey).” Journal Article.  
1210 *Proc Natl Acad Sci U S A* 119 (4). <https://doi.org/10.1073/pnas.2110930119>.
- 1211 Symmons, Robert. 2005. “New Density Data for Unfused and Fused Sheep Bones, and a Preliminary  
1212 Discussion on the Modelling of Taphonomic Bias in Archaeofaunal Age Profiles.” Journal Article. *Journal  
1213 of Archaeological Science* 32 (11): 1691–98. <https://doi.org/10.1016/j.jas.2005.05.011>.
- 1214 Tchernov, Eitan, and Liora Kolska Horwitz. 1991. “Body Size Diminution Under Domestication: Uncon-  
1215 scious Selection in Primeval Domesticates.” Journal Article. *Journal of Anthropological Archaeology* 10:  
1216 54–75.
- 1217 Team, Stan Development. 2022. “Stan Modeling Language Users Guide and Reference Manual, Version  
1218 2.29.” Generic. <https://mc-stan.org>.
- 1219 Thissen, Laurens, Hadi Özbal, Ayla Türkekul Biyik, Fokke Gerritsen, and Rana Özbal. 2010. “The Land  
1220 of Milk? Approaching Dietary Preferences of Late Neolithic Communities in NW Anatolia.” Journal  
1221 Article. *Leiden Journal of Pottery Studies* 26: 157–72.
- 1222 Todd, Lawrence C. 1983. “The Horner Site: Taphonomy of an Early Holocene Bison Bonebed.” Thesis.
- 1223 Trentacoste, A., A. Nieto-Espinet, S. Guimaraes, B. Wilkens, G. Petrucci, and S. Valenzuela-Lamas. 2021.  
1224 “New Trajectories or Accelerating Change? Zooarchaeological Evidence for Roman Transformation of  
1225 Animal Husbandry in Northern Italy.” Journal Article. *Archaeol Anthropol Sci* 13 (1): 25. <https://doi.org/10.1007/s12520-020-01251-7>.

- 1227 Twiss, Katheryn C. 2019. *The Archaeology of Food*. Book. <https://doi.org/10.1017/9781108670159>.
- 1228 Twiss, Katheryn C., and Nerissa Russell. 2009. "Taking the Bull by the Horns: Ideology, Masculinity, and  
1229 Cattle Horns at atalhöyük (Turkey)." Journal Article. *Paléorient* 35 (2): 19–32.
- 1230 Uerpmann, Margarethe, and Hans-Peter Uerpmann. 1994. "Animal Bones." Book Section. In *Qala'at Al-*  
1231 *Bahrain*, edited by Flemming Hojlund and H. Hellmuth Andersen, 1: The Northern City Wall and the  
1232 Islamic Fortress:417–54. Jutland Archaeological Society Publications. Aarhus: Aarhus University Press.
- 1233 Vigne, Jean-Denis. 2011. "Le Mouton (Ovis Aries)." Book Section. In *Shillourokambos: Un Établissement*  
1234 *Néolithique Pré-Céramique à Chypre: Les Fouilles Du Secteur i*, edited by Jean Guilaine, François Briois,  
1235 and Jean-Denis Vigne, 1021–38. Paris: Editions Errance.
- 1236 Watkins, Trevor. 1996. "Excavations at Pinarbaşı: The Early Stages." Book Section. In *On the Surface:*  
1237 *Çatalhöyük 1993-1995*, edited by Ian Hodder, 47–57. Cambridge: McDonald Institute for Archaeological  
1238 Research.
- 1239 Weinstock, Jacobo. 2006. "Environment, Body Size and Sexual Dimorphism in Late Glacial Reindeer."  
1240 Book Section. In *Recent Advances in Ageing and Sexing Animal Bones*, edited by Deborah Ruscillo,  
1241 247–53. Oxford: Oxbow Books.
- 1242 Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>.
- 1243 Widga, Chris. 2006. "Niche Variability in Late Holocene Bison: A Perspective from Big Bone Lick, KY."  
1244 Journal Article. *Journal of Archaeological Science* 33 (9): 1237–55. <https://doi.org/10.1016/j.jas.2005.12.011>.
- 1245 Wolfhagen, Jesse. 2020. "Re-Examining the Use of the LSI Technique in Zooarchaeology." Journal Article.  
1246 *Journal of Archaeological Science* 123: 105254. <https://doi.org/10.1016/j.jas.2020.105254>.
- 1247 Wolfhagen, Jesse, and Max D. Price. 2017. "A Probabilistic Model for Distinguishing Between Sheep and  
1248 Goat Postcranial Remains." Journal Article. *Journal of Archaeological Science: Reports* 12: 625–31.  
1249 <https://doi.org/10.1016/j.jasrep.2017.02.022>.
- 1250 Wolfhagen, Jesse, Katheryn C. Twiss, Jacqui A. Mulville, and G. Arzu Demirergi. 2021. "Examining  
1251 Caprine Management and Cattle Domestication Through Biometric Analyses at Çatalhöyük East (North  
1252 and South Areas)." Book Section. In *Peopling the Landscape of Çatalhöyük: Reports from the 2009–2017*  
1253 *Seasons*, edited by Ian Hodder, 181–98. Los Angeles: Cotsen Institute of Archaeology Press.
- 1254 Wolverton, Steve. 2008. "Harvest Pressure and Environmental Carrying Capacity: An Ordinal-Scale Model  
1255 of Effects on Ungulate Prey." Journal Article. *American Antiquity* 73 (2): 179–99.
- 1256 Wright, Elizabeth, and Sarah Viner-Daniels. 2015. "Geographical Variation in the Size and Shape of the  
1257 European Aurochs (*Bos Primigenius*).". Journal Article. *Journal of Archaeological Science* 54: 8–22.

- 1260        <https://doi.org/10.1016/j.jas.2014.11.021>.
- 1261        Zeder, Melinda A. 2012. "Pathways to Animal Domestication." Book Section. In *Biodiversity in Agriculture: Domestication, Evolution, and Sustainability*, edited by P. Gepts, T. R. Famula, and R. L. Bettinger, 227–59. Cambridge: Cambridge University Press.
- 1264        Zeder, Melinda A., and Brian Hesse. 2000. "The Initial Domestication of Goats (*Capra Hircus*) in the Zagros Mountains 10,000 Years Ago." Journal Article. *Science* 287 (5461): 2254–57.
- 1266        Zeder, Melinda A., and Ximena Lemoine. 2020. "A Method for Constructing Demographic Profiles in Sus Scrofa Using Logarithm Size Index Scaling." Journal Article. *Journal of Archaeological Science* 116.
- 1268        <https://doi.org/10.1016/j.jas.2020.105115>.
- 1269        Zohary, D., Eitan Tchernov, and Liora Kolska Horwitz. 1998. "The Role of Unconscious Selection in the Domestication of Sheep and Goats." Journal Article. *Journal of Zoology, London* 245: 129–35.

1271 **Appendix 1 (Supplemental Table 6)**

1272 ***Posterior Summary Tables for Overall and Site-Level Model Parameters: Simulated Assem-***  
 1273 ***blages***

**Table 10:** Posterior Fit Summaries for Model Parameters (Single Assemblage Simulation)

Parameter	Posterior Mean	Posterior Median	Posterior SD	5% Quantile	95% Quantile	$\hat{R}$	Effective Sample Size (Bulk)	Effective Sample Size (Tail)
Single Assemblage Model $\theta_1$	0.16	0.15	0.04	0.10	0.23	1	6049	3074
Single Assemblage Model $\theta_2$	0.55	0.55	0.09	0.40	0.69	1	5087	2939
Single Assemblage Model $\theta_3$	0.29	0.29	0.09	0.16	0.45	1	4702	2688
Single Assemblage Model $\mu_1$	-0.10	-0.10	0.02	-0.12	-0.07	1	2580	2348
Single Assemblage Model $\mu_2$	-0.08	-0.08	0.01	-0.10	-0.05	1	2367	2360
Single Assemblage Model $\mu_3$	-0.01	-0.01	0.02	-0.04	0.01	1	2550	2714
Single Assemblage Model $\sigma_1$	0.05	0.05	0.00	0.04	0.06	1	6153	3082
Single Assemblage Model $\sigma_2$	0.04	0.04	0.00	0.04	0.05	1	5620	3156
Single Assemblage Model $\sigma_3$	0.04	0.04	0.00	0.04	0.05	1	6217	3368

**Table 11:** Posterior Fit Summaries for Model Parameters (Multisite Simulation)

Parameter	Posterior Mean	Posterior Median	Posterior SD	5% Quantile	95% Quantile	$\hat{R}$	Effective Sample Size (Bulk)	Effective Sample Size (Tail)
Multisite Model $\theta_1$	0.16	0.16	0.02	0.12	0.19	1	4144	3323
Multisite Model $\theta_2$	0.47	0.47	0.05	0.38	0.56	1	3529	3180
Multisite Model $\theta_3$	0.37	0.37	0.06	0.28	0.47	1	3246	3407
Multisite Model $\mu_1$	-0.10	-0.10	0.04	-0.16	-0.04	1	1539	2146
Multisite Model $\mu_2$	-0.09	-0.09	0.04	-0.15	-0.02	1	1523	2236
Multisite Model $\mu_3$	-0.01	-0.01	0.04	-0.07	0.06	1	1574	2062
Multisite Model $\sigma_1$	0.05	0.05	0.00	0.04	0.05	1	7355	3379
Multisite Model $\sigma_2$	0.04	0.04	0.00	0.04	0.05	1	5039	3218
Multisite Model $\sigma_3$	0.04	0.04	0.00	0.03	0.04	1	5405	3278

**Table 12:** Posterior Fit Summaries for Site-Level Model Parameters (Multisite Simulation)

Parameter	Posterior	Posterior	Posterior	5%	95%	$\hat{R}$	Effective	Effective
	Mean	Median	SD	Quantile	Quantile		Sample	Sample
							Size	Size
						(Bulk)		(Tail)
Site 1 $\theta_1$	0.16	0.16	0.04	0.11	0.23	1	5304	3451
Site 2 $\theta_1$	0.49	0.50	0.10	0.33	0.66	1	5358	3448
Site 3 $\theta_1$	0.34	0.34	0.11	0.17	0.52	1	5653	3503
Site 4 $\theta_1$	0.17	0.17	0.04	0.11	0.25	1	5950	3302
Site 5 $\theta_1$	0.45	0.45	0.11	0.27	0.62	1	4972	3478

**Table 12:** Posterior Fit Summaries for Site-Level Model Parameters (Multisite Simulation) (*continued*)

Parameter	Posterior	Posterior	Posterior	5%	95%	$\hat{R}$	Effective	Effective
	Mean	Median	SD	Quantile	Quantile		Sample	Sample
							Size (Bulk)	Size (Tail)
Site 6 $\theta_1$	0.38	0.37	0.11	0.20	0.58	1	5416	3485
Site 7 $\theta_1$	0.14	0.14	0.03	0.09	0.20	1	2975	3311
Site 8 $\theta_1$	0.34	0.34	0.09	0.18	0.49	1	4815	3327
Site 9 $\theta_1$	0.52	0.52	0.09	0.37	0.68	1	4927	3496
Site 10 $\theta_1$	0.18	0.17	0.04	0.12	0.26	1	3525	3322
Site 11 $\theta_1$	0.46	0.47	0.10	0.29	0.62	1	4115	2988
Site 12 $\theta_1$	0.36	0.35	0.10	0.19	0.54	1	5188	3581
Site 13 $\theta_1$	0.16	0.16	0.04	0.10	0.22	1	3174	3437
Site 14 $\theta_1$	0.47	0.47	0.07	0.34	0.59	1	4010	3347
Site 15 $\theta_1$	0.38	0.38	0.07	0.26	0.50	1	4407	3581
Site 1 $\theta_2$	0.16	0.16	0.04	0.11	0.22	1	5425	3617
Site 2 $\theta_2$	0.59	0.59	0.10	0.41	0.73	1	4464	2976
Site 3 $\theta_2$	0.25	0.24	0.10	0.10	0.43	1	4712	2977
Site 4 $\theta_2$	0.16	0.16	0.04	0.10	0.24	1	6403	3254
Site 5 $\theta_2$	0.54	0.55	0.11	0.35	0.71	1	4958	3252
Site 6 $\theta_2$	0.30	0.29	0.12	0.12	0.51	1	5537	3137
Site 7 $\theta_2$	0.16	0.16	0.04	0.10	0.22	1	4024	3302
Site 8 $\theta_2$	0.53	0.53	0.10	0.35	0.69	1	3380	3491
Site 9 $\theta_2$	0.32	0.31	0.10	0.16	0.50	1	3935	3376
Site 10 $\theta_2$	0.19	0.19	0.04	0.13	0.27	1	2665	3226
Site 11 $\theta_2$	0.57	0.57	0.09	0.41	0.71	1	4313	3469
Site 12 $\theta_2$	0.24	0.24	0.09	0.10	0.41	1	5093	3552
Site 13 $\theta_2$	0.25	0.24	0.07	0.16	0.37	1	1556	2133
Site 14 $\theta_2$	0.61	0.62	0.08	0.48	0.73	1	2398	3166
Site 15 $\theta_2$	0.14	0.13	0.05	0.06	0.23	1	4515	3380

**Table 12:** Posterior Fit Summaries for Site-Level Model Parameters (Multisite Simulation) (*continued*)

Parameter	Posterior	Posterior	Posterior	5%	95%	$\hat{R}$	Effective	Effective
	Mean	Median	SD	Quantile	Quantile		Sample	Sample
							Size (Bulk)	Size (Tail)
Site 1 $\theta_3$	0.14	0.14	0.04	0.08	0.19	1	3812	3688
Site 2 $\theta_3$	0.46	0.47	0.11	0.29	0.63	1	5744	3098
Site 3 $\theta_3$	0.40	0.39	0.11	0.22	0.58	1	6001	3199
Site 4 $\theta_3$	0.13	0.13	0.04	0.06	0.19	1	2707	3291
Site 5 $\theta_3$	0.38	0.38	0.12	0.18	0.57	1	5785	3678
Site 6 $\theta_3$	0.49	0.49	0.13	0.30	0.71	1	5195	3418
Site 7 $\theta_3$	0.14	0.14	0.03	0.08	0.19	1	2906	3264
Site 8 $\theta_3$	0.33	0.33	0.11	0.15	0.52	1	4235	3443
Site 9 $\theta_3$	0.53	0.53	0.12	0.34	0.72	1	4324	3168
Site 10 $\theta_3$	0.15	0.15	0.04	0.09	0.21	1	3329	3058
Site 11 $\theta_3$	0.38	0.38	0.10	0.21	0.55	1	4363	3451
Site 12 $\theta_3$	0.47	0.48	0.11	0.30	0.65	1	4804	2672
Site 13 $\theta_3$	0.14	0.14	0.04	0.08	0.20	1	2801	3183
Site 14 $\theta_3$	0.31	0.31	0.07	0.19	0.43	1	3858	3683
Site 15 $\theta_3$	0.55	0.55	0.08	0.43	0.68	1	4182	3722
Site 1 $\mu_1$	-0.09	-0.08	0.02	-0.12	-0.05	1	3147	3290
Site 2 $\mu_1$	-0.09	-0.09	0.03	-0.13	-0.04	1	4013	3495
Site 3 $\mu_1$	0.08	0.08	0.02	0.05	0.11	1	3571	3474
Site 4 $\mu_1$	-0.30	-0.30	0.02	-0.34	-0.27	1	3258	3185
Site 5 $\mu_1$	-0.10	-0.10	0.02	-0.13	-0.07	1	3723	3459
Site 6 $\mu_1$	-0.12	-0.12	0.02	-0.15	-0.10	1	3439	3181
Site 7 $\mu_1$	-0.09	-0.09	0.02	-0.13	-0.05	1	3748	3231
Site 8 $\mu_1$	0.07	0.07	0.02	0.04	0.10	1	3605	3549
Site 9 $\mu_1$	-0.34	-0.34	0.02	-0.37	-0.32	1	2618	3402
Site 10 $\mu_1$	-0.10	-0.11	0.02	-0.13	-0.08	1	3046	3398

**Table 12:** Posterior Fit Summaries for Site-Level Model Parameters (Multisite Simulation) (*continued*)

Parameter	Posterior	Posterior	Posterior	5%	95%	$\hat{R}$	Effective	Effective
	Mean	Median	SD	Quantile	Quantile		Sample	Sample
							Size (Bulk)	Size (Tail)
Site 11 $\mu_1$	-0.08	-0.08	0.02	-0.11	-0.05	1	3756	3689
Site 12 $\mu_1$	-0.09	-0.09	0.03	-0.14	-0.05	1	4229	3256
Site 13 $\mu_1$	0.09	0.09	0.02	0.05	0.12	1	3577	3764
Site 14 $\mu_1$	-0.32	-0.32	0.02	-0.36	-0.29	1	3862	3616
Site 15 $\mu_1$	-0.09	-0.09	0.02	-0.12	-0.06	1	3308	3433
Site 1 $\mu_2$	-0.07	-0.07	0.02	-0.10	-0.04	1	3139	3517
Site 2 $\mu_2$	-0.07	-0.07	0.03	-0.12	-0.03	1	4132	3698
Site 3 $\mu_2$	0.10	0.10	0.02	0.06	0.13	1	3566	3580
Site 4 $\mu_2$	-0.29	-0.29	0.02	-0.32	-0.26	1	3144	3192
Site 5 $\mu_2$	-0.08	-0.08	0.02	-0.11	-0.06	1	3635	3549
Site 6 $\mu_2$	-0.11	-0.11	0.02	-0.14	-0.08	1	3292	3486
Site 7 $\mu_2$	-0.07	-0.07	0.02	-0.11	-0.04	1	3633	3061
Site 8 $\mu_2$	0.08	0.09	0.02	0.06	0.11	1	3554	3601
Site 9 $\mu_2$	-0.33	-0.33	0.02	-0.36	-0.30	1	2394	3112
Site 10 $\mu_2$	-0.09	-0.09	0.02	-0.11	-0.06	1	2887	3457
Site 11 $\mu_2$	-0.06	-0.06	0.02	-0.09	-0.03	1	3668	3467
Site 12 $\mu_2$	-0.07	-0.07	0.03	-0.12	-0.03	1	4224	3519
Site 13 $\mu_2$	0.10	0.10	0.02	0.07	0.13	1	3588	3845
Site 14 $\mu_2$	-0.31	-0.31	0.02	-0.34	-0.27	1	3825	3615
Site 15 $\mu_2$	-0.07	-0.07	0.02	-0.10	-0.04	1	3054	3549
Site 1 $\mu_3$	-0.01	-0.01	0.02	-0.04	0.03	1	3005	3538
Site 2 $\mu_3$	0.02	0.02	0.03	-0.02	0.07	1	3982	3477
Site 3 $\mu_3$	0.20	0.20	0.02	0.16	0.23	1	3683	3643
Site 4 $\mu_3$	-0.22	-0.23	0.02	-0.26	-0.19	1	3666	3573
Site 5 $\mu_3$	0.17	0.17	0.02	0.14	0.20	1	3046	3372

**Table 12:** Posterior Fit Summaries for Site-Level Model Parameters (Multisite Simulation) (*continued*)

Parameter	Posterior	Posterior	Posterior	5%	95%	$\hat{R}$	Effective	Effective
	Mean	Median	SD	Quantile	Quantile		Sample	Sample
							Size (Bulk)	Size (Tail)
Site 6 $\mu_3$	-0.05	-0.06	0.03	-0.09	0.00	1	3589	3281
Site 7 $\mu_3$	0.00	-0.01	0.04	-0.05	0.06	1	4161	3492
Site 8 $\mu_3$	0.15	0.15	0.03	0.11	0.20	1	3669	3493
Site 9 $\mu_3$	-0.26	-0.27	0.03	-0.30	-0.22	1	3611	3648
Site 10 $\mu_3$	0.14	0.14	0.04	0.07	0.20	1	3534	2869
Site 11 $\mu_3$	0.00	0.00	0.02	-0.04	0.03	1	4004	3837
Site 12 $\mu_3$	0.00	0.00	0.03	-0.04	0.05	1	4345	3491
Site 13 $\mu_3$	0.16	0.16	0.02	0.13	0.19	1	3543	3362
Site 14 $\mu_3$	-0.23	-0.23	0.02	-0.27	-0.20	1	3913	3876
Site 15 $\mu_3$	0.17	0.17	0.02	0.14	0.20	1	3493	3140
Site 1 $\sigma_1$	0.05	0.05	0.00	0.04	0.05	1	6814	3498
Site 2 $\sigma_1$	0.05	0.05	0.00	0.04	0.05	1	6055	3324
Site 3 $\sigma_1$	0.05	0.05	0.00	0.04	0.05	1	6527	3465
Site 4 $\sigma_1$	0.05	0.05	0.00	0.04	0.05	1	6019	3341
Site 5 $\sigma_1$	0.05	0.05	0.00	0.04	0.05	1	6164	3238
Site 6 $\sigma_1$	0.05	0.05	0.00	0.04	0.05	1	5797	3189
Site 7 $\sigma_1$	0.05	0.05	0.00	0.04	0.05	1	6146	3283
Site 8 $\sigma_1$	0.05	0.05	0.00	0.04	0.05	1	5800	3265
Site 9 $\sigma_1$	0.05	0.05	0.00	0.04	0.05	1	6374	3280
Site 10 $\sigma_1$	0.05	0.05	0.01	0.04	0.05	1	6445	3225
Site 11 $\sigma_1$	0.05	0.05	0.00	0.04	0.05	1	6186	3486
Site 12 $\sigma_1$	0.05	0.05	0.00	0.04	0.05	1	6500	3482
Site 13 $\sigma_1$	0.05	0.05	0.00	0.04	0.05	1	5610	3059
Site 14 $\sigma_1$	0.05	0.05	0.00	0.04	0.05	1	6082	3516
Site 15 $\sigma_1$	0.05	0.05	0.00	0.04	0.05	1	5356	3515

**Table 12:** Posterior Fit Summaries for Site-Level Model Parameters (Multisite Simulation) (*continued*)

Parameter	Posterior	Posterior	Posterior	5%	95%	$\hat{R}$	Effective	Effective
	Mean	Median	SD	Quantile	Quantile		Sample	Sample
							Size (Bulk)	Size (Tail)
Site 1 $\sigma_2$	0.04	0.04	0.00	0.04	0.05	1	4285	3288
Site 2 $\sigma_2$	0.04	0.04	0.00	0.03	0.05	1	4846	3415
Site 3 $\sigma_2$	0.04	0.04	0.00	0.03	0.05	1	4497	3379
Site 4 $\sigma_2$	0.04	0.04	0.00	0.04	0.05	1	4278	3329
Site 5 $\sigma_2$	0.04	0.04	0.00	0.03	0.05	1	4497	3378
Site 6 $\sigma_2$	0.04	0.04	0.00	0.03	0.05	1	4255	3939
Site 7 $\sigma_2$	0.04	0.04	0.00	0.03	0.05	1	3878	3332
Site 8 $\sigma_2$	0.04	0.04	0.00	0.03	0.05	1	4095	3456
Site 9 $\sigma_2$	0.04	0.04	0.00	0.03	0.05	1	4661	3448
Site 10 $\sigma_2$	0.04	0.04	0.00	0.03	0.05	1	4130	3547
Site 11 $\sigma_2$	0.04	0.04	0.00	0.03	0.05	1	4440	3362
Site 12 $\sigma_2$	0.04	0.04	0.00	0.03	0.05	1	4216	3590
Site 13 $\sigma_2$	0.04	0.04	0.00	0.03	0.05	1	4757	3407
Site 14 $\sigma_2$	0.04	0.04	0.00	0.04	0.05	1	4436	3141
Site 15 $\sigma_2$	0.04	0.04	0.00	0.03	0.05	1	4796	3449
Site 1 $\sigma_3$	0.04	0.04	0.00	0.03	0.05	1	4378	3465
Site 2 $\sigma_3$	0.04	0.04	0.00	0.03	0.05	1	4342	3351
Site 3 $\sigma_3$	0.04	0.04	0.00	0.03	0.05	1	4301	3492
Site 4 $\sigma_3$	0.04	0.04	0.00	0.03	0.05	1	4638	3445
Site 5 $\sigma_3$	0.04	0.04	0.00	0.03	0.04	1	4687	3613
Site 6 $\sigma_3$	0.04	0.04	0.00	0.03	0.05	1	4418	3216
Site 7 $\sigma_3$	0.04	0.04	0.00	0.03	0.05	1	4111	3079
Site 8 $\sigma_3$	0.04	0.04	0.00	0.03	0.05	1	4403	3249
Site 9 $\sigma_3$	0.04	0.04	0.00	0.03	0.05	1	4651	3346
Site 10 $\sigma_3$	0.04	0.04	0.00	0.03	0.05	1	4459	3625

**Table 12:** Posterior Fit Summaries for Site-Level Model Parameters (Multisite Simulation) (*continued*)

Parameter	Posterior	Posterior	Posterior	5%	95%	$\hat{R}$	Effective	Effective
	Mean	Median	SD	Quantile	Quantile		Sample	Sample
							Size	Size
Site 11 $\sigma_3$	0.04	0.04	0.00	0.03	0.05	1	4625	3360
Site 12 $\sigma_3$	0.04	0.04	0.00	0.03	0.05	1	4316	3605
Site 13 $\sigma_3$	0.04	0.04	0.00	0.03	0.05	1	4486	3579
Site 14 $\sigma_3$	0.04	0.04	0.00	0.03	0.04	1	4498	3565
Site 15 $\sigma_3$	0.04	0.04	0.00	0.03	0.04	1	4484	3238

<sup>1274</sup> **Appendix 2 (Supplemental Table 7)**

<sup>1275</sup> ***Posterior Summary Tables for Overall and Site-Level Model Parameters: Archaeological Case Studies***

**Table 13:** Posterior Fit Summaries for Model Parameters (Pinarbaşı B Sheep)

Parameter	Posterior Mean	Posterior Median	Posterior SD	5% Quantile	95% Quantile	$\hat{R}$	Effective Sample Size (Bulk)	Effective Sample Size (Tail)
Pinarbaşı B Sheep $\theta_1$	0.87	0.87	0.04	0.79	0.93	1	6174	2962
Pinarbaşı B Sheep $\theta_2$	0.08	0.08	0.04	0.02	0.15	1	5319	3352
Pinarbaşı B Sheep $\theta_3$	0.05	0.05	0.04	0.01	0.13	1	6247	3246
Pinarbaşı B Sheep $\mu_1$	-0.13	-0.13	0.02	-0.17	-0.09	1	3559	3130
Pinarbaşı B Sheep $\mu_2$	-0.10	-0.10	0.02	-0.14	-0.06	1	3389	3209
Pinarbaşı B Sheep $\mu_3$	-0.03	-0.03	0.03	-0.07	0.01	1	3569	2908
Pinarbaşı B Sheep $\sigma_1$	0.06	0.06	0.01	0.05	0.07	1	5239	3121
Pinarbaşı B Sheep $\sigma_2$	0.05	0.04	0.00	0.04	0.05	1	10190	2813
Pinarbaşı B Sheep $\sigma_3$	0.05	0.05	0.00	0.04	0.05	1	9809	2766

**Table 14:** Posterior Fit Summaries for Model Parameters (Northwest Anatolian Cattle)

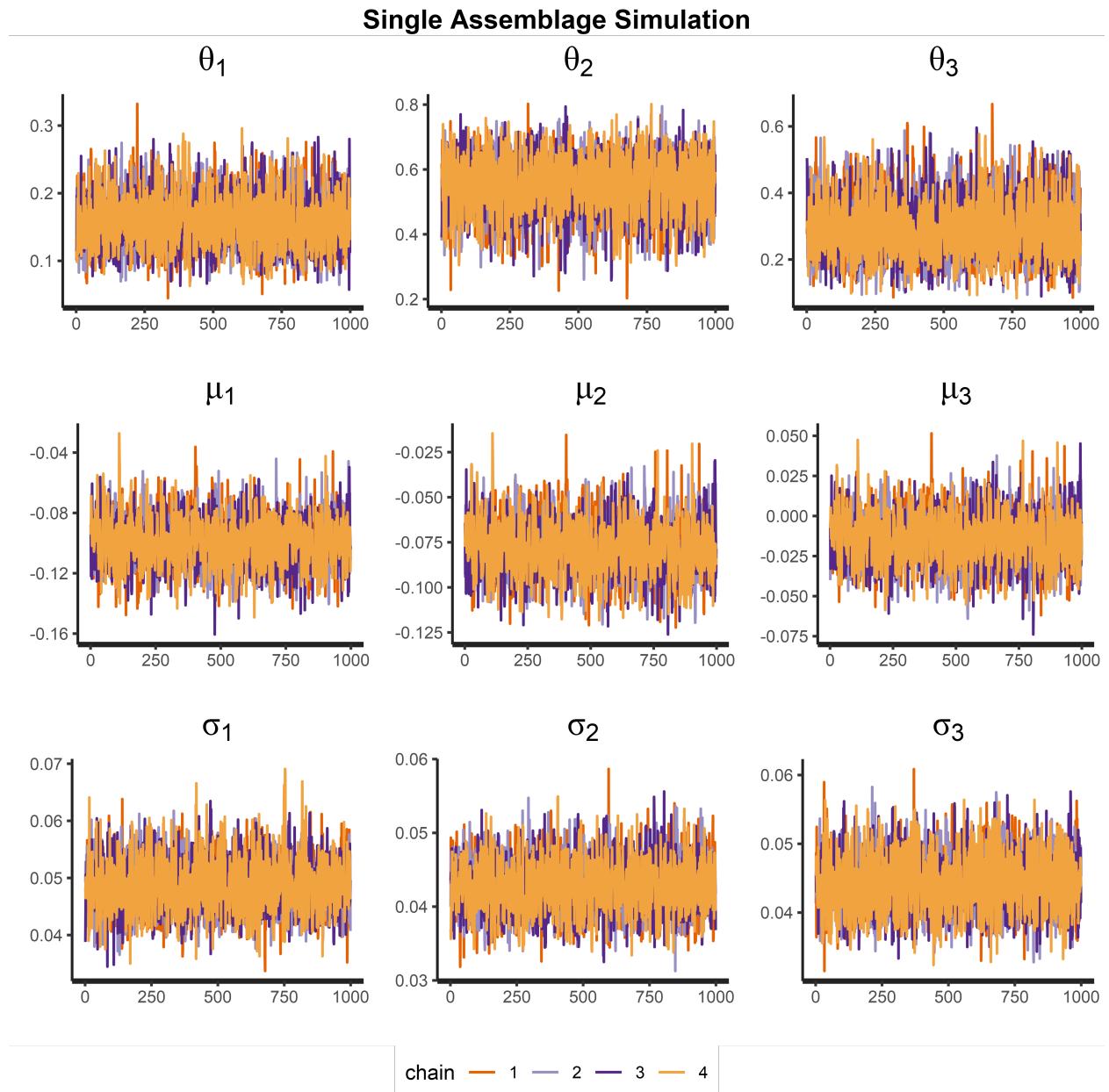
Parameter	Posterior Mean	Posterior Median	Posterior SD	5% Quantile	95% Quantile	$\hat{R}$	Effective Sample Size (Bulk)	Effective Sample Size (Tail)
NW Anatolian Cattle $\theta_1$	0.17	0.16	0.05	0.09	0.26	1	1746	2189
NW Anatolian Cattle $\theta_2$	0.67	0.68	0.06	0.56	0.77	1	1739	2254
NW Anatolian Cattle $\theta_3$	0.16	0.16	0.05	0.09	0.24	1	1685	2126
NW Anatolian Cattle $\mu_1$	-0.26	-0.26	0.03	-0.32	-0.21	1	1901	2118
NW Anatolian Cattle $\mu_2$	-0.20	-0.20	0.03	-0.25	-0.15	1	1767	2050
NW Anatolian Cattle $\mu_3$	-0.05	-0.05	0.03	-0.10	0.00	1	1841	2292
NW Anatolian Cattle $\sigma_1$	0.06	0.05	0.01	0.05	0.06	1	4127	3408
NW Anatolian Cattle $\sigma_2$	0.05	0.05	0.00	0.04	0.05	1	2513	2582
NW Anatolian Cattle $\sigma_3$	0.05	0.05	0.00	0.04	0.05	1	2941	3005

**Table 15:** Posterior Fit Summaries for Site-Level Model Parameters (Northwest Anatolian Cattle)

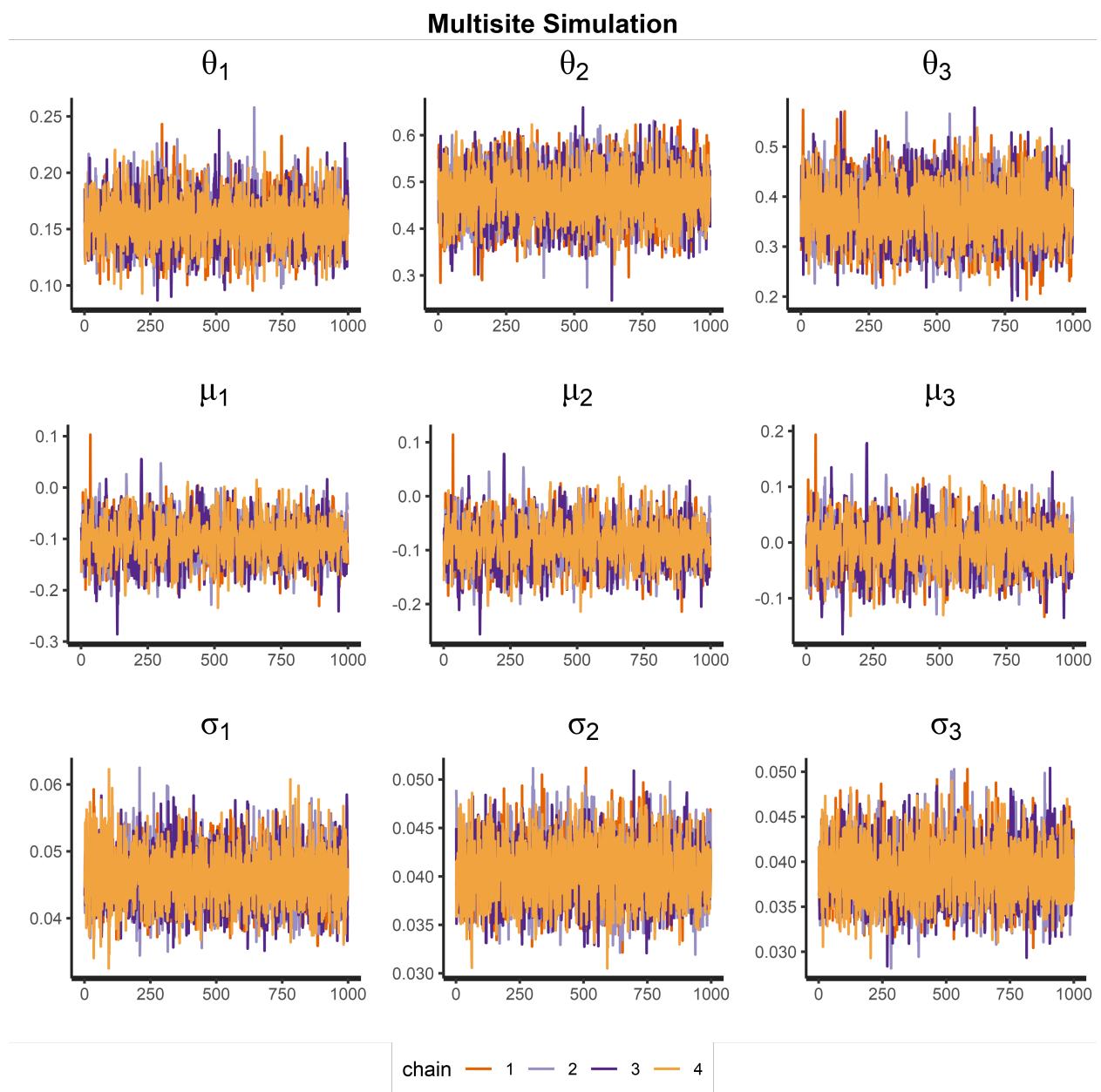
Parameter	Posterior Mean	Posterior Median	Posterior SD	5% Quantile	95% Quantile	$\hat{R}$	Effective Sample Size (Bulk)	Effective Sample Size (Tail)
Barcın $\theta_1$	0.33	0.33	0.04	0.27	0.40	1	3932	3536
Neolithic İlpınar $\theta_1$	0.52	0.52	0.06	0.42	0.60	1	2322	2896
Menteşe $\theta_1$	0.15	0.15	0.05	0.08	0.24	1	2190	3336
Chalcolithic İlpınar $\theta_1$	0.13	0.13	0.02	0.10	0.17	1	4136	3519
Barcın $\theta_2$	0.75	0.75	0.06	0.64	0.84	1	1115	2651
Neolithic İlpınar $\theta_2$	0.12	0.12	0.06	0.03	0.22	1	1027	2511
Menteşe $\theta_2$	0.22	0.22	0.06	0.14	0.33	1	4965	3208
Chalcolithic İlpınar $\theta_2$	0.63	0.63	0.07	0.52	0.74	1	3136	3198
Barcın $\theta_3$	0.14	0.14	0.05	0.07	0.23	1	2259	2766
Neolithic İlpınar $\theta_3$	0.07	0.07	0.02	0.04	0.10	1	3781	3215
Menteşe $\theta_3$	0.73	0.73	0.04	0.65	0.79	1	2483	3228
Chalcolithic İlpınar $\theta_3$	0.21	0.20	0.04	0.14	0.28	1	2087	2936
Barcın $\mu_1$	-0.40	-0.40	0.03	-0.44	-0.36	1	2568	3173
Neolithic İlpınar $\mu_1$	-0.24	-0.24	0.02	-0.28	-0.21	1	2022	2658
Menteşe $\mu_1$	-0.25	-0.25	0.03	-0.30	-0.20	1	2377	2891
Chalcolithic İlpınar $\mu_1$	-0.29	-0.29	0.03	-0.33	-0.25	1	2270	3036
Barcın $\mu_2$	-0.18	-0.18	0.02	-0.22	-0.14	1	2045	2507
Neolithic İlpınar $\mu_2$	-0.21	-0.21	0.02	-0.24	-0.17	1	1803	2453
Menteşe $\mu_2$	-0.19	-0.19	0.02	-0.23	-0.15	1	2074	2711
Chalcolithic İlpınar $\mu_2$	-0.24	-0.24	0.02	-0.27	-0.21	1	1916	2416
Barcın $\mu_3$	-0.02	-0.02	0.03	-0.07	0.03	1	2032	2723
Neolithic İlpınar $\mu_3$	-0.07	-0.07	0.03	-0.12	-0.03	1	1581	2871
Menteşe $\mu_3$	-0.04	-0.04	0.03	-0.09	0.01	1	2239	2545
Chalcolithic İlpınar $\mu_3$	-0.09	-0.09	0.02	-0.13	-0.06	1	2041	2233
Barcın $\sigma_1$	0.06	0.06	0.01	0.05	0.07	1	4302	3717
Neolithic İlpınar $\sigma_1$	0.06	0.06	0.01	0.05	0.07	1	2964	3419
Menteşe $\sigma_1$	0.06	0.06	0.01	0.05	0.07	1	3950	3607
Chalcolithic İlpınar $\sigma_1$	0.06	0.05	0.01	0.05	0.07	1	3839	3731
Barcın $\sigma_2$	0.04	0.04	0.01	0.03	0.06	1	2503	2941
Neolithic İlpınar $\sigma_2$	0.07	0.07	0.01	0.05	0.08	1	1429	1427
Menteşe $\sigma_2$	0.05	0.05	0.01	0.04	0.06	1	3077	3188
Chalcolithic İlpınar $\sigma_2$	0.04	0.04	0.01	0.03	0.05	1	1552	2146
Barcın $\sigma_3$	0.05	0.05	0.01	0.04	0.06	1	2702	3106
Neolithic İlpınar $\sigma_3$	0.05	0.05	0.01	0.04	0.05	1	3433	3506
Menteşe $\sigma_3$	0.05	0.05	0.00	0.04	0.05	1	2786	3165
Chalcolithic İlpınar $\sigma_3$	0.05	0.05	0.00	0.04	0.05	1	2704	3247

<sup>1277</sup> Appendix 3 (Supplemental Figures 1-4)

<sup>1278</sup> *Traceplots of Posterior Distributions of Overall Model Parameters*

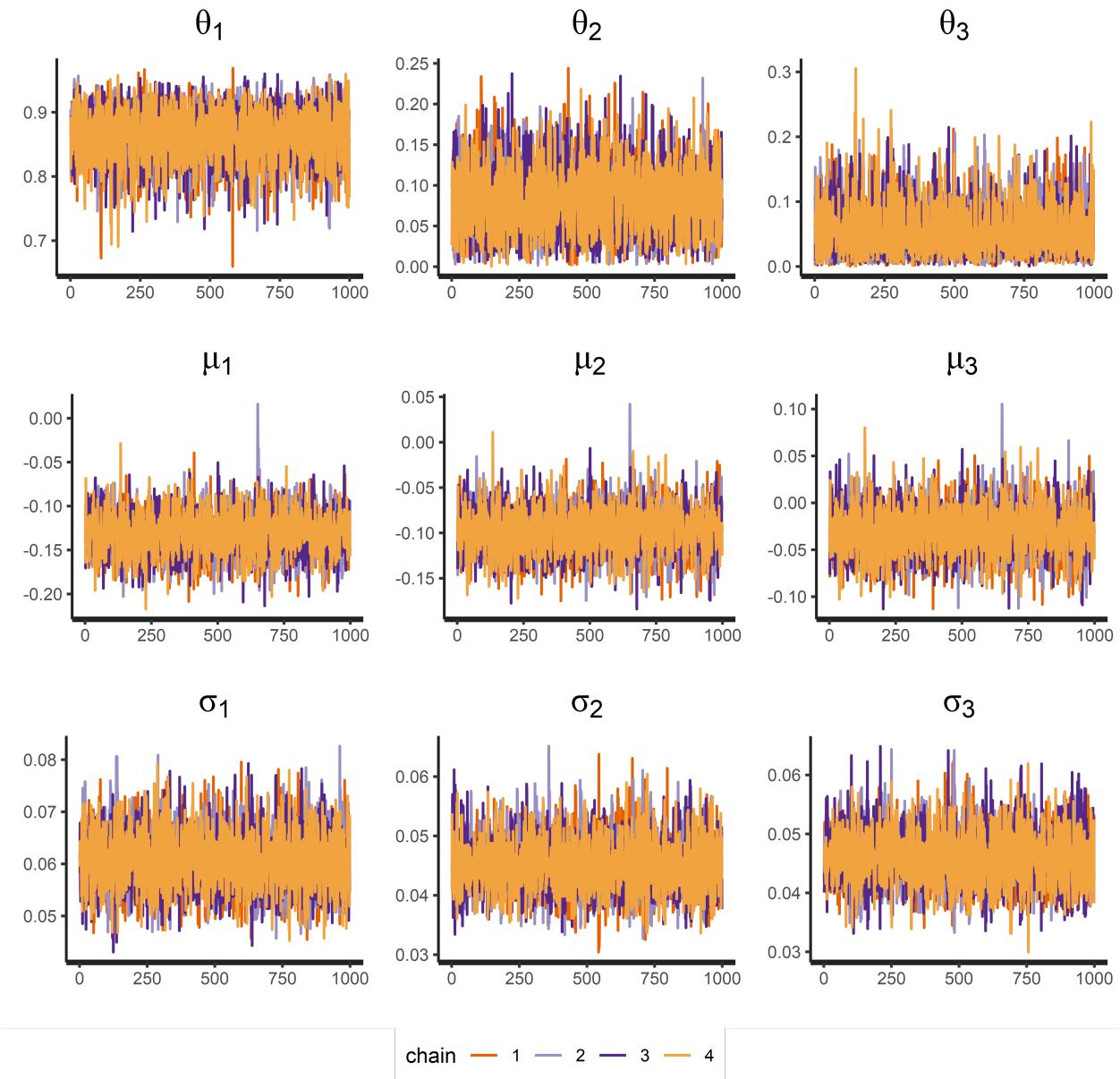


**Figure 15:** Traceplots of Model Parameters (Single Assemblage Simulation)



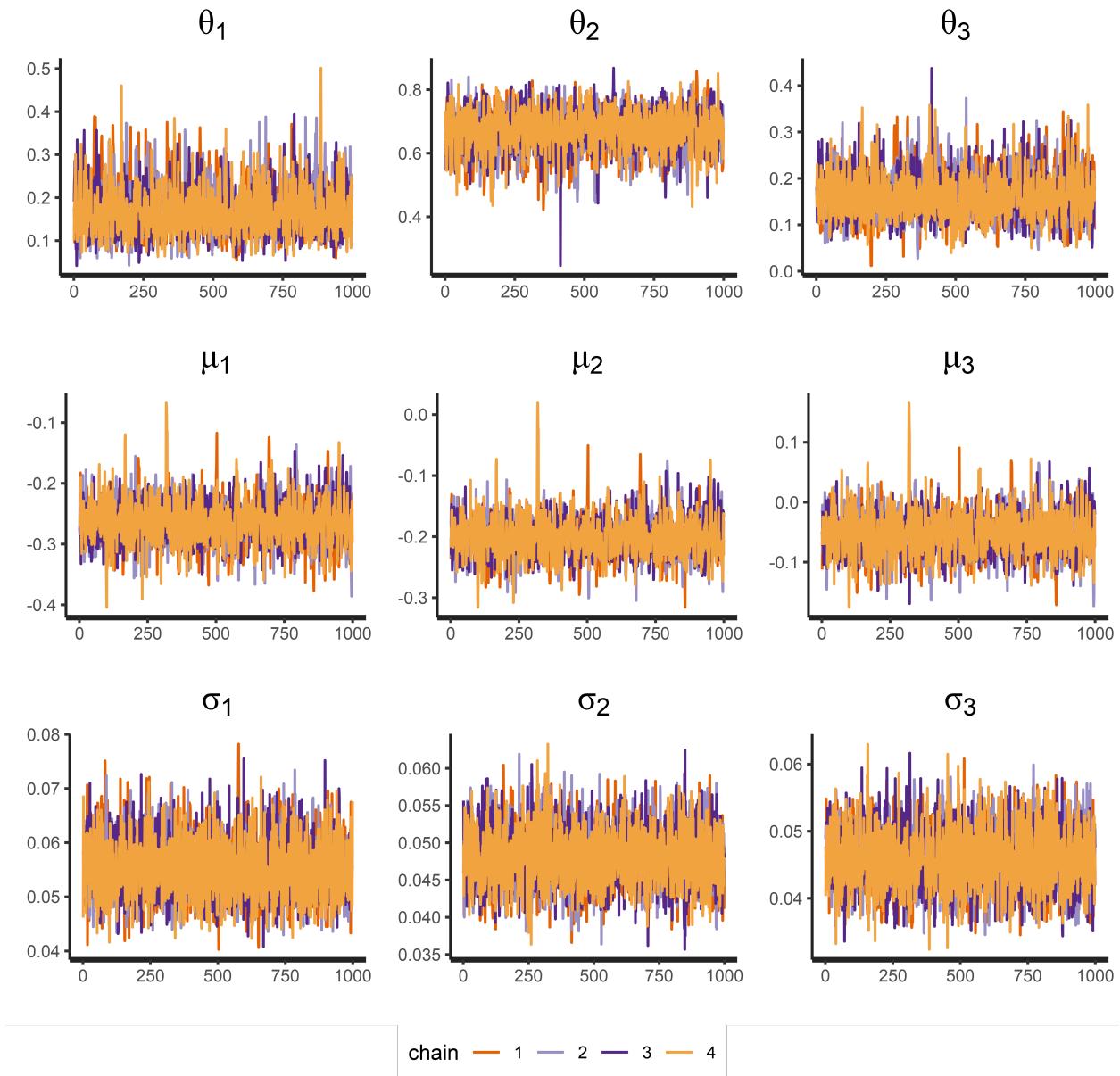
**Figure 16:** Traceplots of Model Parameters (Multisite Simulation)

### Pınarbaşı B Sheep



**Figure 17:** Traceplots of Model Parameters (Pınarbaşı B Sheep)

### Northwest Anatolian Cattle

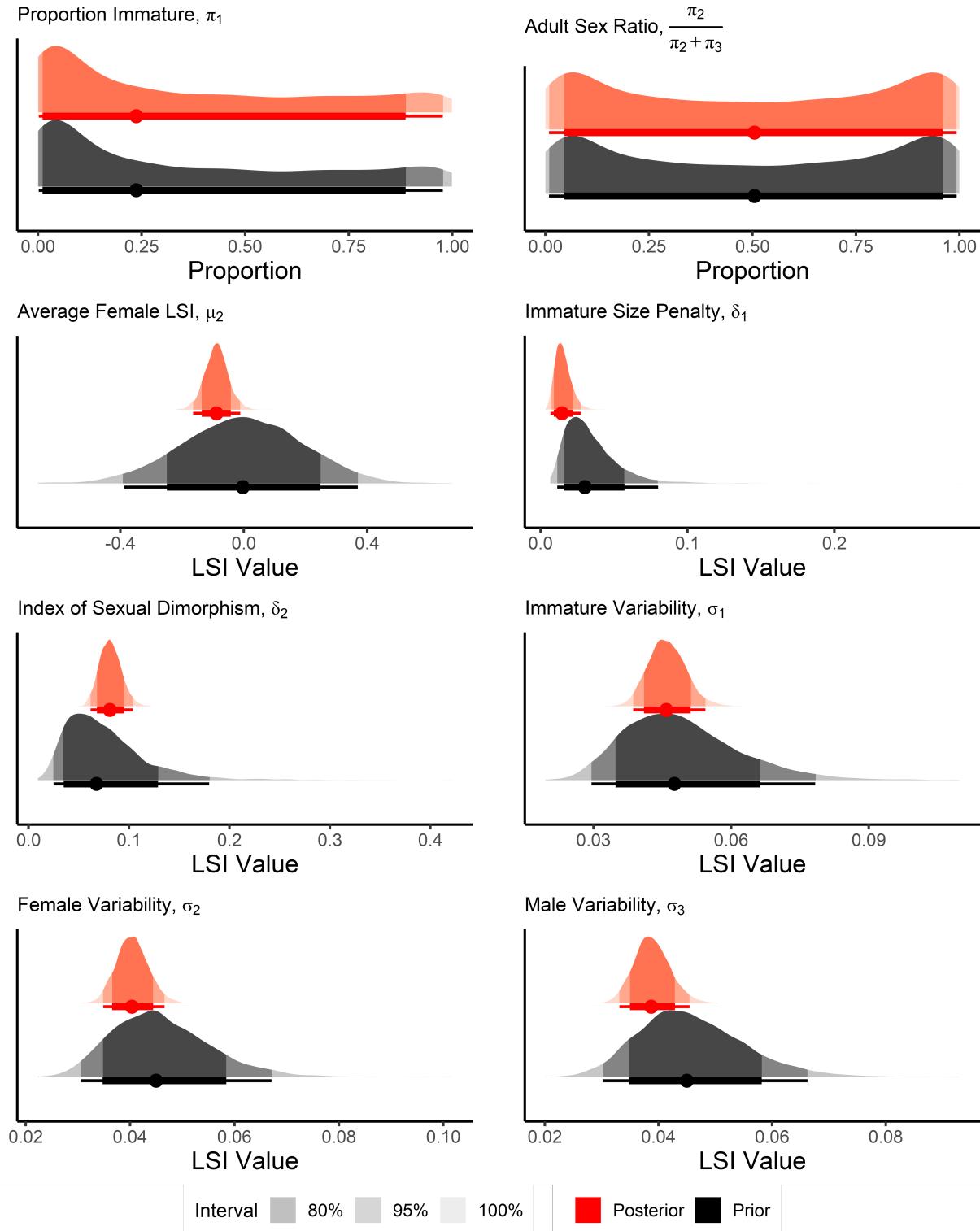


**Figure 18:** Traceplots of Model Parameters (Northwest Anatolian Cattle)

<sup>1279</sup> Appendix 4 (Supplemental Figures 5-7)

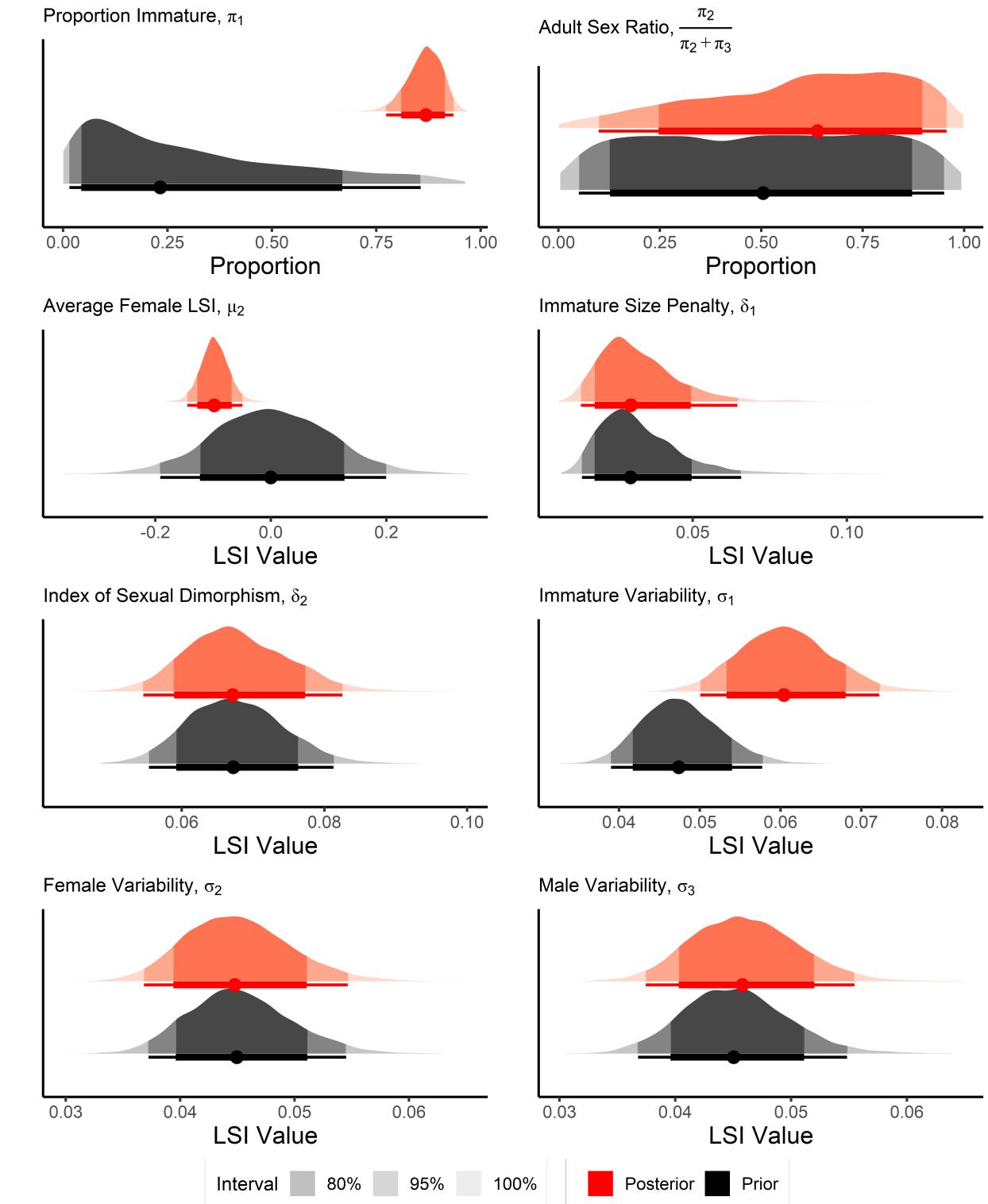
<sup>1280</sup> *Prior-Posterior Comparisons for Multisite, Pinarbaşı B, and NW Anatolian Model Hyper-*  
<sup>1281</sup> *Parameters*

### Multisite Simulation



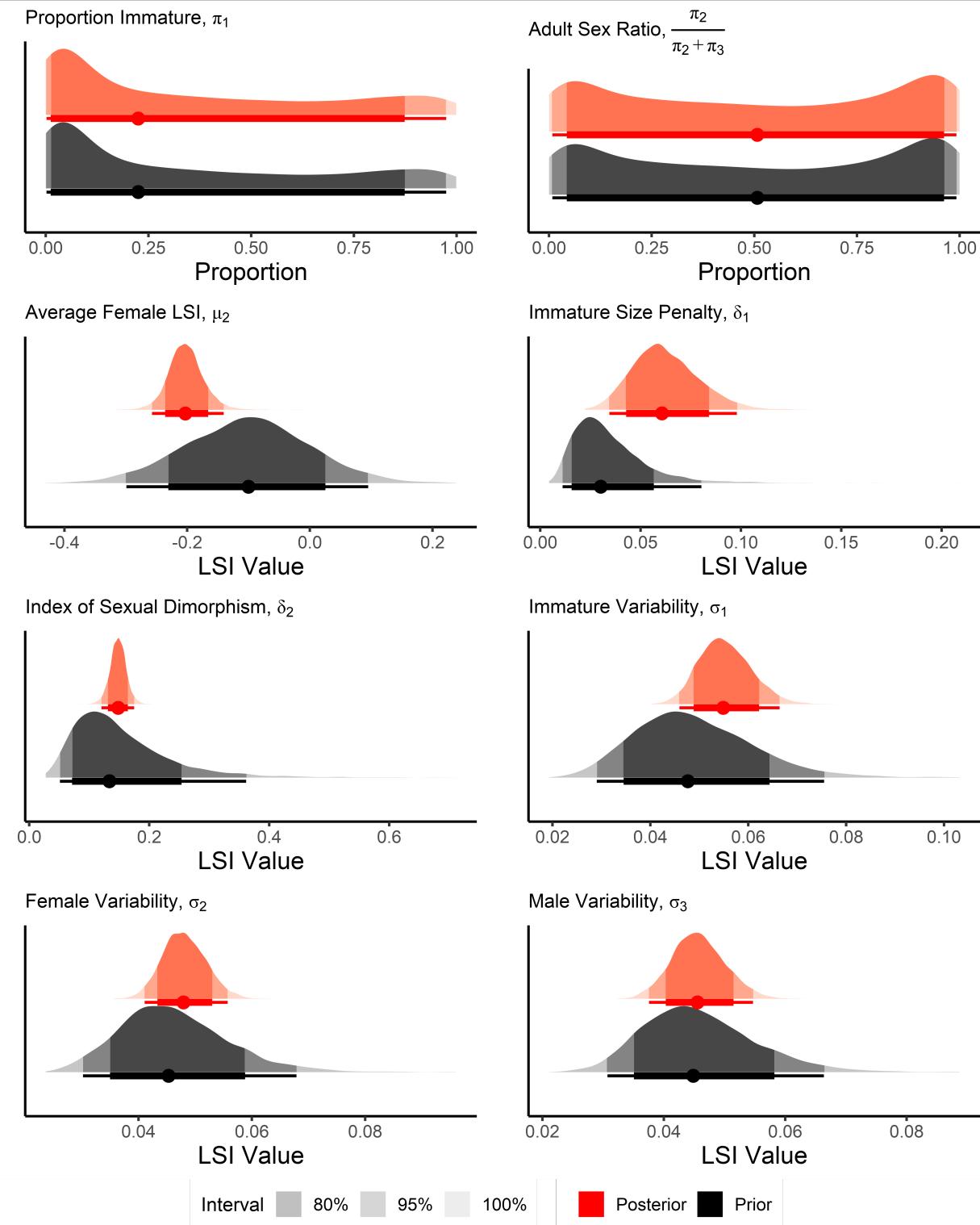
**Figure 19:** Prior-Posterior Comparison of Multisite Simulation Model Hyper-Parameters

### Pinarbaşı B Sheep



**Figure 20:** Prior-Posterior Comparison of Pinarbaşı B Sheep Model Hyper-Parameters

## Northwest Anatolian Cattle



**Figure 21:** Prior-Posterior Comparison of NW Anatolian Cattle Model Hyper-Parameters

1282 ***Model Supplement: A Bayesian Multilevel Mixture Model for***  
1283 ***Zooarchaeological Measurements***

1284 The Bayesian model developed for this paper describes assemblages of faunal measurements as a mixture  
1285 of immature animals, (adult-sized) females, and (adult-sized) males that have distinct average body sizes  
1286 and expected variation around that average size. The model uses multiple measured dimensions (e.g.,  
1287 humerus distal breadth “humerus Bd,” radius proximal breadth “radius Bp,” abbreviations following Driesch  
1288 1976), which are first converted into a logarithmic size index, or LSI, values with a natural logarithm base  
1289 (Meadow 1999; Wolfhagen 2020). LSI observations from measurement sets are grouped together within  
1290 a specimen to create individuals grouped into defined “element portions” that serve as the basis for the  
1291 mixture model analysis. Element portions relate to categories used for element fusion (e.g., distal humerus)  
1292 to relate biometry and mortality profiles; specimens that contain multiple element portions—like complete  
1293 limb bones—are grouped into the latest-fusing element portion (compare to “skeletal part type” in Breslawski  
1294 2022).

1295 The multilevel structure of the model uses partial pooling to allow the parameters of the mixture model  
1296 to vary between element portions while resisting overfitting. These element portion-specific parameters are  
1297 related to each other through hyper-parameters, which describe the average value of the model parameters  
1298 and the variability of model parameters across element portions (Wolfhagen 2020). The following sections  
1299 describe the details of the multilevel mixture model, including the model’s likelihood, the ways that the direct  
1300 observations of measurements and demographic data are used by the model to account for measurement error,  
1301 and the development of prior distributions for the model’s hyper-parameters and for parameters that govern  
1302 the model’s multilevel structure. Finally, this supplement provides the full sets of equations for the single-site  
1303 and multisite Bayesian multilevel mixture models and the prior distribution definitions used in the model  
1304 applications described in the main text.

1305 ***1. Definition of the Bayesian Multilevel Model***

1306 The central likelihood of the mixture model uses parameters that are specific to each element portion. These  
1307 parameters include the relative proportions for the different animal groups: immature animals, females, and  
1308 males ( $\pi_1, \pi_2, \pi_3$ ), the average size for each group ( $\mu_1, \mu_2, \mu_3$ ), and the standard deviation for each group  
1309 ( $\sigma_1, \sigma_2, \sigma_3$ ). For each element portion, immature animals are described with the first set of parameters ( $\pi_1,$   
1310  $\mu_1, \sigma_1$ ), adult-sized females with the second set of parameters ( $\pi_2, \mu_2, \sigma_2$ ), and adult-sized males with the  
1311 third set of parameters ( $\pi_3, \mu_3, \sigma_3$ ). This results in both a set of parameters that describe the composition

1312 of the assemblage (of measurements from that element portion) and an equation to estimate the probability  
1313 that a particular specimen comes from an immature, adult female, or adult male individual.

1314 *Mixture Model Likelihood Equation:*

$$\begin{aligned} P(x|\pi_1, \pi_2, \pi_3, \mu_1, \mu_2, \mu_3, \sigma_1, \sigma_2, \sigma_3) = \\ \pi_1 * \text{Normal}(x, \mu_1, \sigma_1) + \\ \pi_2 * \text{Normal}(x, \mu_2, \sigma_2) + \\ \pi_3 * \text{Normal}(x, \mu_3, \sigma_3) \end{aligned} \quad (1)$$

1315 In addition to a specimen's LSI value, the model needs two additional observed variables to address the  
1316 potential presence of immature animals in the assemblage. First, an indicator variable Immature[specimen]  
1317 describes whether the specimen could be from an immature animal based on the body part and the fusion  
1318 characteristics (1 = potentially immature, 0 = cannot be immature). Data from known-age Shetland sheep  
1319 show that specimens killed at younger than one year of age are significantly smaller than those killed at  
1320 older ages, regardless of fusion status (Popkin et al. 2012). Thus, any measurement from an element with  
1321 an unfused epiphysis or from an element that does not fuse or could fuse before one year of age is considered  
1322 potentially immature. Measurements from specimens with fused epiphyses that fuse after one year of age  
1323 are considered ineligible to be immature, so the model does not consider that probability (it considers  $\pi_1 =$   
1324 0 for fitting that specimen).

1325 Second, the proportion of specimens from an element portion that could be immature ( $\text{proportion}_{\text{immature}}$ )  
1326 determines how to re-weight the mixture components ( $\pi_1$ ,  $\pi_2$ , and  $\pi_3$ ) for potentially-immature specimens  
1327 from that element portion. The mixture components describe the entire assemblage for an element portion  
1328 (a combination of potentially-immature and non-immature specimens), meaning that if  $\pi_1 = 0.25$  we should  
1329 expect 25% of the specimens to be from immature animals. If every specimen could be immature—say, for  
1330 specimens from an early-fusing element—then the mixture components do not need to be re-weighted. If,  
1331 however, only half of the specimens could be immature, then the mixture components must be re-weighted  
1332 to ensure that the whole-assemblage proportions are correct. In such a case, we would expect half of the  
1333 potentially-immature animals to be from immature animals if  $\pi_1 = 0.25$  for the whole assemblage and  
1334 we know that  $\text{proportion}_{\text{immature}} = (\frac{0.25}{0.50} = 0.50)$ ; this same re-weighting would make it less likely that  
1335 potentially-immature animals are from adult-sized female or adult-sized male animals. The code includes  
1336 checks to ensure that  $\pi_1$  can never exceed 1.00 after accounting for the proportion of immature specimens in  
1337 cases where there are very few potentially-immature specimens and/or a relatively high expected proportion  
1338 of immature animals in the assemblage.

1339 **2. Measurement Error and Observations**

1340 The model estimates measurement error for different observed quantities that are input into the model.  
1341 Measurements on both the archaeological specimens and the standard values used to calculate LSI values  
1342 are assumed to have a 1% measurement error (Breslawski and Byers 2015; Popkin et al. 2012: Figure  
1343 6). This 1% value comes from an evaluation of the Breslawski and Byers (2015) measurement data, where  
1344 the average standard deviation of repeated measurements on bison radius proximal breadth measurements  
1345 was 1.1% the average value of the measurement. This means that each measurement is given a standard  
1346 deviation based on the observed value, which is used to estimate the “modeled” measurement value based on  
1347 the observation. These modeled measurements are then used to calculate the LSI value for that observation  
1348 ( $LSI_{Measurement}$ ).

1349 Because specimens can have multiple measured dimensions that are included in the mixture model on  
1350 them (e.g., a distal humerus with both Bd and BT observations or a complete radius with Bp and Bd  
1351 observations), the mixture model calculates specimen-specific LSI values ( $LSI_{Specimen}$ ) that are related to  
1352 the observed measurement-specific LSI values ( $LSI_{Measurement}$ ).  $LSI_{Measurement}$  values are the “observations”  
1353 with a standard deviation of 0.01 (in  $LSI_e$  scale) based on intra-individual variation of  $LSI_e$  values for the  
1354 Popkin et al. (2012) sheep using the *Ovis orientalis* female standard animal (FMC 57951) from Uerpmann  
1355 and Uerpmann (1994, Table 12).

1356 *Observation Error Equations for Measurements:*

$$\begin{aligned}\sigma_{measurement} &= Measurement_{observed} * 0.01 \\ Measurement_{observed} &\sim \text{Normal}(Measurement_{modeled}, \sigma_{measurement}) \\ \sigma_{reference} &= Reference_{observed} * 0.01 \\ Reference_{observed} &\sim \text{Normal}(Reference_{modeled}, \sigma_{reference}) \\ LSI_{measurement} &= \log_e(Measurement_{modeled}) - \log_e(Reference_{modeled}) \\ LSI_{measurement} &\sim \text{Normal}(LSI_{specimen}, 0.01)\end{aligned}\tag{2}$$

1357 The model also uses observations of sex ratios and fusion rates to estimate assemblage-level demographic  
1358 proportions. This allows relevant data to inform the model about the expected relative proportions of  
1359 different animal groups while still allowing these proportions to vary across different element portions. These  
1360 observations are interpreted as binomial data: counts of some quantity (e.g., immature specimens) out of a  
1361 total count of relevant specimens (e.g., total ageable specimens); this approach lets the model incorporate the  
1362 uncertainty caused by small sample sizes. The observation of the average proportion of immature specimens

1363  $(\mu_{\pi_1})$  is based on the fusion rate of proximal and middle phalanges (the number of unfused phalanges  
 1364  $N_{\text{Unfused}}$  out of the total number of phalanges with fusion data  $N_{\text{Ageable}}$ ), which fuse at around the same  
 1365 time as the estimated time that animals reach adult body size in the Shetland sheep population (Popkin et  
 1366 al. 2012). The observation of the average adult sex ratio—the proportion of females among mature animals  
 1367  $\left(\frac{\mu_{\pi_2}}{\mu_{\pi_2} + \mu_{\pi_3}}\right)$ —is based on the sex ratio of fused pelvises (the number of female pelvises  $N_{\text{Female}}$  out of the  
 1368 total number of pelvises with a sex assignment  $N_{\text{Sexable}}$ ). In each case, the number of observable specimens  
 1369 determines the measurement error using the binomial distribution. While this paper uses these quantities  
 1370 to estimate the relevant hyper-parameters, relevant observations from other elements can be incorporated  
 1371 into the model in the same fashion if there is a clear sense of the total number of specimens that could have  
 1372 potentially been immature or female.

1373 *Observation Error Equations for Demographic Estimates:*

$$\begin{aligned}
 N_{\text{unfused}} &\sim \text{Binomial}(N_{\text{ageable}}, \mu_{\pi_1}) \\
 N_{\text{female}} &\sim \text{Binomial}(N_{\text{sexable}}, \frac{\mu_{\pi_2}}{\mu_{\pi_2} + \mu_{\pi_3}})
 \end{aligned} \tag{3}$$

1374 ***3. Prior Distributions***

1375 The prior distributions in this model are focused on describing previous beliefs about the value of the mix-  
 1376 ture model hyper-parameters, as the element portion-specific parameters are derived from these distributions.  
 1377 Mixture modeling performs well with unconstrained parameters values because it is more straightforward to  
 1378 estimate variation across element portions, meaning that constrained parameters—those where the range of  
 1379 possible values depends on the values of other parameters—must first be transformed into related uncon-  
 1380 strained parameters (Betancourt 2017) . The following sub-sections describe the necessary transformations  
 1381 for different sets of the mixture model parameters, describing the unconstrained parameters that can be  
 1382 modeled and the transformations that result in the mixture model parameters. While these sections use the  
 1383 mixture model parameter notations, prior distributions are for the ‘central tendency’ hyper-parameter for  
 1384 the described unconstrained parameter.

1385 It is important to remember that for all of these prior distributions are arbitrary choices made by the  
 1386 researcher, regardless of whether the distributions are based on specific animal populations or on reference  
 1387 priors. Other researchers could and should use different prior distributions to best reflect their intuition  
 1388 about likely parameter values for particular case studies. This also highlight the importance of reporting  
 1389 the prior distributions used in a Bayesian analysis to ensure replicability. Examining the implications of  
 1390 different prior distributions is an important step in the development of Bayesian models, one that should be

1391 regularly tested even before models are fit to datasets (Gelman, Vehtari, et al. 2020).

1392 **3.1 Mixture Proportion Priors**

1393 Prior distributions for the mixture proportions reflect our prior beliefs about the relative proportions of the  
1394 three animal groups in the assemblage (immature, adult-sized females, and adult-sized males). The three  
1395 mixture proportions ( $\pi_1, \pi_2, \pi_3$ ) are a three-value unit simplex, meaning that the values are constrained as  
1396 a group to sum up to one. Thus, the simplex can be described by only two variables because the third value  
1397 cannot vary once those two values are known. The model uses two unconstrained variables ( $\theta_1$  and  $\theta_2$ ) to  
1398 describe the  $\pi$  values. These  $\theta$  values are related back to  $\pi$  values using a ‘stick-breaking’ transformation that  
1399 iteratively estimates the relative proportions of the simplex taken up by each  $\theta$  value (Team 2022: Section  
1400 10.7).

1401 *Stick-Breaking Transformations:*

$$\begin{aligned}\pi_1 &= \text{logit}^{-1}(\theta_1 + \log(0.5)) \\ \pi_2 &= (1 - \pi_1) * \text{logit}^{-1}(\theta_2 + \log(1)) \\ \pi_3 &= 1 - (\pi_1 + \pi_2)\end{aligned}\tag{4}$$

1402 The  $\theta_1$  value can be directly related to the  $\pi_1$  value using the first line of the stick-breaking transformation,  
1403 meaning that one can examine the associated  $\pi_1$  estimate for a given  $\theta_1$  value. Within the stick-breaking  
1404 transformation,  $\theta_2$  relates to the relative proportions of  $\pi_2$  and  $\pi_3$  after  $\pi_1$  has been estimated, which is  
1405 effectively the adult sex ratio. Just as we could examine the expected  $\pi_1$  estimates from a distribution of  
1406  $\theta_1$  values, we can thus use expected adult sex ratios ( $\frac{\pi_2}{\pi_2 + \pi_3}$ ) estimates from a particular prior distribution  
1407 for  $\theta_2$ . Relating these  $\theta$  values back to observable phenomena makes it easier to define reasonable prior  
1408 distribution definitions for the parameters from domain expertise (see Section 2.2 and Figure 2 of the main  
1409 text).

1410 **3.2 Average Body Size and Size Variability Priors**

1411 While the average body sizes of the different components ( $\mu_1, \mu_2, \mu_3$ ) are not intrinsically linked in the  
1412 same way that  $\pi$  values are, the model still requires some structure to aid interpretability. Bayesian mixture  
1413 models that are fit using Markov Chain Monte Carlo (MCMC) methods, like the model in this paper, can  
1414 suffer from an issue called “label switching” if  $\mu$  values are not ordered in some way (Jasra, Holmes, and  
1415 Stephens 2005). MCMC methods rely on running multiple “chains”—separate iterations of the model that  
1416 are independently fit and then combined together—to show that the results are independent of the initial

conditions. Label switching describes a scenario where different chains fit the data well, but the parameter labels relate to different specimens (e.g., smaller specimens are assigned to  $\mu_1$  in one chain and to  $\mu_2$  in another). To avoid label switching, the average body sizes are strictly ordered, meaning that  $\mu_1 < \mu_2 < \mu_3$  must be maintained. Note that this only affects average values, individual immature specimens can still be larger than female specimens or male specimens and individual female specimens can be larger than male specimens. This ordering is achieved by only estimating  $\mu_2$  directly (average  $LSI_e$  value for females) and estimating the average  $LSI_e$  value for immature and male animals with offsets ( $\delta_1, \delta_2$ ) from the female average. The  $\delta$  values must be positive to maintain the ordering of the  $\mu$  values, so each  $\delta$  is modeled in a log-transformed space.

Offset Equations for  $\mu$  Values:

$$\begin{aligned}\mu_1 &= \mu_2 - \delta_1 \\ \mu_3 &= \mu_2 + \delta_2\end{aligned}\tag{5}$$

Conceptually, this expression of animal body size defines female animals as the standard “body size” that is subject to various selective pressures, with the offset for male animals reflecting sex-specific pressures on males. This interpretation of body size broadly fits the general pressures affecting adult body size in females and males across many ungulate taxa, including domestic herd animals (Pérez-Barbería, Gordon, and Pagel 2002; Tchernov and Horwitz 1991). The body size offset between immature animals and adult-sized females ( $\delta_1$ ) is admittedly an *ad hoc* definition rather than one under strict biological constraints, as it can be affected by the age immature animals reach before being killed (Gillis et al. 2014). The computational advantages of this definition arguably outweigh the awkwardness of the definition, however. Further, evaluation of  $\delta_1$  and  $\delta_2$  values across different sites could conceivably highlight variation in the timing of the killing of immature animals ( $\delta_1$ ) and the degree of adult sexual dimorphism ( $\delta_2$ ); both variables can be related to models of hunting intensity, animal domestication, and herd management (Gillis et al. 2014; Marom and Bar-Oz 2013; Zeder and Hesse 2000).

The  $LSI_e$  size variability of animals within a group ( $\sigma_1, \sigma_2, \sigma_3$ ) is a key variable in the Bayesian mixture model. The values of these standard deviation parameters play a major role in ensuring that the mixture components reflect biological entities rather than overfitting to specific sample noise. Coefficients of variance (CVs) for raw mammal bone measurements from a single sex have been found to be relatively consistent (Davis 1996; Popkin et al. 2012). When transforming these measurements using a logarithm, this produces consistent standard deviations of the transformed measurement values (Wolfhagen 2020: Figure 1), suggesting that  $\sigma$  values should be relatively stable across elements. While  $\sigma$  parameters from different

1446 animal groups within the model are not directly related to each other, the values still need some transfor-  
1447 mations to be modeled consistently by the multilevel model. These values must be positive, which conflicts  
1448 with the multilevel model's need for unconstrained variables. To achieve this, the model uses the same  
1449 log-transformation technique used for size offsets to create an unconstrained parameter,  $\log_e \sigma$ , that is then  
1450 transformed to actual  $\sigma$  values after estimating variation across element portions.

1451 **3.3 Developing Priors from a Simulated Prior Assemblage**

1452 Relevant prior estimates of the biologically-relevant parameters were derived from the biometry of 356 known-  
1453 age and known-sex Shetland sheep described in Popkin et al. (2012) (see Section 2.2). Castrated individuals  
1454 were included as males and animals killed under one year of age were considered immature; this resulted in an  
1455 assemblage of 48 immature animals, 164 females, and 144 males/castrates. The element portion definitions  
1456 and included measurements are shown Table 4 of the main text. From the 2848 element portions in the full  
1457 assemblage, 150 immature, female, and male element portions were randomly selected to create an assemblage  
1458 of 450 element portions for analysis. LSI<sub>e</sub> values are calculated using the *Ovis orientalis* female standard  
1459 animal (FMC 57951) from Uerpmann and Uerpmann (1994: Table 12). These specimens were modeled using  
1460 a Bayesian multilevel mixture model that used their known identities to estimate the biologically relevant  
1461 parameters directly ( $\mu_2, \delta_1, \delta_2, \sigma_1, \sigma_2, \sigma_3$ ). The resulting hyper-parameters, which average across anatomical  
1462 variation in the parameter values, are used as a baseline for defining prior distributions of the parameter in  
1463 the archaeological model (see Figure 3 in the main text).

1464 The prior distributions used in this Bayesian multilevel mixture model on the reference population are  
1465 more straightforward. While the same transformations to create unconstrained parameters are necessary  
1466 (e.g., modeling average size as  $\mu_2$  with offsets for immature and male animals), the definition of these prior  
1467 distributions can be broadly described as weakly-informative priors (Gelman et al. 2008). These weakly-  
1468 informative prior distributions are reasonable in this case—and not in the archaeological case—because all  
1469 the mixture model parameters have direct observations rather than relying on latent state estimations. That  
1470 is, parameters like the size difference between males and females ( $\delta_2$ ) and the size variability in female animals  
1471 ( $\sigma_2$ ) can be directly observed because the group identities of every specimen are known. With these direct  
1472 observations, the prior distributions have a more muted influence on the resulting posterior distributions.  
1473 This does not mean that the prior distributions have no effect, however, which is why objective priors can  
1474 have undesirable impacts on modeling results (Gabry et al. 2019).

$$\begin{aligned}
& \mu_2 \sim \text{Normal}(-0.1, 0.1) \\
& \log \delta_1 \sim \text{Normal}(-3, 0.5) \\
& \log \delta_2 \sim \text{Normal}(-3, 0.5) \\
& \log \sigma_1 \sim \text{Normal}(-3, 0.1) \\
& \log \sigma_2 \sim \text{Normal}(-3, 0.1) \\
& \log \sigma_3 \sim \text{Normal}(-3, 0.1) \\
& \sigma_{\text{Element}}[1, 2, 3, 4, 5, 6] \sim \text{Half-Normal}(0, 0.05)
\end{aligned} \tag{6}$$

1476 The average LSI<sub>e</sub> value for females ( $\mu_2$ ) is likely to vary across contexts in reaction to different selective  
 1477 pressures, both anthropogenic and ecological (e.g., Davis 1981; Manning et al. 2015; Wright and Viner-  
 1478 Daniels 2015). While the posterior distribution is extremely focused on this specific population, there is no  
 1479 reason to think that this value should be centered at any particular value since that relates to the standard  
 1480 animal used (Meadow 1999; Wolfhagen 2020). Therefore, the prior distribution used in archaeological models  
 1481 for  $\mu_2$  uses a larger standard deviation,  $\mu_2 \sim \text{Normal}(0, 0.1)$ , to encompass likely LSI<sub>e</sub> values (Figure 3A  
 1482 in the main text). Under this definition, there is a 95% probability that the  $\mu_2$  value lies within the range  
 1483 of -0.20 and 0.20 on the LSI<sub>e</sub> scale, which translates to roughly 82-122% the size of the standard animal's  
 1484 measurement.

1485 For the average size difference between immature and female animals ( $\delta_1$ ), the narrowness of the posterior  
 1486 distribution likely reflects the fact that immature animals in the sample cover a narrow age range. Animals  
 1487 killed under one year of age span only 36 days and the youngest animals are nearly half a year old (178-214  
 1488 days: Popkin et al. 2012). Thus while the posterior results provide a useful starting point for estimating  
 1489 this offset, there is a good potential for larger  $\delta_1$  values (i.e., greater size differences between immature and  
 1490 adult-sized female animals) in other contexts that could include animals killed at a younger age (Figure 3B  
 1491 in the main text). To capture this possibility, the archaeological model uses a prior distribution with a larger  
 1492 standard deviation and a slightly higher center,  $\log \delta_1 \sim \text{Normal}(-3.5, 0.4)$ , which results in an average size  
 1493 difference of 0.03 on the LSI<sub>e</sub> scale and a 95% probability that the size difference is between 0.01 and 0.07.  
 1494 This translates into expecting the average body size of immature animals in an assemblage being 3% smaller  
 1495 than the average body size of adult-sized female animals, but also plausibly believing that this size difference  
 1496 could range from 1-7% smaller.

1497 The average size difference between adult males and females ( $\delta_2$ ), also known as the index of sexual

1498 dimorphism (Fernández and Monchot 2007), is likely to be under stricter biological control than the other  
1499 “average body size” parameters in the model. This does not mean that this difference could not vary between  
1500 contexts, however. Some models of animal domestication argue that initial domestication removed sexual  
1501 selective pressures on male body size, reducing sexual dimorphism (e.g., Tchernov and Horwitz 1991). In  
1502 a similar fashion, specialized hunting strategies could also reduce sexual dimorphism by targeting large-  
1503 bodied males, for example (Zeder 2012; Proaktor, Coulson, and Milner-Gulland 2007; Milner, Nilsen, and  
1504 Andreassen 2007). Again, the posterior distribution of the extent of sexual dimorphism in the Shetland  
1505 sheep population provides a useful starting point to describe a prior distribution for the model (Figure 3C  
1506 in the main text). Increasing the standard deviation of the distribution slightly,  $\log \delta_2 \sim \text{Normal}(-2.7, 0.1)$ ,  
1507 produces a distribution centered at 0.07 LSI<sub>e</sub> units with a 95% probability that the value is between 0.06-0.08,  
1508 translating to the average male being 6-9% larger than the average female relative to a standard measurement.  
1509 The smaller standard deviation in the prior distribution of  $\delta_2$  than for  $\delta_1$  reflects our understanding that  
1510 the extent of sexual dimorphism, as a biological phenomenon, is less likely to have extreme values than  
1511 the average size difference between immature and female animals, since  $\delta_2$  is unaffected by the specific age  
1512 structure of the assemblage.

1513 As in the average body size parameters, prior distributions for the size variability model parameters  
1514 are developed from the Bayesian model of known-identity Shetland sheep measurements. The resulting  $\sigma$   
1515 hyper-parameters provide a baseline for establishing hyper-parameter prior distributions in archaeological  
1516 cases. Figure 3D-F of the main text shows the posterior distributions of these  $\sigma$  hyper-parameters in both  
1517 the log-transformed values and associated LSI<sub>e</sub> values. Average size variability within an element portion for  
1518 immature animals ( $\sigma_1$ ) is higher, on average, than for females ( $\sigma_2$ ) and males ( $\sigma_3$ ). The immature category  
1519 includes both male and female animals, so larger size variability makes sense; again, it is possible that  $\sigma_1$  is  
1520 relatively low in this population relative to other contexts given the narrow age range of immature animals  
1521 in the Shetland sheep population. Unlike the average body size parameters, there are not compelling reasons  
1522 to believe that size variability parameters for females and males ( $\sigma_2$  and  $\sigma_3$ ) should vary widely in different  
1523 contexts given the consistency of coefficients of variation in mammals broadly (Davis 1996). Thus, the results  
1524 of this analysis are used for the prior distributions of  $\log \sigma_2$  and  $\log \sigma_3$ , while the prior distribution of  $\log \sigma_1$   
1525 is given an increased standard deviation and slightly increased average value. Overall, however, these prior  
1526 distributions suggest that the average size variability within an element portion is between 0.04-0.06 for  
1527 females and males and is between 0.04-0.05 for immature animals. Note that even though  $\sigma_2$  and  $\sigma_3$  have  
1528 the same prior distributions, these values can still vary from each other in different contexts.

1529 **4. Multilevel Structure of the Model**

1530 The previous section described prior distributions that describe the *average* value for different mixture model  
 1531 parameters across all element portions. To create parameter estimates that are specific to different element  
 1532 portions, it is necessary to estimate the *variation* around these average values that different parameters  
 1533 can have among different element portions. The model uses a Multivariate Normal definition of the model  
 1534 parameters to allow for correlations between different parameters; effectively, the possibility that multiple  
 1535 model parameters will covary from element portion to element portion. To do this, each hyper-parameter  
 1536 has an associated  $\sigma_{\text{element}}$  parameter that describes inter-element variation in parameter values. The model  
 1537 uses a non-centered parameterization, wherein the Multivariate Normal distribution is centered at zero to  
 1538 calculate offsets,  $\nu_{\text{element}}$ , that are added to the average hyper-parameters to calculate model parameters for  
 1539 each element portion. This definition provides computational stability and makes it more straightforward  
 1540 to incorporate other levels of multilevel structure.

1541 *Equations for Defining Inter-Element Variation (Multilevel Modeling):*

$$\nu_{\text{Element}} \sim \text{MultivariateNormal} \left( \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}, \Sigma_{\text{Element}} \right)$$

$$\Sigma_{\text{Element}} = \begin{pmatrix} \sigma_{\text{Element}}[1] & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_{\text{Element}}[8] \end{pmatrix} \rho_{\text{Element}} \begin{pmatrix} \sigma_{\text{Element}}[1] & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_{\text{Element}}[8] \end{pmatrix}$$

$$\rho_{\text{Element}} = LKJcorr(2)$$

$$\theta_1[\text{Element}] = \theta_1 + \nu_{\text{Element}}[1] \quad (7)$$

$$\theta_2[\text{Element}] = \theta_2 + \nu_{\text{Element}}[2]$$

$$\mu_2[\text{Element}] = \mu_2 + \nu_{\text{Element}}[3]$$

$$\log \delta_1[\text{Element}] = \log \delta_1 + \nu_{\text{Element}}[4]$$

$$\log \delta_2[\text{Element}] = \log \delta_2 + \nu_{\text{Element}}[5]$$

$$\log \sigma_1[\text{Element}] = \log \sigma_1 + \nu_{\text{Element}}[6]$$

$$\log \sigma_2[\text{Element}] = \log \sigma_2 + \nu_{\text{Element}}[7]$$

$$\log \sigma_3[\text{Element}] = \log \sigma_3 + \nu_{\text{Element}}[8]$$

1542 The multilevel structure used to allow variation in parameter estimates across element portions can

1543 also be expanded to create multisite models that can directly compare sex-specific biometric estimates  
 1544 alongside the age/sex composition of different assemblages. Such comparisons can highlight variation in  
 1545 herd management strategies or diachronic body size trends related to population turnover (e.g., Arbuckle  
 1546 and Atici 2013; Arbuckle et al. 2016). To do this, an additional multilevel structure can be applied to the  
 1547 same mixture model parameters, using  $\sigma_{Site}$  rather than  $\sigma_{Element}$  parameters. However, an additional set of  
 1548 multilevel structure parameters,  $\sigma_{Interaction}$ , are also necessary to ensure that elemental variation is different  
 1549 at different sites (e.g., the difference between  $\mu_2$  for the distal humerus and  $\mu_2$  for the distal radius is not  
 1550 necessarily the same at different sites). Again, weakly-informative priors are appropriate for both sets of  
 1551 parameters. Each additional term is included in the sum to create specific mixture model parameter values.  
 1552

*Example of Parameter Definition for Inter-Site and Inter-Element Variation:*

$$\theta_1[Site, Element] = \theta_1 + \nu_{Site}[Site] + \nu_{Element}[Element] + \nu_{Interaction}[Site, Element]$$

1553 The inclusion of multiple sites changes the definition of the ‘grand mean’ variable ( $\theta_1$  in the example  
 1554 equation) from a site-level estimate to an overall mean across the sites and elements. These parameter esti-  
 1555 mates thus describe the average composition of the entire set of assemblages. The details of the assemblages  
 1556 included in the analyses would affect how useful these estimates are for interpretation. Assemblage-specific  
 1557 estimates can be calculated for each model parameter by adding the relevant  $\nu_{site}$  estimate to the ‘grand  
 1558 mean’ parameter, which would again act to describe the average composition of the assemblage regardless  
 1559 of its elemental composition.

1560 Prior distributions for  $\sigma_{Element}$  values (and  $\sigma_{Site}$  and  $\sigma_{Interaction}$  values in multisite models) are weakly-  
 1561 informative priors based on the scale of the parameter and the expectation for variation for the parameter.  
 1562 For example, there is likely more variation in  $\theta$  parameters—that govern the relative composition of imma-  
 1563 ture, female, and male animals—among element portions than variation in  $\sigma$  parameters that govern size  
 1564 variability within each group. Similarly, it is expected that average body sizes of females  $\mu_2$  will vary more  
 1565 between sites  $\sigma_{Site}[3]$  than between elements within a site  $\sigma_{Element}[3]$ . The impacts of these prior distri-  
 1566 bution definitions were evaluated using prior predictive checking, suggesting that these prior distributions  
 1567 allow enough variability to encompass reasonable size estimates without providing too much prior support  
 1568 to implausible or impossible values (see Section 6 of the Model Supplement).

*Prior Distributions for Element-level Variation (Multilevel Component):*

$$\begin{aligned} \sigma_{Element}[1, 2] &\sim \text{Half-Normal}(0, 0.5) \\ \sigma_{Element}[3, 4, 5, 6, 7, 8] &\sim \text{Half-Normal}(0, 0.05) \end{aligned} \tag{8}$$

1570      *Prior Distributions for Site-level Variation (Multilevel Component):*

$$\begin{aligned}\sigma_{\text{Site}}[1, 2] &\sim \text{Half-Normal}(0, 0.5) \\ \sigma_{\text{Site}}[3, 4, 5] &\sim \text{Half-Normal}(0, 0.1) \\ \sigma_{\text{Site}}[6, 7, 8] &\sim \text{Half-Normal}(0, 0.05) \\ \rho_{\text{Site}} &= LKJcorr(2)\end{aligned}\tag{9}$$

1571      *Prior Distributions for Interaction Effect (Multilevel Component):*

$$\begin{aligned}\sigma_{\text{Interaction}}[1, 2] &\sim \text{Half-Normal}(0, 0.25) \\ \sigma_{\text{Interaction}}[3, 4, 5] &\sim \text{Half-Normal}(0, 0.1) \\ \sigma_{\text{Interaction}}[6, 7, 8] &\sim \text{Half-Normal}(0, 0.05) \\ \rho_{\text{Interaction}} &= LKJcorr(2)\end{aligned}\tag{10}$$

1572    **5. Prior Distributions for the Model Hyper-Parameters (Simulations and  
1573    Archaeological Cases)**

1574    *Prior Distribution Definitions for the Single Assemblage Simulation Model Hyper-Parameters*

$$\begin{aligned}\theta_1 &\sim \text{Normal}(-0.5, 1.5) \\ \theta_2 &\sim \text{Normal}(0.0, 1.5) \\ \mu_2 &\sim \text{Normal}(0.0, 0.1) \\ \log \delta_1 &\sim \text{Normal}(-3.5, 0.4) \\ \log \delta_2 &\sim \text{Normal}(-2.7, 0.1) \\ \log \sigma_1 &\sim \text{Normal}(-3.05, 0.1) \\ \log \sigma_2 &\sim \text{Normal}(-3.10, 0.1) \\ \log \sigma_3 &\sim \text{Normal}(-3.10, 0.1)\end{aligned}\tag{11}$$

*Prior Distribution Definitions for the Multisite Simulation Model Hyper-Parameters*

$$\begin{aligned}
\theta_1 &\sim \text{Normal}(-0.5, 1.5) \\
\theta_2 &\sim \text{Normal}(0.0, 1.5) \\
\mu_2 &\sim \text{Normal}(0.0, 0.2) \\
\log \delta_1 &\sim \text{Normal}(-3.5, 0.5) \\
\log \delta_2 &\sim \text{Normal}(-2.7, 0.2) \\
\log \sigma_1 &\sim \text{Normal}(-3.05, 0.1) \\
\log \sigma_2 &\sim \text{Normal}(-3.10, 0.1) \\
\log \sigma_3 &\sim \text{Normal}(-3.10, 0.1)
\end{aligned} \tag{12}$$

*Prior Distribution Definitions for the Pinarbaşı B Sheep Model Hyper-Parameters*

$$\begin{aligned}
\theta_1 &\sim \text{Normal}(-0.5, 1.5) \\
\theta_2 &\sim \text{Normal}(0.0, 1.5) \\
\mu_2 &\sim \text{Normal}(0.0, 0.1) \\
\log \delta_1 &\sim \text{Normal}(-3.5, 0.4) \\
\log \delta_2 &\sim \text{Normal}(-2.7, 0.1) \\
\log \sigma_1 &\sim \text{Normal}(-3.05, 0.1) \\
\log \sigma_2 &\sim \text{Normal}(-3.10, 0.1) \\
\log \sigma_3 &\sim \text{Normal}(-3.10, 0.1)
\end{aligned} \tag{13}$$

$$\begin{aligned}
\theta_1 &\sim \text{Normal}(-0.5, 1.5) \\
\theta_2 &\sim \text{Normal}(0.0, 1.5) \\
\mu_2 &\sim \text{Normal}(-0.1, 0.1) \\
\log \delta_1 &\sim \text{Normal}(-3.5, 0.5) \\
\log \delta_2 &\sim \text{Normal}(-2.0, 0.5) \\
\log \sigma_1 &\sim \text{Normal}(-3.05, 0.1) \\
\log \sigma_2 &\sim \text{Normal}(-3.10, 0.1) \\
\log \sigma_3 &\sim \text{Normal}(-3.10, 0.1)
\end{aligned} \tag{14}$$

1578 **6. Simulating Assemblages from the Prior Distributions (Prior Predictive  
1579 Checking)**

1580 Prior predictive checks are a critical component of Bayesian model development workflows, ensuring that  
1581 reasonable prior definitions are chosen (Gabry et al. 2019; Gelman, Vehtari, et al. 2020). This process  
1582 uses the model's prior distribution definitions to simulate data, which can then be evaluated against domain  
1583 knowledge and observed data. This is particularly important when dealing with model parameters that are  
1584 difficult to examine in isolation, like parameters that govern inter-element variation ( $\sigma_{\text{Element}}$ ) in a multilevel  
1585 model structure. Prior predictive checking is an iterative process, informing researchers about the potential  
1586 consequences of their prior distribution definitions; in particular, it can highlight how excessively imprecise  
1587 definitions can provide considerable prior weight on implausible and even impossible values for data (Gabry  
1588 et al. 2019: Figure 4). Thus, prior predictive checking allows researchers to create more accurate summaries  
1589 of their domain knowledge but also produces more efficient MCMC performance because less time is spent  
1590 evaluating parameter values that are inconsistent with even cursory prior knowledge about the problem being  
1591 modeled.

1592 Prior predictive checks were developed for both the single-assemblage and multisite model fits, using the  
1593 prior distribution definitions used in the sheep simulations in the main text (see Section 5 of the Model  
1594 Supplement). In each simulation, 25  $\text{LSI}_{\text{Specimen}}$  values were calculated for each of the 5 element portions  
1595 based on the relevant prior distributions and model structures; for the multisite simulation, this was done  
1596 for 3 assemblages. To evaluate the feasibility of these prior distribution definitions, the  $\text{LSI}_{\text{Specimen}}$  were  
1597 converted into simulated measurement values based on the reference values of the *Ovis orientalis* female  
1598 standard animal (FMC 57951) from Uerpmann and Uerpmann (1994: Table 12). Table 1 shows how the

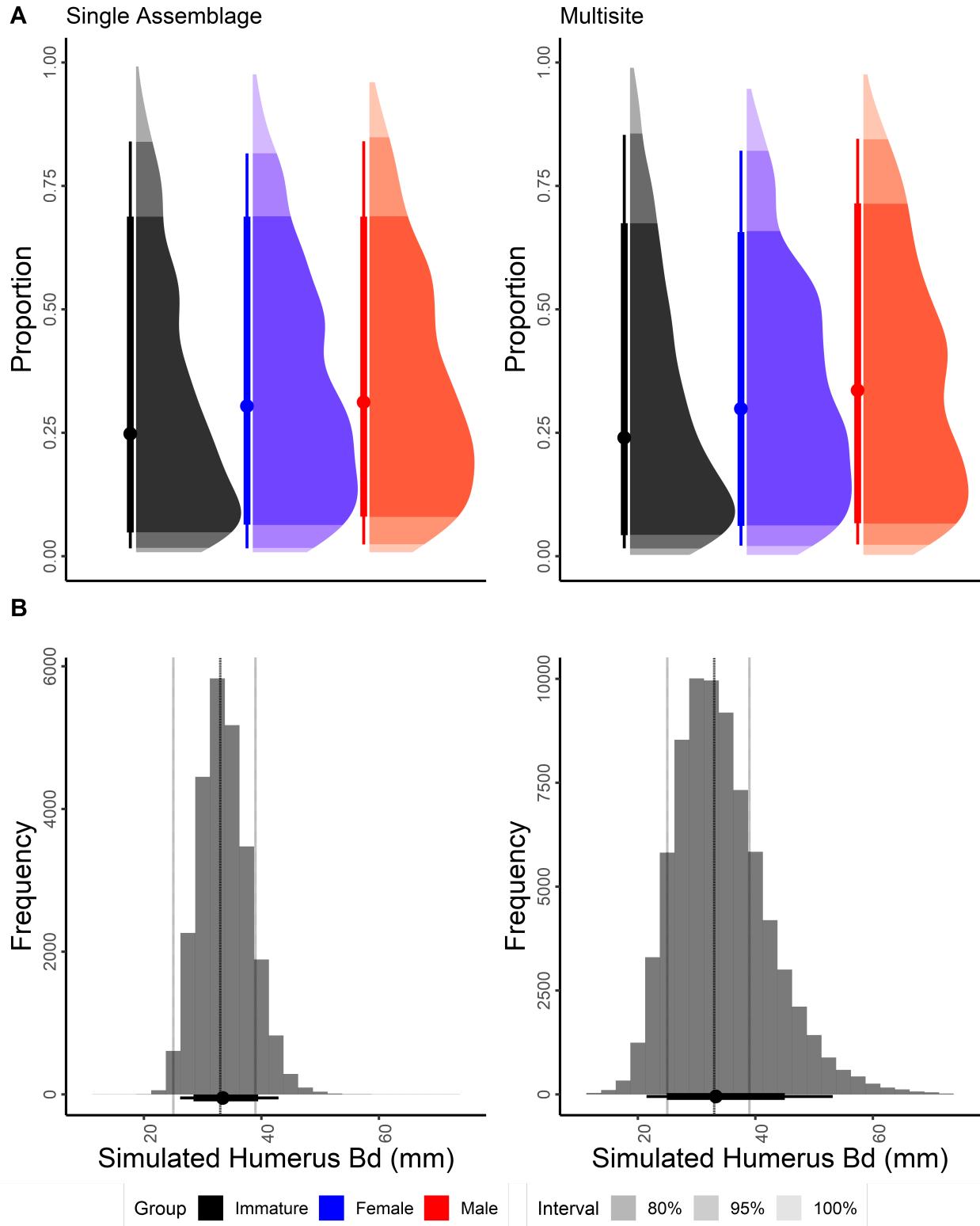
**Table 16:** Values used to convert simulated LSI values into measurements in the prior predictive simulations. Dimension definitions follow von den Driesch (1976). Reference value refers to the female standard mouflon FMC 57951 (Uerpmann and Uerpmann 1994: Table 12).

Element	Portion	Measured Dimension	Reference Value (mm)
Humerus	Bd		33.0
Humerus	BT		29.5
Radius	Bp		33.5
Metacarpus	Bp		25.0
Metatarsus	Bp		22.5
Astragalus	Bd		19.6

1599 five element portions were converted into measurements from specific dimensions, following the equations in  
 1600 Section 2 of this model supplement. Each simulation was run 1000 times, producing 1000 assemblages of  
 1601 relevant simulated measurement values.

1602 Figure 1 shows two results for the single assemblage (left) and multisite (right) prior prediction sim-  
 1603 ulations. The top row shows the simulated proportions of immature, female, and male specimens in the  
 1604 simulated assemblages (Figure 1A). These proportions were sampled directly from the element-specific mix-  
 1605 ture proportion variables, showing the range of potential distributions the model is expecting before seeing  
 1606 any data. The 95% quantiles of the single assemblage model's proportion of immature specimens ranges  
 1607 from 2-84%, for female specimens, and for male specimens. It is noteworthy that the simulated proportion  
 1608 of immature specimens for the single assemblage model does not reach the extreme value seen in the Pinar-  
 1609 bası B sheep data, which may explain some of the long tails in the element-specific compositional estimates  
 1610 (Figure 9 of the main text). The multisite model has similar expected ranges (immature: 2-85%, female: ,  
 1611 male: ).

1612 The bottom row shows the distribution of simulated humerus Bd measurements for the assemblages  
 1613 (Figure 1B). These plots include all simulated specimens, showing the range of sheep humerus Bd values the  
 1614 model expects before seeing any data. Several vertical lines on the plots give a sense of domain knowledge  
 1615 about sheep humerus Bd values. First, the standard value (33.0 mm) from the *Ovis orientalis* female  
 1616 standard animal (FMC 57951) from Uerpmann and Uerpmann (1994: Table 12) is shown in the red dashed  
 1617 line. Second, vertical blue lines show ranges of observed archaeological sheep humerus Bd values: the largest  
 1618 sheep (technically *Ovis orientalis*) humerus Bd (39.0 mm) from the 10th millennium BP site of Körtik Tepe,  
 1619 in southeastern Anatolia (Arbuckle and Özkaya 2006: Table b) and the smallest sheep (*Ovis aries*) humerus  
 1620 Bd (25.0 mm) from the fifth-sixth century CE site of West Stow, United Kingdom (Crabtree 1990: Table  
 1621 29). These plots highlight how the multisite model's structure allows for much more variation in body



**Figure 22:** Single Assemblage and Multisite Prior Predictive Checks. Top row (A): Estimates of the proportion of immature, female, and male specimens. Bottom row (B): Histograms of simulated humerus Bd measurements from the simulations. Vertical lines show the value of the standard reference value (dashed) and two extreme archaeological samples (dotted).

size, especially large measurements. While the 95% quantiles of the single assemblage model's simulated measurements (26-43 mm) do not exactly encapsulate the range of the observed extreme measurements, the multisite model's simulated measurements go well beyond the range (21-53 mm). While one may not expect this full range of measurements in a single assemblage, the results of the prior predictive checks show that the multisite model could plausibly encapsulate variation in body size across diverse assemblages, though possibly at the cost of being somewhat inefficient (i.e., evaluating parameter values that are somewhat beyond reasonable expectations).

These prior predictive checks show that the Bayesian multilevel mixture model's structure is robust for diverse archaeological applications. The chosen prior distribution definitions for the multilevel variation components of the model encapsulate a reasonable range of expected variation and thus do not require extensive retooling as they are applied to new archaeological situations or taxa. Moreover, the multilevel model appears to have enough variability to model drastic changes in body size, making the models' structure relevant for examining broad spatial and temporal variation in animal biometry. While experts are encouraged to use domain knowledge to update and adjust the models to better fit the questions they ask, these models are widely-applicable tools that are suitable for asking many questions about animal body size and the composition of zooarchaeological assemblages.

## References

- Arbuckle, Benjamin S. 2012. "Animals and Inequality in Chalcolithic Central Anatolia." Journal Article. *Journal of Anthropological Archaeology* 31 (3): 302–13. <https://doi.org/10.1016/j.jaa.2012.01.008>.
- Arbuckle, Benjamin S., and Levent Atici. 2013. "Initial Diversity in Sheep and Goat Management in Neolithic South-Western Asia." Journal Article. *Levant* 45 (2): 219–35.
- Arbuckle, Benjamin S., Sarah Whitcher Kansa, Eric C. Kansa, David C. Orton, Canan Cakirlar, Lionel Gourichon, A. Levent Atici, et al. 2014. "Data Sharing Reveals Complexity in the Westward Spread of Domestic Animals Across Neolithic Turkey." Journal Article. *PLoS One* 9 (6): 1–11. <https://doi.org/10.1371/journal.pone.0099845>.
- Arbuckle, Benjamin S., and Theo M. Kassebaum. 2021. "Management and Domestication of Cattle (*Bos Taurus*) in Neolithic Southwest Asia." Journal Article. *Animal Frontiers* 11 (3): 10–19. <https://doi.org/10.1093/af/vfab015>.
- Arbuckle, Benjamin S., and V. Özkaya. 2006. "Animal Exploitation at Körtik Tepe: An Early Aceramic Neolithic Site in Southeastern Turkey." Journal Article. *Paléorient* 32 (2): 113–36.
- Arbuckle, Benjamin S., Max D. Price, Hitomi Hongo, and Banu Öksüz. 2016. "Documenting the Initial

- 1653      Appearance of Domestic Cattle in the Eastern Fertile Crescent (Northern Iraq and Western Iran).”  
1654      Journal Article. *Journal of Archaeological Science* 72: 1–9. <https://doi.org/10.1016/j.jas.2016.05.008>.
- 1655      Baird, Douglas, Denise Carruthers, Andrew Fairbairn, and Jessica A. Pearson. 2011. “Ritual in the Land-  
1656      scape: Evidence from Pınarbaşı in the Seventh-Millennium Cal BC Konya Plain.” Journal Article.  
1657      *Antiquity* 85: 380–94.
- 1658      Baird, Douglas, Andrew Fairbairn, Emma Jenkins, Louise Martin, Caroline Middleton, Jessica A. Pearson,  
1659      Eleni Asouti, et al. 2018. “Agricultural Origins on the Anatolian Plateau.” Journal Article. *Proc Natl  
1660      Acad Sci U S A*. <https://doi.org/10.1073/pnas.1800163115>.
- 1661      Benaglia, Tatiana, Didier Chauveau, David R. Hunter, and Derek Young. 2009. “mixtools: An R Package  
1662      for Analyzing Finite Mixture Models.” *Journal of Statistical Software* 32 (6): 1–29. <http://www.jstatsoft.org/v32/i06/>.
- 1663      Betancourt, Michael. 2017. “A Conceptual Introduction to Hamiltonian Monte Carlo.” Journal Article.  
1664      *arXiv* 1701.02434 [Preprint]: Available from <https://arxiv.org/abs/1701.02434v1>.
- 1665      Breslawski, Ryan P. 2022. “Minimum Animal Units and the Standardized Count Problem.” Journal Article.  
1666      *Journal of Archaeological Method and Theory*. <https://doi.org/10.1007/s10816-022-09563-9>.
- 1667      Breslawski, Ryan P., and David A. Byers. 2015. “Assessing Measurement Error in Paleozoological Osteo-  
1668      metrics with Bison Remains.” Journal Article. *Journal of Archaeological Science* 53: 235–42. <https://doi.org/10.1016/j.jas.2014.10.001>.
- 1669      Buitenhuis, Hjalke. 2008. “Ilipinar: The Faunal Remains from the Late Neolithic and Early Chalcolithic  
1670      Levels.” Book Section. In *Archaeozoology of the Near East VIII*.
- 1671      ———. 2013. “Ilipinar Zooarchaeology.” Dataset. <http://opencontext.org/projects/D297CD29-50CA-4B2C-4A07-498ADF3AF487>. <https://doi.org/https://doi.org/10.6078/M76H4FB>.
- 1672      Çakırlar, Canan. 2013. “Rethinking Neolithic Subsistence at the Gateway to Europe With New Archaeo-  
1673      zoological Evidence from Istanbul.” Book Section. In *Barely Surviving or More Than Enough?: The  
1674      Environmental Archaeology of Subsistence, Specialisation and Surplus Food Production*, edited by Maaike  
1675      Groot, Daphne Lentjes, and Jørn Zeiler, 59–80. Leiden: Sidestone Press.
- 1676      Carty, Angelo, and B. D. Ripley. 2021. *Boot: Bootstrap r (s-Plus) Functions*.
- 1677      Carruthers, Denise. 2005. “Hunting and Herding in Central Anatolian Prehistory: The Sites at Pinarbasi.”  
1678      Book Section. In *Archaeozoology of the Near East VI*.
- 1679      ———. 2006. “Pınarbaşı 1994: Animal Bones.” Dataset. <https://opencontext.org/projects/1677643a-15d0-1b2c-3a35-37f04c765387>.
- 1680      Corporation, Microsoft, and Steve Weston. 2022. *doParallel: Foreach Parallel Adaptor for the 'Parallel'  
1681      Package*. <https://CRAN.R-project.org/package=doParallel>.

- 1686 Corti, Paulo, and David M. Shackleton. 2002. "Relationship Between Predation-Risk Factors and Sexual  
1687 Segregation in Dall's Sheep (*Ovis Dalli Dalli*).” Journal Article. *Canadian Journal of Zoology* 80 (12):  
1688 2108–17. <https://doi.org/10.1139/z02-207>.
- 1689 Crabtree, Pam J. 1990. *West Stow: Early Anglo-Saxon Animal Husbandry*. Book. East Anglian Archaeology.  
1690 Suffolk: Suffolk County Planning Dept.
- 1691 Crema, Enrico R. 2011. "Modelling Temporal Uncertainty in Archaeological Analysis." Journal Article.  
1692 *Journal of Archaeological Method and Theory* 19 (3): 440–61. <https://doi.org/10.1007/s10816-011-9122-3>.
- 1693 Dahl, Gudrun, and Anders Hjort. 1976. *Having Herds: Pastoral Herd Growth and Household Economy*.  
1694 Book. Stockholm Studies in Social Anthropology. Stockholm: University of Stockholm.
- 1695 Daly, K. G., V. Mattiangeli, A. J. Hare, H. Davoudi, H. Fathi, S. B. Doost, S. Amiri, et al. 2021. "Herded  
1696 and Hunted Goat Genomes from the Dawn of Domestication in the Zagros Mountains." Journal Article.  
1697 *Proc Natl Acad Sci U S A* 118 (25). <https://doi.org/10.1073/pnas.2100901118>.
- 1698 Davis, Simon J. M. 1981. "The Effects of Temperature Change and Domestication on the Body Size of Late  
1699 Pleistocene to Holocene Mammals of Israel." Journal Article. *Paleobiology* 7 (1): 101–14.
- 1700 ———. 1982. "Climatic Change and the Advent of Domestication: The Succession of Ruminant Artiodactyls  
1701 in the Late Pleistocene-Holocene in the Israel Region." Journal Article. *Paléorient* 8 (2): 5–15.
- 1702 ———. 1996. "Measurements of a Group of Adult Female Shetland Sheep Skeletons from a Single Flock:  
1703 A Baseline for Zooarchaeologists." Journal Article. *Journal of Archaeological Science* 23: 593–612.
- 1704 ———. 2000. "The Effect of Castration and Age on the Development of the Shetland Sheep Skeleton and  
1705 a Metric Comparison Between Bones of Males, Females and Castrates." Journal Article. *Journal of  
1706 Archaeological Science* 27 (5): 373–90. <https://doi.org/10.1006/jasc.1999.0452>.
- 1707 Davison, A. C., and D. V. Hinkley. 1997. *Bootstrap Methods and Their Applications*. Cambridge: Cambridge  
1708 University Press. <http://statwww.epfl.ch/davison/BMA/>.
- 1709 De Cupere, Bea, Wim Van Neer, Hervé Monchot, Elina Rijmenants, Mircea Udrescu, and Marc Waelkens.  
1710 2005. "Ancient Breeds of Domestic Fowl (*Gallus Gallus f. Domestica*) Distinguished on the Basis of  
1711 Traditional Observations Combined with Mixture Analysis." Journal Article. *Journal of Archaeological  
1712 Science* 32 (11): 1587–97. <https://doi.org/10.1016/j.jas.2005.04.015>.
- 1713 Degerbøl, Magnus. 1970. "The Urus (*Bos Primigenius Bojanus*) and Neolithic Domesticated Cattle (*Bos  
1714 Taurus Domesticus Linné*) in Denmark." Journal Article. *Det Kongelige Danske Videnskabernes Selskab  
1715 Biologiske Skrifter* 17 (1): 5–177.
- 1716 Dong, Zhuan. 1997. "Mixture Analysis and Its Preliminary Application in Archaeology." Journal Article.  
1717 *Journal of Archaeological Science* 24: 141–61.
- 1718

- 1719 Dowle, Matt, and Arun Srinivasan. 2021. *Data.table: Extension of ‘Data.frame’*. <https://CRAN.R-project.org/package=data.table>.
- 1720
- 1721 Driesch, Angela von den. 1976. *A Guide to the Measurement of Animal Bones from Archaeological Sites*.  
1722 Book. Peabody Museum Bulletins. Cambridge, MA: Harvard University.
- 1723 Evershed, Richard P., Sebastian Payne, Andrew G. Sherratt, Mark S. Copley, Jennifer Coolidge, Duska  
1724 Urem-Kotsu, Kostas Kotsakis, et al. 2008. “Earliest Date for Milk Use in the Near East and Southeastern  
1725 Europe Linked to Cattle Herding.” Journal Article. *Nature* 455 (7212): 528–31. <https://doi.org/10.1038/nature07180>.
- 1726
- 1727 Fernández, Hélène, and Hervé Monchot. 2007. “Sexual Dimorphism in Limb Bones of Ibex (*Capra Ibex*  
1728 l.): Mixture Analysis Applied to Modern and Fossil Data.” Journal Article. *International Journal of  
1729 Osteoarchaeology* 17: 479–91. <https://doi.org/10.1002/oa.876>.
- 1730 Fernée, Christianne L., and Konstantinos P. Trimmis. 2021. “Detecting Variability: A Study on the  
1731 Application of Bayesian Multilevel Modelling to Archaeological Data. Evidence from the Neolithic Adri-  
1732 atic and the Bronze Age Aegean.” Journal Article. *Journal of Archaeological Science* 128. <https://doi.org/10.1016/j.jas.2021.105346>.
- 1733
- 1734 Gabry, Jonah, and Rok Cešnovar. 2022. *Cmdstanr: R Interface to ‘CmdStan’*.
- 1735 Gabry, Jonah, Daniel Simpson, Aki Vehtari, Michael Betancourt, and Andrew Gelman. 2019. “Visualization  
1736 in Bayesian Workflow.” Journal Article. *Journal of the Royal Statistical Society, Series A* 182 (2): 389–  
1737 402.
- 1738 Galik, Alfred. 2013. “Barçın Höyük Zooarchaeology.” Dataset. <http://opencontext.org/projects/74749949-4FD4-4C3E-C830-5AA75703E08E>. <https://doi.org/10.6078/M78G8HM0>.
- 1739
- 1740 Gelman, Andrew. 2006a. “Multilevel (Hierarchical) Modeling: What It Can and Cannot Do.” Journal  
1741 Article. *Technometrics* 48 (3): 432–35. <https://doi.org/10.1198/004017005000000661>.
- 1742 ———. 2006b. “Prior Distributions for Variance Parameters in Hierarchical Models (Comment on Article  
1743 by Browne and Draper).” Journal Article. *Bayesian Analysis* 3: 515–34.
- 1744 Gelman, Andrew, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin. 2020.  
1745 *Bayesian Data Analysis*. Book. Third Edition.
- 1746 Gelman, Andrew, Aleks Jakulin, Maria Grazia Pittau, and Yu-Sung Su. 2008. “A Weakly Informative  
1747 Default Prior Distribution for Logistic and Other Regression Models.” Journal Article. *The Annals of  
1748 Applied Statistics* 2 (4): 1360–83. <https://doi.org/10.1214/08-aos191>.
- 1749 Gelman, Andrew, Aki Vehtari, Daniel Simpson, Charles C. Margossian, Bob Carpenter, Yuling Yao, Lauren  
1750 Kennedy, Jonah Gabry, Paul-Christian Bürkner, and Martin Modrák. 2020. “Bayesian Workflow.”  
1751 Journal Article. *ArXiv*.

- 1752 Gerritsen, Fokke, and Rana Özbal. 2019. “Barcin Höyük, a Seventh Millennium Settlement in the Eastern  
1753 Marmara Region of Turkey.” Journal Article. *Documenta Praehistorica* 46: 58–67. <https://doi.org/10.4312/dp.46.4>.
- 1754 Gillis, R., R. M. Arbogast, J. F. Piningre, K. Debue, and J. D. Vigne. 2015. “Prediction Models for Age-  
1755 at-Death Estimates for Calves, Using Unfused Epiphyses and Diaphyses.” Journal Article. *International  
1756 Journal of Osteoarchaeology* 25 (6): 912–22. <https://doi.org/10.1002/oa.2377>.
- 1757 Gillis, R., I. Carrère, M. Saña Seguí, G. Radi, and J. D. Vigne. 2014. “Neonatal Mortality, Young Calf  
1758 Slaughter and Milk Production During the Early Neolithic of North Western Mediterranean.” Journal  
1759 Article. *International Journal of Osteoarchaeology* 26 (2): 303–13. <https://doi.org/10.1002/oa.2422>.
- 1760 Gourichon, Lionel, and Daniel Helmer. 2008. “Etude de La Faune Neolithique de Mentesé.” Book Section. In  
1761 *Life and Death in a Prehistoric Settlement in Northwest Anatolia. The Ilipinar Excavations, Volume III.*,  
1762 edited by Jacob Roodenberg and Songül Alpaslan Roodenberg, 435–48. Leiden: Nederlands Instituut  
1763 voor he Nabije Oosten.
- 1764 ———. 2013. “Faunal Data from Neolithic Mentesé.” Dataset. <https://opencontext.org/projects/42ead4dbbaed-4a58-9b9b-7ec85266d2a9>. <https://doi.org/https://doi.org/10.6078/M7S46PVN>.
- 1765 Grau-Sologestoa, Idoia, and Umberto Albarella. 2019. “The ‘Long’ Sixteenth Century: A Key Period of  
1766 Animal Husbandry Change in England.” Journal Article. *Archaeological and Anthropological Sciences* 11  
1767 (6): 2781–2803. <https://doi.org/10.1007/s12520-018-0723-6>.
- 1768 Grigson, Caroline. 1982. “Sex and Age Determination of Some Bones and Teeth of Domestic Cattle: A  
1769 Review of the Literature.” Book Section. In *Ageing and Sexing Animal Bones from Archaeological Sites*,  
1770 edited by Bob Wilson, Caroline Grigson, and Sebastian Payne, 7–23. BAR British Series. Oxford: British  
1771 Archaeological Reports.
- 1772 ———. 1989. “Size and Sex: Evidence for the Domestication of Cattle in the Near East.” Book Section.  
1773 In *People and Culture in Change: Preceedings of the 2nd Symposium on Upper Paleolithic, Mesolithic,  
1774 and Mesolithic Populations of Europe and the Mediterranean Basin*, edited by Israel Hershkovitz, 77–109.  
1775 BAR International Series. Oxford: Archaeopress.
- 1776 Hammer, Oyvind. 2013. *PAST PAleontological STatistics Reference Manual*. Book.
- 1777 Helmer, Daniel, Lionel Gourichon, Hervé Monchot, Joris Peters, and Maria Saña Segui. 2005. “Identifying  
1778 Early Domestic Cattle from Pre-Pottery Neolithic Sites on the Middle Euphrates Using Sexual Dimor-  
1779 phism.” Book Section. In *First Steps of Animal Domestication: New Archaeozoological Approaches*,  
1780 edited by Jean-Denis Vigne, Joris Peters, and Daniel Helmer, 86–95. Proceedings of the 9th Conference  
1781 of the International Council of Archaeozoology, Durham, August 2002. Oxford: Oxbow Books.
- 1782 Hill, Matthew E., Matthew G. Hill, and Christopher C. Widga. 2008. “Late Quaternary Bison Diminution  
1783

- 1785      on the Great Plains of North America: Evaluating the Role of Human Hunting Versus Climate Change.”  
1786      Journal Article. *Quaternary Science Reviews* 27 (17-18): 1752–71. <https://doi.org/10.1016/j.quascirev>.  
1787      2008.07.002.
- 1788      Hongo, Hitomi, Jessica A. Pearson, Banu Öksüz, and Gülcin Ilgezdi. 2009. “The Process of Ungulate  
1789      Domestication at Çayönü, Southeastern Turkey: A Multidisciplinary Approach Focusing on Bos Sp.  
1790      And *Cervus Elaphus*.” Journal Article. *Anthropozoologica* 44 (1): 63–78.
- 1791      Jasra, A., C. C. Holmes, and D. A. Stephens. 2005. “Markov Chain Monte Carlo Methods and the Label  
1792      Switching Problem in Bayesian Mixture Modeling.” Journal Article. *Statistical Science* 20 (1): 50–67.  
1793      <https://doi.org/10.1214/088342305000000016>.
- 1794      Kabukcu, Ceren. 2017. “Woodland Vegetation History and Human Impacts in South-Central Anatolia  
1795      16,000–6500 Cal BP: Anthracological Results from Five Prehistoric Sites in the Konya Plain.” Journal  
1796      Article. *Quaternary Science Reviews* 176: 85–100. <https://doi.org/10.1016/j.quascirev.2017.10.001>.
- 1797      Karul, Necmi. 2019. “Early Farmers in Northwestern Anatolia in the Seventh Millennium.” Book Section.  
1798      In *Concluding the Neolithic: The Near East in the Second Half of the Seventh Millennium BC*, edited by  
1799      Arkadiusz Marciniak, 269–86. Atlanta, Georgia: Lockwood Press.
- 1800      Kassambara, Alboukadel. 2020. *Ggpubr: 'Ggplot2' Based Publication Ready Plots*. <https://CRAN.R-project.org/package=ggpubr>.
- 1801      Kay, Matthew. 2022. *ggdist: Visualizations of Distributions and Uncertainty*. <https://doi.org/10.5281/zenodo.3879620>.
- 1802      Koch, Paul L. 1986. “Clinal Geographic Variation in Mammals: Implications for the Study of Chronoclines.”  
1803      Journal Article. *Paleobiology* 12 (3): 269–81.
- 1804      Lebennon, Roxanne, and Natalie D. Munro. 2022. “Body Size Variation in a Modern Collection of Mountain  
1805      Gazelle (*Gazella Gazella*) Skeletons.” Journal Article. *Journal of Archaeological Science: Reports* 41.  
1806      <https://doi.org/10.1016/j.jasrep.2021.103285>.
- 1807      Lyman, R. Lee. 2008. *Quantitative Paleozoology*. Book. Cambridge Manuals in Archaeology. Cambridge:  
1808      Cambridge University Press.
- 1809      Madgwick, Richard, and Jacqui Mulville. 2015. “Reconstructing Depositional Histories Through Bone  
1810      Taphonomy: Extending the Potential of Faunal Data.” Journal Article. *Journal of Archaeological Science*  
1811      53: 255–63. <https://doi.org/10.1016/j.jas.2014.10.015>.
- 1812      Manning, Katie, Adrian Timpson, Stephen Shennan, and Enrico Crema. 2015. “Size Reduction in Early  
1813      European Domestic Cattle Relates to Intensification of Neolithic Herding Strategies.” Journal Article.  
1814      *PLoS One* 10 (12): e0141873. <https://doi.org/10.1371/journal.pone.0141873>.
- 1815      Marin, Jean-Michel, Kerrie Mengerson, and Christian P. Robert. 2005. “Bayesian Modelling and Inference  
1816      99

- 1818       on Mixtures of Distributions.” Journal Article. *Handbook of Statistics* 25 (Bayesian Thinking, Modeling  
1819       and Computation): 459–509.
- 1820       Marom, Nimrod, and Guy Bar-Oz. 2013. “The Prey Pathway: A Regional History of Cattle (*Bos Taurus*)  
1821       and Pig (*Sus Scrofa*) Domestication in the Northern Jordan Valley, Israel.” Journal Article. *PLoS One*  
1822       8 (2): e55958. <https://doi.org/10.1371/journal.pone.0055958.g001>.
- 1823       Martín, Patricia, and Ricardo García-González. 2015. “Identifying Sheep (*Ovis Aries*) Fetal Remains in  
1824       Archaeological Contexts.” Journal Article. *Journal of Archaeological Science* 64: 77–87. <https://doi.org/10.1016/j.jas.2015.10.003>.
- 1826       Martín, Patricia, Ricardo García-González, Jordi Nadal, and Josep Maria Vergès. 2015. “Perinatal Ovi-  
1827       caprine Remains and Evidence of Shepherding Activities in Early Holocene Enclosure Caves: El Mirador  
1828       (Sierra de Atapuerca, Spain).” Journal Article. *Quaternary International*. <https://doi.org/10.1016/j.quaint.2015.08.024>.
- 1830       Marwick, Ben. 2017. “Computational Reproducibility in Archaeological Research: Basic Principles and a  
1831       Case Study of Their Implementation.” Journal Article. *Journal of Archaeological Method and Theory* 24  
1832       (2): 424–50. <https://doi.org/10.1007/s10816-015-9272-9>.
- 1833       Marwick, Ben, and Suzanne E. Pilaar Birch. 2018. “A Standard for the Scholarly Citation of Archaeological  
1834       Data as an Incentive to Data Sharing.” Journal Article. *Advances in Archaeological Practice* 6 (02):  
1835       125–43. <https://doi.org/10.1017/aap.2018.3>.
- 1836       McCarthy, Michael A., and P. I. P. Masters. 2005. “Profiting from Prior Information in Bayesian Analyses  
1837       of Ecological Data.” Journal Article. *Journal of Applied Ecology* 42 (6): 1012–19. <https://doi.org/10.1111/j.1365-2664.2005.01101.x>.
- 1839       McElreath, Richard. 2020. *Statistical Rethinking: A Bayesian Course with Examples in r and Stan*. Book.  
1840       Texts in Statistical Science. Boca Raton, FL: CRC Press.
- 1841       McGrory, S., E. M. Svensson, A. Götherström, J. Mulville, A. J. Powell, M. J. Collins, and T. P. O’Connor.  
1842       2012. “A Novel Method for Integrated Age and Sex Determination from Archaeological Cattle Mandibles.”  
1843       Journal Article. *Journal of Archaeological Science* 39 (10): 3324–30. <https://doi.org/10.1016/j.jas.2012.05.021>.
- 1845       Meadow, Richard. 1999. “The Use of Size Index Scaling Techniques for Research on Archaeozoological  
1846       Collections from the Middle East.” Book Section. In *Historia Animalium Ex Ossibus: Beiträge Zur  
1847       Paläoanatomie, Archäologie, Ägyptologie, Ethnologie Und Geschichte Der Tiermedizin*, edited by Cornelia  
1848       Becker, Henriette Manhart, Joris Peters, and Jörg Schibler, 285–300. Leidorf: Verlag Marie.
- 1849       Meier, Jacqueline. 2020. “The Contextual Taphonomy of Middens at Neolithic Kfar HaHoresh.” Journal  
1850       Article. *Journal of Archaeological Science: Reports* 33. <https://doi.org/10.1016/j.jasrep.2020.102531>.

- 1851 Middleton, Caroline. 2018. "The Beginning of Herding and Animal Management: The Early Development  
1852 of Caprine Herding on the Konya Plain, Central Anatolia." Journal Article. *Anatolian Studies* 68: 1–31.  
1853 <https://doi.org/10.1017/s0066154618000017>.
- 1854 Milner, J. M., E. B. Nilsen, and H. P. Andreassen. 2007. "Demographic Side Effects of Selective Hunting in  
1855 Ungulates and Carnivores." Journal Article. *Conserv Biol* 21 (1): 36–47. [https://doi.org/10.1111/j.1523-1739.2006.00591.x](https://doi.org/10.1111/j.1523-<br/>1856 1739.2006.00591.x).
- 1857 Monchot, Hervé, and Jacques Léchelle. 2002. "Statistical Nonparametric Methods for the Study of Fossil  
1858 Populations." Journal Article. *Paleobiology* 28 (1): 55–69. [https://doi.org/http://dx.doi.org/10.1666/0094-8373\(2002\)028%3C0055:SNMFTS%3E2.0.CO;2](https://doi.org/http://dx.doi.org/10.1666/0094-8373(2002)028%3C0055:SNMFTS%3E2.0.CO;2).
- 1860 Monchot, Hervé, Marjan Mashkour, and Jean-Denis Vigne. 2005. "Kernel Smoothing and Mixture Analyses  
1861 for the Determination of the Sex Ratios at Death, at the Beginning of the Domestication of Ungulates." Book Section. In *First Steps of Animal Domestication: New Archaeozoological Approaches*, edited by  
1862 Jean-Denis Vigne, Joris Peters, and Daniel Helmer, 55–60. Oxford: Oxbow Books.
- 1863 Munro, Natalie D., Guy Bar-Oz, and Austin C. Hill. 2011. "An Exploration of Character Traits and Linear  
1864 Measurements for Sexing Mountain Gazelle (*Gazella Gazella*) Skeletons." Journal Article. *Journal of  
1865 Archaeological Science* 38 (6): 1253–65. <https://doi.org/10.1016/j.jas.2011.01.001>.
- 1866 Munro, Natalie D., Roxanne Lebzon, and Lisa Sapir-Hen. 2022. "Revisiting Late Pleistocene-Early  
1867 Holocene Mountain Gazelle (*Gazella Gazella*) Body Size Change in the Southern Levant: A Case for  
1868 Anthropogenic Impact." Journal Article. *PLoS One* 17 (8): e0273024. <https://doi.org/10.1371/journal.pone.0273024>.
- 1869 Otárola-Castillo, Erik, and Melissa G. Torquato. 2018. "Bayesian Statistics in Archaeology." Journal Article.  
1870 *Annual Review of Anthropology* 47 (1). <https://doi.org/10.1146/annurev-anthro-102317-045834>.
- 1871 Otárola-Castillo, Erik, Melissa G. Torquato, Jesse Wolfhagen, Matthew E. Hill, and Caitlin E. Buck. 2022.  
1872 "Beyond Chronology, Using Bayesian Inference to Evaluate Hypotheses in Archaeology." Journal Article.  
1873 *Advances in Archaeological Practice*.
- 1874 Özdogan, Mehmet. 2011. "Archaeological Evidence on the Westward Expansion of Farming Communities  
1875 from Eastern Anatolia to the Aegean and the Balkans." Journal Article. *Current Anthropology* 52 (S4):  
1876 S415–30. <https://doi.org/10.1086/658895>.
- 1877 ———. 2019. "Early Farmers in Northwestern Turkey: What Is New?" Book Section. In *Concluding  
1878 the Neolithic: The Near East in the Second Half of the Seventh Millennium BC*, edited by Arkadiusz  
1879 Marciniak, 307–28. Atlanta, Georgia: Lockwood Press.
- 1880 Payne, Sebastian. 1969. "A Metrical Distinction Between Sheep and Goat Metacarpals." Book Section. In  
1881 *The Domestication and Exploitation of Plants and Animals*, edited by Peter J. Ucko and Geoffrey W.

- 1884 Dimbleby, 295–305. London: Duckworth.
- 1885 ———. 1973. “Kill-Off Patterns in Sheep and Goats: The Mandibles from Aşvan Kale.” Journal Article.
- 1886 *Anatolian Studies* 23: 281–303.
- 1887 Payne, Sebastian, and Gail Bull. 1988. “Components of Variation in Measurements of Pig Bones and Teeth,  
1888 and the Use of Measurements to Distinguish Wild from Domestic Pigs.” Journal Article. *Archaeozoologia*
- 1889 II (1,2): 27–66.
- 1890 Pebesma, Edzer. 2018. “Simple Features for R: Standardized Support for Spatial Vector Data.” *The R  
1891 Journal* 10 (1): 439–46. <https://doi.org/10.32614/RJ-2018-009>.
- 1892 Pérez-Barbería, F. Javier, Iain J. Gordon, and M. Pagel. 2002. “The Origins of Sexual Dimorphism in Body  
1893 Size in Ungulates.” Journal Article. *Evolution* 56 (6): 1276–85.
- 1894 Peters, Joris, Angela von den Driesch, and Daniel Helmer. 2005. “The Upper Euphrates-Tigris Basin:  
1895 Cradle of Agro-Pastoralism?” Book Section. In *The First Steps of Animal Domestication*, edited by  
1896 Jean-Denis Vigne, Joris Peters, and Daniel Helmer, 96–124. Proceedings of the 9th Conference of the  
1897 International Council of Archaeozoology, Durham, August 2002. Oxford: Oxbow Books.
- 1898 Popkin, Peter R. W., Polydora Baker, Fay Worley, Sebastian Payne, and Andy Hammon. 2012. “The Sheep  
1899 Project (1): Determining Skeletal Growth, Timing of Epiphyseal Fusion and Morphometric Variation  
1900 in Unimproved Shetland Sheep of Known Age, Sex, Castration Status and Nutrition.” Journal Article.  
1901 *Journal of Archaeological Science* 39 (6): 1775–92. <https://doi.org/10.1016/j.jas.2012.01.018>.
- 1902 Post, Diane M., Trent S. Armbrust, Eva A. Horne, and Jacob R. Goheen. 2001. “Sexual Segregation  
1903 Results in Differences in Content and Quality of Bison (*Bos Bison*) Diets.” Journal Article. *Journal of  
1904 Mammalogy* 82 (2): 407–13.
- 1905 Pozo, Jose M, Valenzuela-Lamas, Silvia, Trentacoste, Angela, Nieto-Espinet, Ariadna, Guimarães Chiarelli,  
1906 and Silvia. 2021. “Zoolog: Zooarchaeological Analysis with Log-Ratios.” *Journal of Statistical Software*  
1907 In preparation. <https://josempozo.github.io/zoolog/>.
- 1908 Proaktor, G., T. Coulson, and E. J. Milner-Gulland. 2007. “Evolutionary Responses to Harvesting in  
1909 Ungulates.” Journal Article. *J Anim Ecol* 76 (4): 669–78. <https://doi.org/10.1111/j.1365-2656.2007.01244.x>.
- 1911 R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R  
1912 Foundation for Statistical Computing. <https://www.R-project.org/>.
- 1913 Ram, Karthik, and Ben Marwick. 2018. “Building Toward a Future Where Reproducible, Open Science Is  
1914 the Norm.” Book Section. In *The Practice of Reproducible Research: Case Studies and Lessons from the  
1915 Data-Intensive Sciences*, edited by Justin Kitzes, Daniel Turek, and Fatma Deniz, 69–78. Oakland, CA:  
1916 University of California Press.

- 1917 Redding, Richard W. 1984. "Theoretical Determinants of a Herder's Decisions: Modeling Variation in the  
1918 Sheep/Goat Ratio." Book Section. In *Animals and Archaeology*, edited by Juliet Clutton-Brock and  
1919 Caroline Grigson, 3. Early Herders and their Flocks:223–41. BAR International Series. Oxford: British  
1920 Archaeological Reports.
- 1921 Roodenberg, Jacob. 2012a. "Change in Food Production and Its Impact on an Early 6th Millennium  
1922 Community in Northwest Anatolia. The Example of Ilipinar." Journal Article. *Praehistorische Zeitschrift*  
1923 87 (2). <https://doi.org/10.1515/pz-2012-0015>.
- 1924 ———. 2012b. "Ilipinar: A Neolithic Settlement in the Eastern Marmara Region." Book Section. In  
1925 *The Oxford Handbook of Ancient Anatolia (10,000-323 BCE)*, edited by Gregory McMahon and Sharon  
1926 Steadman. <https://doi.org/10.1093/oxfordhb/9780195376142.013.0044>.
- 1927 Roodenberg, Jacob, A. van As, L. Jacobs, and M. H. Wijnen. 2003. "Early Settlement in the Plain of  
1928 Yenisehir (NW Anatolia): The Basal Occupation Layers at Mentese." Journal Article. *Anatolica* XXIX  
1929 (1): 18–58.
- 1930 Rosenthal, Robert. 1979. "The "File Drawer Problem" and Tolerance for Null Results." Journal Article.  
1931 *Psychological Bulletin* 86 (3): 638–41.
- 1932 RStudio Team. 2022. *RStudio: Integrated Development Environment for r*. Boston, MA: RStudio, PBC.  
1933 <http://www.rstudio.com/>.
- 1934 Ruckstuhl, K. E. 2007. "Sexual Segregation in Vertebrates: Proximate and Ultimate Causes." Journal  
1935 Article. *Integr Comp Biol* 47 (2): 245–57. <https://doi.org/10.1093/icb/icm030>.
- 1936 Ruckstuhl, K. E., and P. Neuhaus. 2002. "Sexual Segregation in Ungulates: A Comparative Test of  
1937 Three Hypotheses." Journal Article. *Biological Reviews* 77 (1): 77–96. <https://doi.org/10.1017/s1464793101005814>.
- 1939 Russell, Nerissa, and Janet L. Griffitts. 2013. "Catalhoyuk Worked Bones: South and 4040 Areas." Book  
1940 Section. In *Substantive Technologies at Catalhoyuk: Reports from the 2000-2008 Seasons*, edited by Ian  
1941 Hodder, 277–306. Los Angeles.
- 1942 Said, Sonia, Vincent Tolon, Serge Brandt, and Eric Baubet. 2011. "Sex Effect on Habitat Selection in  
1943 Response to Hunting Disturbance: The Study of Wild Boar." Journal Article. *European Journal of  
1944 Wildlife Research* 58 (1): 107–15. <https://doi.org/10.1007/s10344-011-0548-4>.
- 1945 Sasson, Aharon, and Susan Arter. 2020. "Earliest Utilization of Chicken in Upper California: The Zooar-  
1946 chaeology of Avian Remains from the San Diego Royal Presidio." Journal Article. *American Antiquity*  
1947 85 (3): 516–34. <https://doi.org/10.1017/aaq.2020.27>.
- 1948 Scrucca, Luca, Michael Fop, T. Brendan Murphy, and Adrian E. Raftery. 2016. "mclust 5: Clustering,  
1949 Classification and Density Estimation Using Gaussian Finite Mixture Models." *The R Journal* 8 (1):

- 1950        289–317. <https://doi.org/10.32614/RJ-2016-021>.
- 1951        Smith, A. F. M., and Alan E. Gelfand. 1992. “Bayesian Statistics Without Tears: A Sampling-Resampling  
1952        Perspective.” Journal Article. *The American Statistician* 46 (2): 84–88.
- 1953        South, Andy. 2017a. *Rnaturalearth: World Map Data from Natural Earth*. <https://CRAN.R-project.org/package=rnaturalearth>.
- 1954        ———. 2017b. *Rnaturalearthdata: World Vector Map Data from Natural Earth Used in 'Rnaturalearth'*.  
1956        <https://CRAN.R-project.org/package=rnaturalearthdata>.
- 1957        Speth, John D. 1983. *Bison Kills and Bone Counts: Decision Making by Ancient Hunters*. Book. Chicago:  
1958        University of Chicago Press.
- 1959        ———. 2013. “Thoughts about Hunting: Some Things We Know and Some Things We Don’t Know.”  
1960        Journal Article. *Quaternary International* 297: 176–85. <https://doi.org/10.1016/j.quaint.2012.12.005>.
- 1961        Stan Development Team. 2021. “RStan: The R Interface to Stan.” <https://mc-stan.org/>.
- 1962        Stiner, Mary C. 1990. “The Use of Mortality Patterns in Archaeological Studies of Hominid Predatory  
1963        Adaptations.” Journal Article. *Journal of Anthropological Archaeology* 9: 305–51.
- 1964        Stiner, Mary C., Natalie D. Munro, Hiljke Buitenhuis, G. Duru, and M. Ozbasaran. 2022. “An Endemic  
1965        Pathway to Sheep and Goat Domestication at Asikli Hoyuk (Central Anatolia, Turkey).” Journal Article.  
1966        *Proc Natl Acad Sci U S A* 119 (4). <https://doi.org/10.1073/pnas.2110930119>.
- 1967        Symmons, Robert. 2005. “New Density Data for Unfused and Fused Sheep Bones, and a Preliminary  
1968        Discussion on the Modelling of Taphonomic Bias in Archaeofaunal Age Profiles.” Journal Article. *Journal  
1969        of Archaeological Science* 32 (11): 1691–98. <https://doi.org/10.1016/j.jas.2005.05.011>.
- 1970        Tchernov, Eitan, and Liora Kolska Horwitz. 1991. “Body Size Diminution Under Domestication: Uncon-  
1971        scious Selection in Primeval Domesticates.” Journal Article. *Journal of Anthropological Archaeology* 10:  
1972        54–75.
- 1973        Team, Stan Development. 2022. “Stan Modeling Language Users Guide and Reference Manual, Version  
1974        2.29.” Generic. <https://mc-stan.org>.
- 1975        Thissen, Laurens, Hadi Özbal, Ayla Türkekul Biyik, Fokke Gerritsen, and Rana Özbal. 2010. “The Land  
1976        of Milk? Approaching Dietary Preferences of Late Neolithic Communities in NW Anatolia.” Journal  
1977        Article. *Leiden Journal of Pottery Studies* 26: 157–72.
- 1978        Todd, Lawrence C. 1983. “The Horner Site: Taphonomy of an Early Holocene Bison Bonebed.” Thesis.
- 1979        Trentacoste, A., A. Nieto-Espinet, S. Guimaraes, B. Wilkens, G. Petrucci, and S. Valenzuela-Lamas. 2021.  
1980        “New Trajectories or Accelerating Change? Zooarchaeological Evidence for Roman Transformation of  
1981        Animal Husbandry in Northern Italy.” Journal Article. *Archaeol Anthropol Sci* 13 (1): 25. <https://doi.org/10.1007/s12520-020-01251-7>.

- 1983 Twiss, Katheryn C. 2019. *The Archaeology of Food*. Book. <https://doi.org/10.1017/9781108670159>.
- 1984 Twiss, Katheryn C., and Nerissa Russell. 2009. “Taking the Bull by the Horns: Ideology, Masculinity, and
- 1985 Cattle Horns at atalhöyük (Turkey).” Journal Article. *Paléorient* 35 (2): 19–32.
- 1986 Uerpmann, Margarethe, and Hans-Peter Uerpmann. 1994. “Animal Bones.” Book Section. In *Qala’at Al-*
- 1987 *Bahrain*, edited by Flemming Hojlund and H. Hellmuth Andersen, 1: The Northern City Wall and the
- 1988 Islamic Fortress:417–54. Jutland Archaeological Society Publications. Aarhus: Aarhus University Press.
- 1989 Vigne, Jean-Denis. 2011. “Le Mouton (Ovis Aries).” Book Section. In *Shillourokambos: Un Établissement*
- 1990 *Néolithique Pré-Céramique à Chypre: Les Fouilles Du Secteur i*, edited by Jean Guilaine, François Briois,
- 1991 and Jean-Denis Vigne, 1021–38. Paris: Editions Errance.
- 1992 Watkins, Trevor. 1996. “Excavations at Pinarbaşı: The Early Stages.” Book Section. In *On the Surface:*
- 1993 *Çatalhöyük 1993-1995*, edited by Ian Hodder, 47–57. Cambridge: McDonald Institute for Archaeological
- 1994 Research.
- 1995 Weinstock, Jacobo. 2006. “Environment, Body Size and Sexual Dimorphism in Late Glacial Reindeer.”
- 1996 Book Section. In *Recent Advances in Ageing and Sexing Animal Bones*, edited by Deborah Ruscillo,
- 1997 247–53. Oxford: Oxbow Books.
- 1998 Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>.
- 1999
- 2000 Widga, Chris. 2006. “Niche Variability in Late Holocene Bison: A Perspective from Big Bone Lick, KY.”
- 2001 Journal Article. *Journal of Archaeological Science* 33 (9): 1237–55. <https://doi.org/10.1016/j.jas.2005.12.011>.
- 2002
- 2003 Wolfhagen, Jesse. 2020. “Re-Examining the Use of the LSI Technique in Zooarchaeology.” Journal Article.
- 2004 *Journal of Archaeological Science* 123: 105254. <https://doi.org/10.1016/j.jas.2020.105254>.
- 2005 Wolfhagen, Jesse, and Max D. Price. 2017. “A Probabilistic Model for Distinguishing Between Sheep and
- 2006 Goat Postcranial Remains.” Journal Article. *Journal of Archaeological Science: Reports* 12: 625–31.
- 2007 <https://doi.org/10.1016/j.jasrep.2017.02.022>.
- 2008 Wolfhagen, Jesse, Katheryn C. Twiss, Jacqui A. Mulville, and G. Arzu Demirergi. 2021. “Examining
- 2009 Caprine Management and Cattle Domestication Through Biometric Analyses at Çatalhöyük East (North
- 2010 and South Areas).” Book Section. In *Peopling the Landscape of Çatalhöyük: Reports from the 2009–2017*
- 2011 *Seasons*, edited by Ian Hodder, 181–98. Los Angeles: Cotsen Institute of Archaeology Press.
- 2012 Wolverton, Steve. 2008. “Harvest Pressure and Environmental Carrying Capacity: An Ordinal-Scale Model
- 2013 of Effects on Ungulate Prey.” Journal Article. *American Antiquity* 73 (2): 179–99.
- 2014 Wright, Elizabeth, and Sarah Viner-Daniels. 2015. “Geographical Variation in the Size and Shape of the
- 2015 European Aurochs (*Bos Primigenius*).” Journal Article. *Journal of Archaeological Science* 54: 8–22.

- <sup>2016</sup> [https://doi.org/10.1016/j.jas.2014.11.021.](https://doi.org/10.1016/j.jas.2014.11.021)
- <sup>2017</sup> Zeder, Melinda A. 2012. "Pathways to Animal Domestication." Book Section. In *Biodiversity in Agriculture: Domestication, Evolution, and Sustainability*, edited by P. Gepts, T. R. Famula, and R. L. Bettinger, 227–59. Cambridge: Cambridge University Press.
- <sup>2018</sup> Zeder, Melinda A., and Brian Hesse. 2000. "The Initial Domestication of Goats (*Capra Hircus*) in the Zagros Mountains 10,000 Years Ago." Journal Article. *Science* 287 (5461): 2254–57.
- <sup>2019</sup> Zeder, Melinda A., and Ximena Lemoine. 2020. "A Method for Constructing Demographic Profiles in Sus Scrofa Using Logarithm Size Index Scaling." Journal Article. *Journal of Archaeological Science* 116.
- <sup>2020</sup> [https://doi.org/10.1016/j.jas.2020.105115.](https://doi.org/10.1016/j.jas.2020.105115)
- <sup>2021</sup> Zohary, D., Eitan Tchernov, and Liora Kolska Horwitz. 1998. "The Role of Unconscious Selection in the Domestication of Sheep and Goats." Journal Article. *Journal of Zoology, London* 245: 129–35.
- <sup>2022</sup>