

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

Jesse Wolfhagen^{1,2}

¹Department of Anthropology, Purdue University, West Lafayette, Indiana, USA

²Department of Archaeology, Max Planck Institute for Geoanthropology, Jena, Germany

email: jl.wolfhagen@gmail.com

R Markdown version last compiled on 2023-05-18

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.*

¹ 1. INTRODUCTION

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

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

²¹ 1.1 *Morphometric Sex Determination in Zooarchaeology*

²² Some biometric methods to determine the sex of an animal bone are multivariate—using combinations of
²³ measurements with bivariate plots or discriminant functions to predict the sex of archaeological specimens
²⁴ based on distributions of known-sex specimens (e.g., Munro, Bar-Oz, and Hill 2011; Speth 1983; Widga 2006).
²⁵ These methods typically combine dimensions from different planes of an element (e.g., the breadth and depth
²⁶ of a distal articular end) to produce patterns that can be separated by a ‘cut point’ between males and females,
²⁷ either visually in the case of bivariate plots or algorithmically in the case of discriminant functions. The
²⁸ analytical requirement that multiple dimensions of a bone be preserved in measurable condition, even on the
²⁹ same end of an element, may make it difficult to apply these methods to more heavily processed assemblages.
³⁰ Further, specimens from animals that died before reaching adult body size may be misclassified as females,
³¹ particularly for dimensions 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
42 that the biometry of the archaeological population is the same as the reference population. However, this
43 implication is an untenable one in most cases, as animal biometry typically varies spatially and temporally due
44 to population-level intra-taxonomic genetic differences caused by adaptation to local climates and ecologies
45 (e.g., Koch 1986; Davis 1982; Wright and Viner-Daniels 2015; Hill, Hill, and Widga 2008; Lebzon and
46 Munro 2022). Biometric variation in wild and domesticated taxa have also been attributed to anthropogenic
47 pressures as a result of herding decisions or hunting pressure (e.g., Arbuckle and Kassebaum 2021; Manning
48 et al. 2015; Trentacoste et al. 2021; Grau-Sologestoa and Albarella 2019), though harvest pressure has also
49 been attributed to biometric changes in wild taxa (e.g., Wolverton 2008; Munro, Lebzon, and Sapir-Hen
50 2022). These environmental and anthropogenic pressures may affect males and females differently (e.g.,
51 Tchernov and Horwitz 1991; Zohary, Tchernov, and Horwitz 1998); pressures that reduce sexual dimorphism
52 could interfere with analyses, as more specimens may be indeterminate or misclassified. Biometric variation
53 between populations can complicate efforts to estimate changes in the demographic (ontogenetic age and sex)
54 composition of assemblages over time. Further, efforts to control for ontogenetic age (e.g., removing unfused
55 specimens or those from early-fusing elements) distorts the relationship between the analyzed specimens and
56 the rest of the assemblage, decreasing our ability to make reliable inferences about the entire assemblage
57 (Zeder and Hesse 2000).

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

65 spatial and temporal scales. What is necessary is a way to estimate the ontogenetic age and sex composition
66 of a faunal assemblage that preserves the uncertainty inherent in the process.

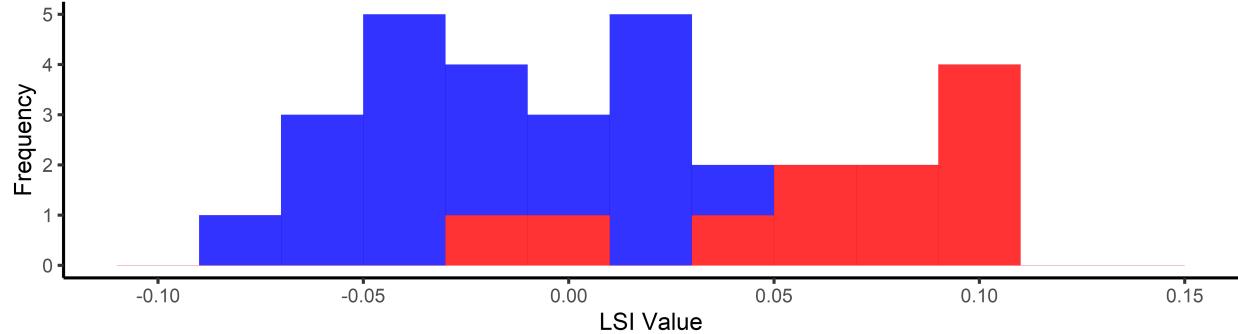
67 **1.2 Mixture Modeling in Zooarchaeology**

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

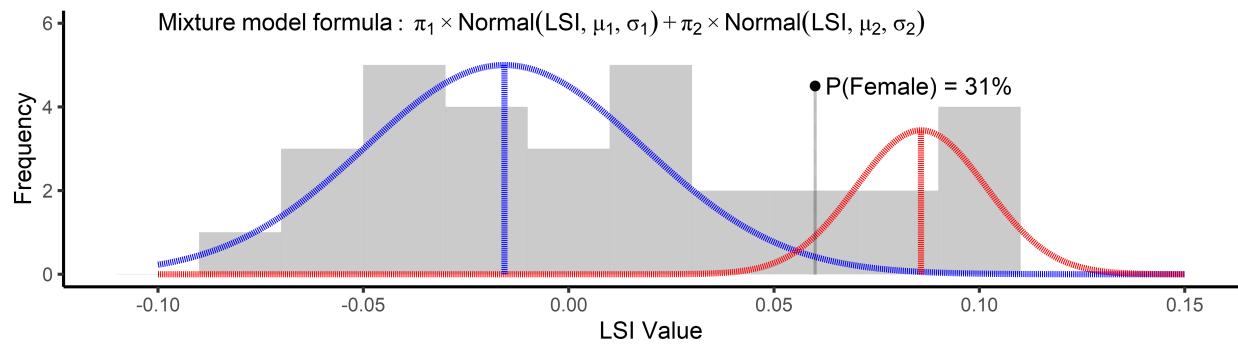
79 Conceptually, a mixture model can be thought of as a “latent state” or “missing data” problem: we know
80 that measured specimens come from particular animal groups, but that information has been lost (Marin,
81 Mengersen, and Robert 2005). If we knew every specimen’s group identity, then the calculation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

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

A Adult Pig Tibia Bd (Known Identities)



B Mixture Model Fit



C Simulated Histograms

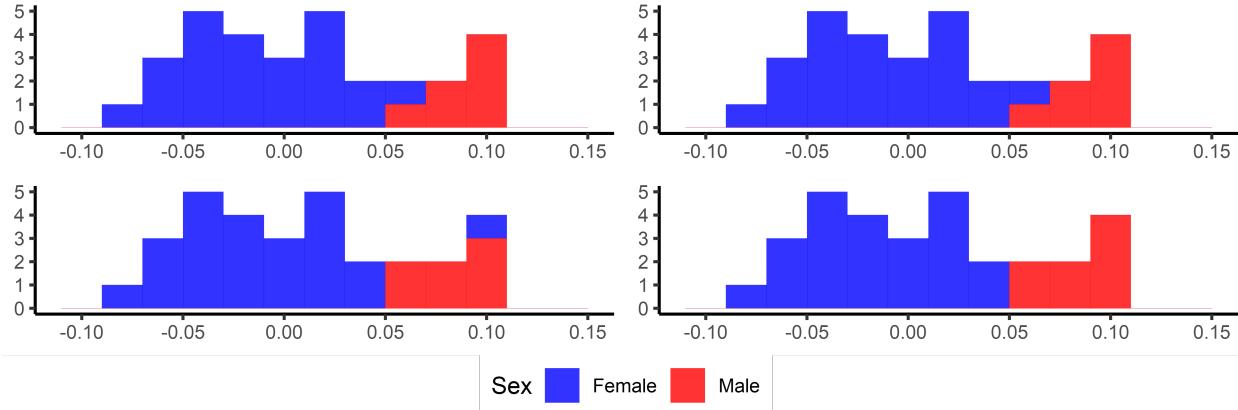


Figure 1: Walkthrough of the mixture modeling procedure using pig Tibia Bd measurements from Zeder and Lemoine (2022). A: the distribution of Tibia Bd LSI values (standard value: 33.5 mm; Hongo and Meadow 2000) of adult specimens, shaded by known sex (females = blue, males = red). B: The result of a 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 probability is used to simulate a ‘true’ identity. Importantly, by leaving the mixture model results as
98 specimen-specific probabilities of being female or male, mixture model results retain the uncertainty of the
99 sex determination process: that is, a specimen with a 51% probability of being female is not treated as
100 equivalent to one with a 95% probability of being female.

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

112 The reasons for the patchy application of mixture modeling in zooarchaeology are less straightforward.
113 High-profile early case studies of mixture modeling report size variability parameters (standard deviation
114 σ) that vary widely and include very small values for some groups (e.g., De Cupere et al. 2005; Monchot,
115 Mashkour, and Vigne 2005; Vigne 2011). These results suggest that the very flexibility that is a great
116 strength of mixture modeling is actually identifying ‘groups’ that are not necessarily consistent with biological
117 expectations (e.g., that the results are “overfitted” to the observed data). Such extreme differences in
118 the standard deviation of different groups can result in counterintuitive implications: specimens may be
119 considered more likely to come from the broad distribution (the one with the larger σ parameter) than the
120 narrow distribution even when the value is more extreme than the narrow distribution’s mean (e.g., is larger
121 than a larger mean or smaller than a smaller mean). Returning to the mixture model example in Figure 1
122 can explain this issue more clearly. Table 1 shows the mixture model parameters for the two components:
123 the standard deviation (σ) for females is more than twice the standard deviation for males ($\sigma_1 = 0.034$, $\sigma_2 =$
124 0.016). As such, higher numbers beyond the observed range will be considered likely females: an LSI_e value
125 of 0.173 (Tibia Bd value: 39.83 mm) is more likely to be a female than a male using the mixture model’s
126 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 π	Average Size μ	Size Variability σ
Female	0.76	-0.016	0.034
Male	0.24	0.086	0.016

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

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

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

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

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

¹⁸³ **2. A BAYESIAN MULTILEVEL MIXTURE MODEL FOR ZOOARCHAEO-**
¹⁸⁴ **LOGICAL MEASUREMENTS**

¹⁸⁵ The Bayesian model developed for this paper improves on standard mixture modeling for zooarchaeological
¹⁸⁶ measurements in four distinct ways. First, it addresses complications caused by measurements from unfused
¹⁸⁷ specimens and post-fusion growth by modeling three groups within the mixture: immature animals, (adult-
¹⁸⁸ sized) females, and (adult-sized) males, each with distinct size parameters. Second, the multilevel structure
¹⁸⁹ allows the model to balance bias due to aggregation and overfitting from small sample sizes. Third, the
¹⁹⁰ Bayesian foundation of the model provides an avenue for synthesizing information about the ontogenetic
¹⁹¹ age and sex composition of the assemblage from non-metrical data (e.g., fusion rates, sex ratios based
¹⁹² on morphological data) to inform the results of the mixture model. Finally, researchers can create prior
¹⁹³ distributions for mixture model parameters from prior assemblages or other sources, ensuring that the
¹⁹⁴ mixture model results are biologically interpretable. This section outlines these benefits; specific details of
¹⁹⁵ the model are described in a Model Supplement and in the analytical code (available at the project's GitHub
¹⁹⁶ page).

¹⁹⁷ Observed measurements from different dimensions (e.g., humerus distal breadth “Humerus Bd,” radius
¹⁹⁸ proximal breadth “Radius Bp,” abbreviations following Driesch 1976) are first converted to logarithmic size
¹⁹⁹ index (LSI) values using a natural logarithm base to take advantage of the normalization LSI standardization
²⁰⁰ provides (Meadow 1999; Wolfhagen 2020). The model uses a single LSI value per specimen, so any specimens
²⁰¹ with multiple observed dimensions are first summarized by estimating the mean LSI value from the observed
²⁰² dimensions; these specimen-level LSI values are the basis for the mixture model. These specimen-level LSI
²⁰³ values can be clustered into different element portions—partial or complete elements that are the basic
²⁰⁴ categorical unit of an analysis [e.g., “distal humerus” or “first phalanx”; compare to “skeletal part type” in
²⁰⁵ Breslawski (2023)]. 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
Element Portion	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 (2023).
Dimension	Specific type of observed measurement (e.g., "humerus distal breadth") on a specimen. Dimension definitions typically follow von den Driesch (1976).
Measured Assemblage	Assemblage of measured specimens from a defined number of element portions of a specific taxon.
Modeled Assemblage	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).
Full Assemblage	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

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

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

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

242 Relevant information from a prior distribution can inform an analyst about reasonable values for the
243 model’s hyper-parameters, which can be summarized as prior distributions. Creating informed prior distri-
244 butions allows the model to ensure that the hyper-parameters have biologically interpretable results (e.g.,
245 average size and size variability parameters for a measurement that align with reasonable expectations for a
246 taxon). The multilevel structure of the model then ensures that mixture model parameters can vary between
247 different element portions while still being informed by these hyper-parameters to maintain biological in-
248 terpretability, even with small numbers of observations. Prior distributions draw explicit links between our

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

2.2 Developing Prior Distributions from a Prior Assemblage

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

The mixture proportions summarize the composition of the assemblage and mediate the relative likelihoods of the different animal groups, adjusting a specimen's membership probabilities. Prior distributions for the mixture components reflect our prior beliefs about the relative proportions of immature, adult-sized females, and adult-sized males in the assemblage. Instead of estimating the prior belief for each of these three related categories, the model uses two prior distributions to estimate independent variables: the proportion of immature animals (π_1) and the adult sex ratio estimated through the relative proportion of adult females ($\frac{\pi_2}{\pi_2 + \pi_3}$) (see Model Supplement for more details). The following examples show some of the flexibility researchers have when describing their prior belief about the proportion of immature animals or the adult sex ratio in an assemblage; the Model Supplement shows the mathematical details necessary to create relevant

281 prior distributions for a model.

282 Figure 2 shows three examples of prior distributions for one of these mixture proportion concepts (pro-
283 portion of immature or adult sex ratio) that reflect different expectations based on prior knowledge. Note
284 that the Bayesian mixture model uses the observations of fusion rates and morphological sex ratios as ob-
285 served data, so these prior distributions reflect knowledge prior to even those observations. Figure 2A shows
286 a relatively broad, or uncertain, prior distribution where a researcher doesn't believe that the proportion is
287 extreme (i.e., that the assemblage is either dominated by or bereft of immature animals or that the sex ratio
288 is dominated by females or males) but has little opinion otherwise. Figure 2B shows a somewhat inverse
289 situation, where the researcher is confident that the proportion (either the proportion of immature animals
290 or the adult sex ratio) is at either one extreme or the other but isn't sure which extreme it is. Figure
291 2C displays a scenario where the researcher is confident that the proportion is centered around 67% before
292 looking at the faunal data, presumably based on prior research or other contextual information.

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

307 Prior predictive checks, shown in the Model Supplement, show the implications of the prior distribution
308 definitions used in the model. The results show that the chosen prior distributions do not exclude the
309 possibilities of extreme values in mixture components and cover a wide range of potential size measurements.
310 On the other hand, these model definitions are not so broad as to include much prior weight on biologically
311 impossible values (i.e., impossibly large measurements) that would slow down how quickly the model runs
312 because it must evaluate the fit of extremely poorly fitting data. Such behavior could also produce biologically
313 implausible results of actual model fits when only a small amount of data are available.

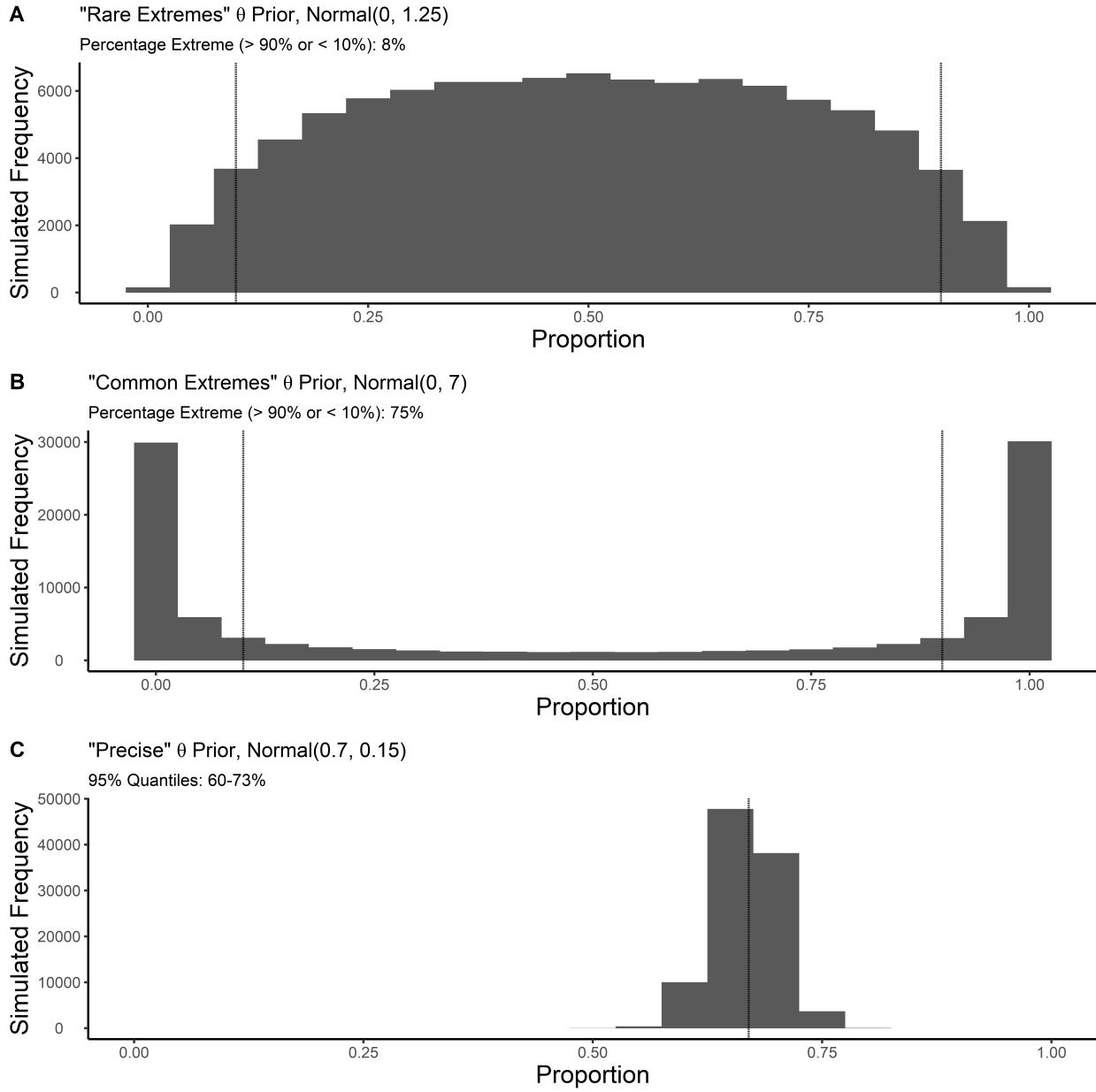


Figure 2: Examples of prior distributions for demographic parameters (proportion of immature animals or adult sex ratio) with different expectations. A: A distribution where extreme values are considered unlikely but otherwise most values are about equally as likely. B: A distribution where extremely 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.

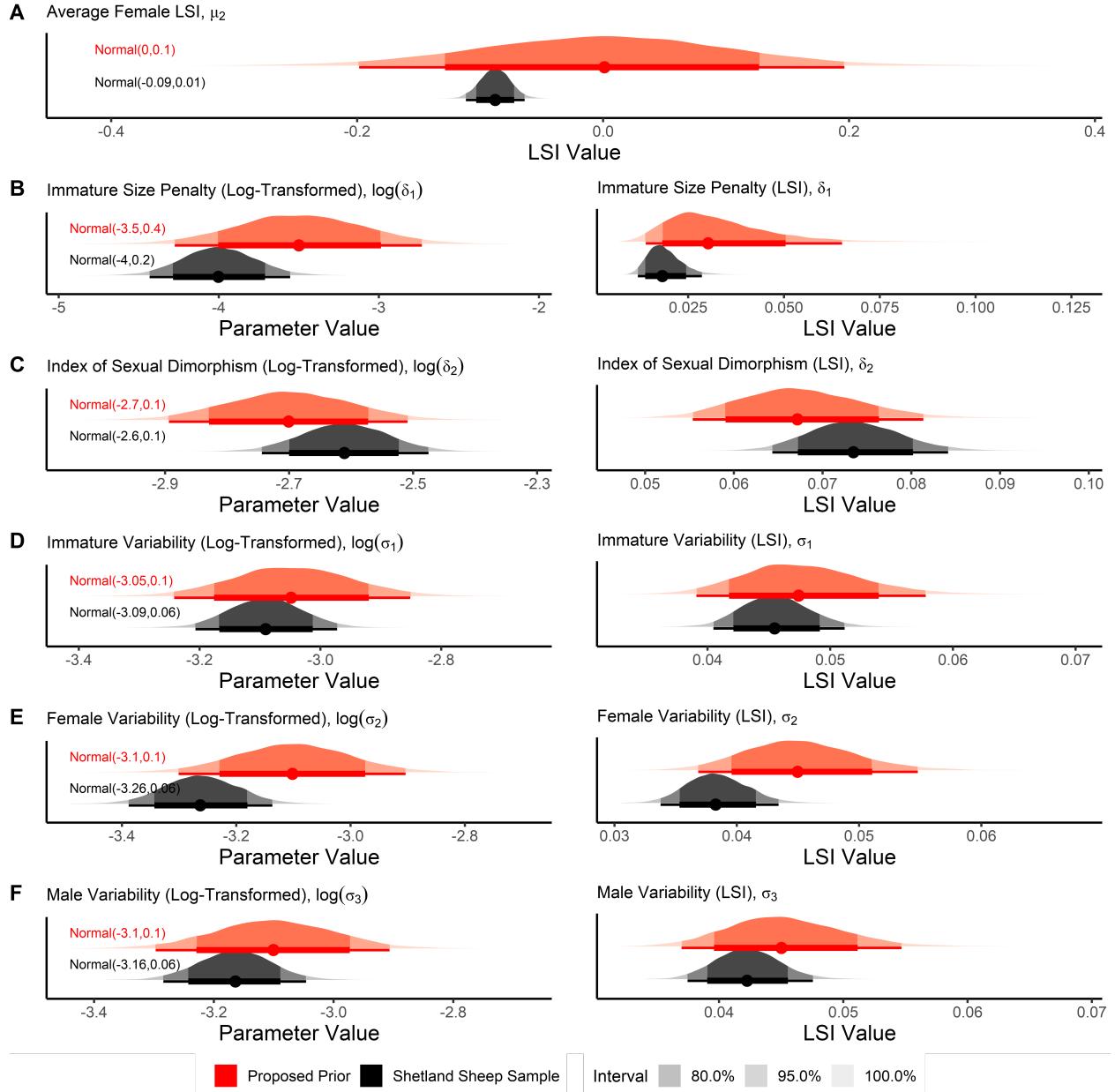


Figure 3: Posterior distributions of model hyper-parameters from a sample of known-identity sheep specimens (red) and proposed prior distributions for mixture model applications (black). Prior distributions describe (A) average female body size, in LSIE units; (B) the size difference between average 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.

314 This section describes one approach for defining prior distributions of mixture models that are relevant
315 for a wide range of zooarchaeological cases, particularly if researchers do not have strong preconceptions
316 about the relevant parameters from prior research. It is important to remember that a model's prior dis-
317 tributions are choices made by the researcher to fit particular research questions, regardless of whether the
318 distributions are informed by advice on reference priors, prior assemblages, or mathematical summaries of
319 existing research. Zooarchaeological assemblages resulting from catastrophic kills would be expected to have
320 a different ontogenetic age and sex composition compared to assemblages derived from sustained hunting
321 or herding take-off (e.g., lyman_1987?). Similarly, other research contexts may provide an analyst with
322 different prior expectations about animal size variability and overall biometry. Other reference populations,
323 particularly those from other taxa, could also be used to create prior assemblages and help determine the
324 limits on biological feasibility. Researchers could and should adapt their prior distributions to best reflect
325 their intuition about likely parameter values for their research context. Regardless of the prior distributions
326 a researcher uses, it is crucial to formally describe the prior distributions that is used in a Bayesian analysis
327 to ensure replicability. Further, researchers should examine the implications of different candidate prior
328 distributions while developing a Bayesian model to test a research question; prior distributions should be
329 regularly tested even before models are fit to datasets (Gelman, Carlin, et al. 2020).

330 ***2.3 Extending the Multilevel Analysis to Multiple Sites***

331 The multilevel structure of the model that allows parameters to vary across element portions can also be used
332 to extend the modeling approach to examine multiple assemblages at once. Combining multiple assemblages
333 into a single model allows researchers to investigate regional variation in herd management strategies or
334 outline diachronic trends in body size that may relate to population turnover (e.g., Arbuckle and Atici 2013;
335 Arbuckle et al. 2016). It also allows researchers to model diachronic changes over the course of a multi-
336 period site's occupation, as each occupation layer can be defined as a separate assemblage (e.g., Hongo et al.
337 2009; Wolfhagen et al. 2021). By including the assemblages in the same model, estimates share the same
338 hyper-parameters, which improves the precision of these estimates and allows researchers to directly compare
339 assemblage-specific parameters by using contrasts. Further, adopting this structure provides the foundation
340 for more sophisticated analyses that test specific hypotheses about variation in biometric or compositional
341 parameters, such as spatiotemporal autocorrelation in body size.

342 An important consequence of extending the model to evaluate multiple sites at once is that the interpreta-
343 tion of the overall hyper-parameters that the researcher inputs into the model changes. Instead of describing
344 the overall estimates for a specific assemblage, these hyper-parameters now describe a 'grand mean' of the
345 parameter value for all the included assemblages. These overall summaries could be interpretively useful if,

346 for instance, all the assemblages come from a discrete archaeological culture or region. In other scenarios,
347 however, the interpretation of these overall hyperparameters may be less meaningful than comparisons of
348 assemblage-specific estimates that still account for anatomical variation within each assemblage (see Model
349 Supplement for more details).

350 **2.4 Interpreting Model Results: Measured, Modeled, and Full Assemblages**

351 The results of the Bayesian multilevel mixture model include specimen-specific membership probabilities
352 (π_{Specimen}) based on the mixture model parameters. While these membership probabilities can be used to
353 calculate “critical size limits” where the largest membership probability shifts from one group to another (e.g.,
354 Monchot and Léchelle 2002), they can also be used to simulate assemblages of known-group specimens to
355 examine age/sex-stratified estimates of body part representation and sex-stratified fusion rates. Membership
356 probabilities (π_{Specimen}) are used to simulate the specimen’s identity by sampling from the probabilities using
357 a multinomial distribution; in each posterior sample, a single simulated assemblage is created, resulting in a
358 distribution of simulated assemblages with known age/sex assignments (Crema 2012). The characteristics of
359 these assemblages can then be used to summarize the overall assemblage or identify differences in composition
360 based on element types, fusion states, sub-assemblage features, or other pertinent factors that a researcher
361 is interested in examining in relation to the composition of the assemblage.

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

371 We can avoid this bias by formalizing the relationship between the *measured* and *modeled* assemblages by
372 stating that the measured assemblage is a sample of the modeled assemblage, wherein inclusion is governed
373 by a specimen’s *measurability*—the preservation of specific bony portions that allow for biometric measure-
374 ment(s). If we assume that measurability is unrelated to a specimen’s ontogenetic age or sex, then we can
375 assume that the measured assemblage is a random sample of the modeled assemblage. Thus, an unmeasured
376 specimen will have the same model parameters (mixture proportions π , average size μ , and size variability σ)
377 of the measured specimens from the same element portion. Crucially, this means that we can include unmea-

sured specimens in our simulated assemblages by using the relevant mixture proportions π (adjusted for the specimen's fusion data as necessary) as that specimen's membership probabilities (π_{Specimen}). Leveraging the multilevel structure of the model further, we can assume that the overall mixture model hyper-parameters for *modeled* element portions are equally valid for *unmodeled* element portions. The Bayesian multilevel mixture model estimates hyper-parameters that describe the average value (μ_{Element}) and expected variability (σ_{Element}) of mixture model parameters for element portions; these hyper-parameters can be used to estimate the relevant mixture model parameters of an unobserved element portion (Gelman, Carlin, et al. 2020; McElreath 2020). The resulting parameters, then, could be used to estimate π_{Specimen} membership probabilities for the unmodeled (and unmeasured) specimens, as in the first extension, creating an estimate of the composition of the *full* assemblage.

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

2.5 Computational Details of the Bayesian Analysis

The Bayesian multilevel mixture model is written in Stan, version 2.30.0 (Team 2022). All analyses in this paper use R version 4.1.2 (2021-11-01), in Rstudio 2023.3.0.386 (Cherry Blossom) (R Core Team 2022; RStudio Team 2022); Table 3 lists the packages, versions, and citations for the packages used in the analytical scripts. The model Stan code and analytical R code necessary to replicate and apply the analyses in this paper are freely available in a GitHub page and Open Science Framework page. The files include a copy of the Shetland sheep data file from the supplemental files published in Popkin et al. (2012) and archaeological datasets for the case studies downloaded from OpenContext (Buitenhuis 2013; Carruthers 2006; Galik 2013; Gourichon and Helmer 2013). The analytical code includes two script files—a script for replication and one

410 for application. The R markdown file (“ZooarchMixMod.Rmd”) file replicates the entire analytical workflow
 411 of the paper, with a specific seed set to ensure exact replicability of the submitted manuscript. Another set
 412 of scripts to standardize the analytical workflow for faunal datasets structured like the OpenContext faunal
 413 datasets used in these case studies, see the GitHub for more details. All scripts (R and Stan) are released
 414 under the MIT license and figures are released as CC-BY to encourage reuse and reproducibility (Marwick
 415 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, 2021
Data aggregation, analysis, and multi-core processing	readxl	1.4.0	Wickham & Bryan, 2022
Data aggregation, analysis, and multi-core processing	parallel	4.1.2	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	1.38	Xie, 2015, 2022
Bayesian analysis and summarization	cmdstanr	0.4.0	Gabry & Cešnovar, 2022
Bayesian analysis and summarization	rstan	2.21.3	Stan Development Team, 2021
Mixture model analysis and standard animal measurements	mixtools	1.2.0	Benaglia, et al., 2009
Mixture model analysis and standard animal measurements	zoolog	0.4.1	Pozo, et al., 2021
Visualization	Cairo	1.5.15	Urbanek & Horner, 2022
Visualization	ggplot2	3.3.5	Wickham, 2016
Visualization	ggdist	3.1.1	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	1.0.7	Pebesma, 2018

416 3. TESTING THE BAYESIAN MULTILEVEL MIXTURE MODEL

417 Two sets of tests are used to evaluate different aspects of the Bayesian multilevel mixture model. First,
 418 the accuracy of the model’s ability to reconstruct the age and sex composition of assemblages is tested
 419 using simulated faunal assemblages of known age and sex from the Shetland sheep population. This test
 420 evaluates both the single-assemblage model and the multi-assemblage model. Second, two archaeological
 421 case studies showcase the applicability of the model to archaeological data and the added insights gained
 422 from adopting Bayesian multilevel mixture models. The simulated assemblage case study and the single
 423 assemblage archaeological case study use sheep (*Ovis aries*) measurements, with standard measurements

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

424 coming from a female wild sheep (*Ovis orientalis* FMC 57951: Uerpmann and Uerpmann 1994: Table 12).
 425 The multiple assemblage case study uses cattle (*Bos taurus*) measurements, with standard measurements
 426 coming from a wild female aurochs (*Bos primigenius* “Ullerslev”: Degerbøl 1970). Two dimensions of the
 427 standard cow (Scapula GLP: 89 mm; and Calcaneus GB: 46 mm) were not included in the ‘zoolog’ output
 428 and were included manually, drawn from the referenced source.

429 **3.1 Simulated Assemblages**

430 A series of simulated assemblages of known age and sex composition are created from the Shetland sheep
 431 population by randomly drawing element portions (and all associated measurements) from the total assem-
 432 blage without replacement. Table 4 describes the measured dimensions included in the simulation analyses
 433 from the 10 element portions. The first test, using a single-assemblage model, uses 150 element portions
 434 from the Shetland sheep population where every element portion has an equal probability of being selected.
 435 There is no guarantee, however, that the element portions have equal representation or even that all element
 436 portions are present in the simulated assemblage, which better approximates archaeological assemblages.
 437 The result of this first simulation produces an assemblage of 231 measurements from 125 individual animals.
 438 Using the same procedure, the second test creates 15 simulated assemblages that are analyzed in a single
 439 multi-assemblage model. Demographic observations for phalanx fusion rates and pelvis sex ratios were also
 440 simulated from the Shetland sheep population using the same underlying probabilities as the measurement
 441 assemblages. Table 5 describes the sample sizes of the measurement assemblages, including any manipula-
 442 tions to the measurement values. The specific elemental composition and measurements of the assemblages,

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

Note:

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

* Size increased for males only

along with the simulated demographic observations, used in both simulations can be recovered from the replication script with the recorded random seed (see also Supplemental Tables S1-S3); using another random seed would provide a conceptual replication of new assemblages drawn from the same underlying populations.

While the simulated assemblages are derived from the same Shetland sheep population that was sampled to create the ‘prior assemblage’ (see Section 2.2), there are several key justifications of this double use. First, the prior distributions used in the model differ from the results from the ‘prior assemblage’ (Figure 3); prior predictive checks of the single assemblage and multisite models show that the prior distributions are flexible enough to allow a wide range of potential assemblages (see Model Supplement). Logistically, the Popkin et al. (2012) population is the most complete fully-published assemblage of standard measurements, particularly including immature, adult female, and adult males; Davis (1996); Davis (2000) describes similar sheep, but does not include any immature specimens. Finally, the simulated assemblages vary in sample size, and some assemblages are manipulated to vary in average body size and expected composition from the original Shetland sheep population. These modifications are important to try to avoid issues of ‘prior mimicry’, as seen in survivorship modeling (e.g., Millard 2006); this also stresses, however, the importance of developing additional sources of ‘prior assemblages’ to test or develop relevant prior distributions.

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

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

470 Compositional accuracy does not have the same structure as parametric accuracy because there is no
471 underlying population value for composition: there is only the true number of immature, adult female,
472 and adult male specimens in the measured and modeled assemblages. Again, though, it is important to
473 understand accuracy in the context of overfitting and underfitting. Overfitted results, wherein credible
474 intervals about the number of immature, adult female, and adult male specimens contain the true abundances
475 at a lower rate than designed (e.g., fewer than 95% of the 95% credible intervals), could lead to a researcher
476 declaring an imbalance in the demographic composition of an element where one does not exist (or is even
477 imbalanced in the opposite direction). Underfitted results, by contrast, would mean that a researcher is
478 unable to identify an imbalance where one exists because the credible intervals are too wide. It is important,
479 then, to use the simulations to understand the kinds of errors the model is prone to making so that researchers
480 avoid overinterpretation.

481 Modeled assemblages were created for the single-assemblage simulation and the multi-assemblage sim-
482 ulation by assuming that measured specimens represent 20% of the overall assemblage and sampling more
483 specimens from the Shetland sheep population to create the remaining 80% of the assemblage. For example,
484 in the single-assemblage simulation with 150 measured specimens, this means sampling 600 more specimens
485 from the Shetland sheep population to create a total modeled assemblage of 750 specimens. Specimens could
486 not be repeatedly sampled, though multiple specimens could be from the same individual. As described in
487 Section 2.4, unmeasured specimens use the relevant π parameters for the element portion. For the multi-
488 site simulation, this potentially includes element portions where there are no relevant measurements in the
489 specific assemblage.

490 Because the “grand mean” parameters in the multisite simulation no longer represent the same thing
491 as in the single assemblage model (see Model Supplement), the prior distributions must also be changed to
492 reflect different expectations. Again, the goal of these prior distribution definitions is to prevent extreme

493 overfitting so that parameter estimates are biologically feasible (Gelman et al. 2008). In general, the centers
494 of the distributions stayed the same, but the uncertainty was increased to reflect the fact that there's
495 less certainty about biologically feasible values for multiple populations, especially if there is size variation
496 expected between the assemblages. These prior distributions are listed in the Model Supplement.

497 **3.2 Archaeological Case Studies**

498 The Bayesian multilevel mixture model is applied to two archaeological case studies to showcase its utility
499 for both interpreting a single assemblage and examining multiple assemblages. In both case studies, the
500 sheep and cattle measurements have been previously published on OpenContext and the general zooarchaeo-
501 logical summaries of the assemblages have been published as well (Buitenhuis 2008, 2013; Carruthers 2005,
502 2006; Galik 2013; Gerritsen and Özbal 2019; Gourichon and Helmer 2008, 2013). Again, LSI_e values are
503 calculated using the same standard animal as the simulation analysis for the single assemblage analysis, the
504 *Ovis orientalis* female standard animal (FMC 57951) from Uerpman and Uerpman (1994, Table 12) and
505 the *Bos primigenius* female standard animal (“Ullerslev”: Degerbøl 1970; Grigson 1989), operationalized
506 through ‘zoolog’ functions (Pozo et al. 2021). Alongside metric data, the OpenContext faunal tables pro-
507 vide demographic data that can be used to observe relevant estimates of the age and sex composition of the
508 assemblages. The goal of applying the mixture model to these assemblages, then, is to use the metric data
509 to improve estimates of the age and sex composition of the assemblage, biometric estimates, and sex-specific
510 fusion rates.

511 **3.2.1 Single Assemblage: Biometric Analysis of Sheep from 7th Millennium BCE Central
512 Anatolia (Pinarbaşı B)**

513 The site of Pinarbaşı, located in the Konya Plain of central Turkey, consists of a series of rock shelter and
514 open-air sites at the foothills of the Karadağ volcanic region and Lake Hotamış and its associated wetlands
515 (Baird et al. 2011; Kabukcu 2017). This case study examines the Pinarbaşı B late Neolithic occupation,
516 which is dated to the second half of the 7th millennium BCE and includes a large number of domesticated
517 sheep and goat remains (Baird et al. 2011; Carruthers 2005). Carruthers (2005) analyzed fauna from the
518 1994-1995 excavations by Trevor Watkins (Watkins 1996), interpreting the presence of fetal sheep remains
519 and other juvenile remains in the assemblage as evidence for herders penning sheep on-site. The Neolithic
520 assemblage was thus described as the result of seasonal occupation by sheep and goat herders during the
521 lambing season and the fall, with culling in the spring possibly focused on young males (Carruthers 2005).
522 This analysis makes several claims that can be evaluated with the Bayesian multilevel mixture model: the
523 dominance of immature remains and a female-dominated adult sex ratio.

524 The Bayesian multilevel mixture model for the late Neolithic Pinarbaşı B assemblage uses 44 sheep
 525 measurements from 44 specimens (see Table 6; Supplemental Table S4). In addition to these measurements,
 526 the observed proportion of immature animals from unfused first and second phalanges is 59 / 62 (95%),
 527 including specimens identified to sheep and to sheep/goat. There are 0 observed sheep (or sheep/goat)
 528 pelvis bones with sex identifications; this is entered into the model by having an observed adult sex ratio for
 529 the assemblage of 0 / 0 (females / females + males). All data come from the Pinarbaşı faunal assemblage
 530 uploaded to OpenContext, focusing only on specimens in the Site B Neolithic contexts (Carruthers 2006).
 531 The Pinarbaşı B sheep model uses the same prior distribution definitions for the model hyper-parameters as
 532 the single assemblage simulation since both models, even though the sheep body sizes likely differ between
 533 the two populations, showcase the flexibility of the standard prior distribution definitions.

Table 6: Elemental composition of the Pinarbaşı B assemblage. Dimensions definitions follow von den Diesch (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

534 **3.2.2 Multiple Assemblages: Biometric Analysis of Cattle from 7th-6th Millennium BCE**
 535 **Northwest Anatolia (Barçın Höyük, İlipinar Höyük, Menteşe Höyük)**

536 Understanding the development of Neolithic communities in northwestern Anatolia has long been of interest
 537 for researchers interested in studying the spread of agricultural lifeways from southwest Asia into Europe (e.g.,
 538 Çakırlar 2013; Karul 2019; Özdogan 2011, 2019). Agricultural communities first appear in the Marmara
 539 region in the mid-seventh millennium BCE in sites like Barçın Höyük (Gerritsen and Özbal 2019; Karul
 540 2019). The domestic animal economies of these Late Neolithic and Early Chalcolithic communities appears
 541 to be focused on cattle and caprine (sheep and goat) herding, rather than pig husbandry (Buitenhuis 2008;
 542 Çakırlar 2013; Gourichon and Helmer 2008). Milk residues on pottery recovered from these sites suggest
 543 these communities regularly consumed milk, potentially orienting herd management strategies of sheep,
 544 goats, and particularly cattle to specialize in milk production (Evershed et al. 2008; Thissen et al. 2010).

545 Four archaeological components from three sites are used in this case study, located near Lake İznik

546 and on the Yenişehir Plain in the Bursa province of Turkey (Figure 4). The Neolithic layers from Barçın
547 Höyük (Phase VI) is the earliest of these assemblages, with occupation roughly from 6500-6000 cal BCE;
548 excavations revealed a subsistence economy focused on cereal agriculture and the herding of cattle, sheep, and
549 goat (Galik 2013; Gerritsen and Özbal 2019). Menteşe Höyük is located approximately five km west of Barçın
550 Höyük on the Yenişehir Plain; the three Neolithic layers at the site date to 5800-5600 cal BCE (Gourichon
551 and Helmer 2013; Roodenberg et al. 2003). Previous faunal analysis of the Neolithic assemblage identified
552 animal economies that shifted from predominantly cattle to sheep herding over the course of the occupation
553 (Gourichon and Helmer 2008). İlipinar Höyük is located near Lake İznik, separated from the Yenişehir Plain
554 by a mountain ridge (Roodenberg 2012a). The Neolithic/Early Chalcolithic occupation of the site spanned
555 6200-5400 cal BCE (Buitenhuis 2013); the assemblage is divided into two sub-assemblages (Neolithic İlipinar
556 = Phases X-VII, 6000-5700 cal BCE; Chalcolithic İlipinar = Phases VI-V, 5600-5400 cal BCE), marked by
557 the introduction of mudbrick architecture and expanded storage (Roodenberg 2012a, 2012b). Sheep and
558 goat are common in the earlier assemblages of the site, with cattle becoming predominate in later phases of
559 the site (Buitenhuis 2008; Roodenberg 2012a). Notably for this biometric analysis, Buitenhuis (2008) notes
560 cattle body sizes are stable throughout site's occupation.

561 The Northwest Anatolian cattle bone assemblages consist of 614 measured specimens spread unevenly
562 across the four components (Barçın Höyük N = 67, Menteşe Höyük N = 45, Neolithic İlipinar N = 249,
563 Chalcolithic İlipinar N = 253; Supplemental Table S5). All measured *Bos* remains were included in the
564 analysis, rather than separating out those identified as aurochs (*Bos primigenius*, N = 3) or identified only
565 to *Bos* spp. (N = 134) in the İlipinar Höyük dataset; all specimens were only labeled as “*Bos*” in the Menteşe
566 Höyük dataset. Table 7 shows the composition of the four Northwest Anatolian measurement assemblages.
567 Demographic observations of the proportion of immature animals and the adult sex ratio for each assemblage
568 describe these parameters. For the four Northwest Anatolian assemblages, estimates of the assemblage-level
569 proportion of immature specimens based on the fusion rates of proximal and middle phalanges for cattle
570 specimens are 28 / 87 (32%) for Barçın Höyük, 28 / 184 (15%) for Neolithic İlipinar, 8 / 25 (32%) for Menteşe
571 Höyük, and 9 / 89 (10%) for Chalcolithic İlipinar. The observed adult sex ratios (females / females + males)
572 based on cattle pelvis morphology are 3 / 4 (75%) for Barçın Höyük, 0 / 0 for Neolithic İlipinar, 0 / 0 for
573 Menteşe Höyük, and 3 / 5 (60%) for Chalcolithic İlipinar. As in the Pınarbaşı B example, observations of 0 /
574 0 impart no information onto the prior distribution of the adult sex ratio. All demographic and measurement
575 data come from the OpenContext datasets (Buitenhuis 2013; Galik 2013; Gourichon and Helmer 2013); the
576 associated RMarkdown file (“ZooarchMixMod.Rmd”) includes the steps for data processing and analysis.

577 Previous syntheses of the Late Neolithic and Early Chalcolithic animal economies in Northwest Anatolia
578 provide several prior inferences about the age and sex structure of cattle bone assemblages that can be

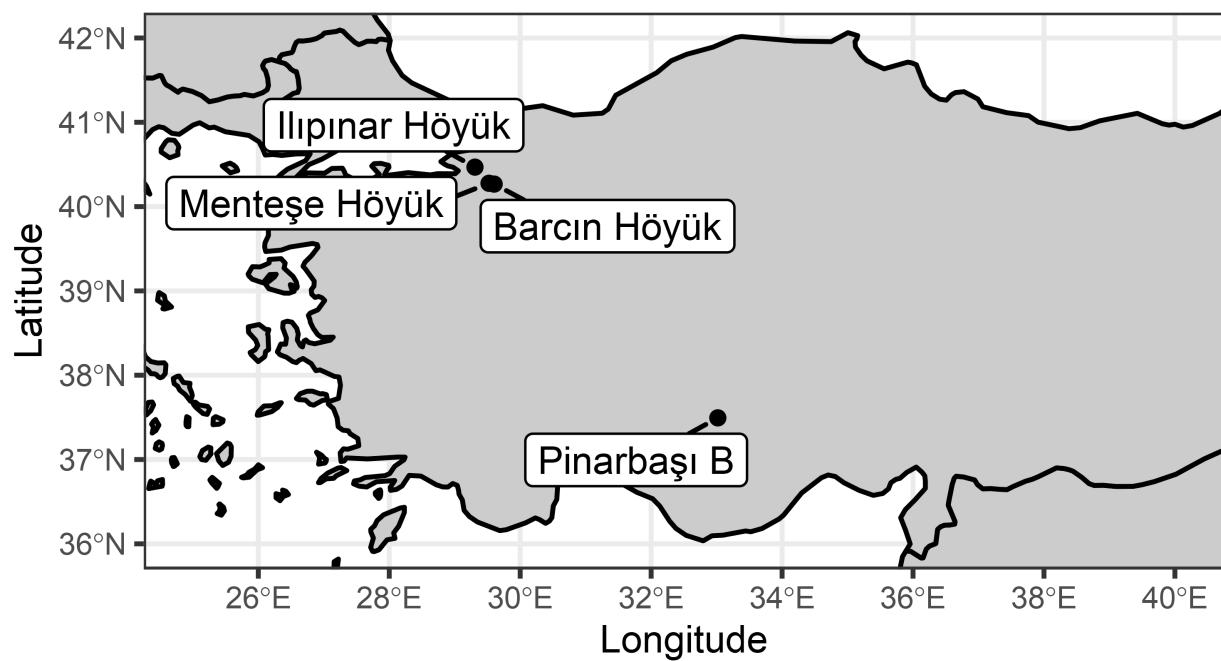


Figure 4: Map of archaeological sites included in this analysis

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

Element Portion	Barcın Höyük	Ilıpınar Höyük (Late Neolithic/Transitional)	Ilıpınar 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
Total	67	262	296	46

evaluated with the results of the Bayesian multilevel mixture model. First, the general cultural continuity of the assemblages suggests that the biometry and composition of cattle bone assemblages may be similar at the sites, having been produced by similar processes (Çakırlar 2013; Özdoğan 2019); Buitenhuis (2008, 312) explicitly states that there is no size change among cattle bones across the Ilıpınar assemblage. Second, the widespread evidence of milk consumption from pottery residue analyses from these sites and others in the region (Evershed et al. 2008; Thissen et al. 2010) has led some researchers to argue that cattle were managed for milk production (Gourichon and Helmer 2008; Roodenberg 2012a). Gourichon and Helmer (2008, 440) argue that the cattle tooth eruption and wear data at Menteşe indicate exploitation focused on milk consumption; one consequence of this pattern should be female-dominated adult sex ratios, including higher fusion rates for later-fusing elements among females than males (Zeder and Hesse 2000). The multilevel modeling results can be used to evaluate the feasibility of these inferences by examining posterior distributions of relevant parameters and simulations of sex-specific fusion rates.

Because this application is a multisite model and deals with a different taxon than the original simulations, the prior distributions for the model hyper-parameters are redefined to reflect different expectations of biological feasibility. While the multisite simulation provides useful prior distribution definitions for most of the parameters, two other parameters (average body size of females μ_2 and index of sexual dimorphism $\log_e(\delta_2)$) should be changed because of different expectations modeling cattle rather than sheep. The change in the prior distribution definition of μ_2 reflects the fact that the standard measurements for cattle come from an aurochs female (Degerbøl 1970), which is expected to be larger than the domestic cattle females in the assemblages. Cattle are expected to be more sexually dimorphic than sheep, which is reflected in increasing the average expected value of $\log_e(\delta_2)$, resulting in an expectation of 0.14 LSI_e units between males and females on average. This is slightly lower than the index of sexual dimorphism seen in the Degerbøl (1970)

601 aurochs specimens [Grigson (1989): Figure 2, which uses LSI₁₀; the equivalent size difference is 0.06 on
602 the LSI₁₀ scale], though domestic cattle may be expected to be less sexually dimorphic than their wild
603 counterparts (e.g., Tchernov and Horwitz 1991); these prior distribution definitions are listed in the Model
604 Supplement.

605 4. RESULTS

606 Bayesian models that use Monte Carlo methods, like the ones used here, rely on convergence diagnostics to
607 ensure that the results (posterior distributions) have converged to the target distribution—that the results
608 are not unduly affected by the random starting position of the analysis. To do this, analysts run multi-
609 ple independent chains of the model—starting from different initial values—then evaluate how similar the
610 chains are to one another using different diagnostic criteria (e.g., **gelman_rubin_1992?**). Supplemental
611 Tables S6-S11 show the posterior estimates of the (overall and site-specific) model hyper-parameters and
612 diagnostic criteria (R-hat and effective sample size). These results are consistent with the model successfully
613 converging, as R-hat values are 1.01 and effective sample sizes are greater than 100x the number of chains
614 (**vehtari_et al._2021?**). Trace plots show the value of a parameter at each posterior sample, with each
615 chain overlain on top of each other; a converged model should have no directionality across the length of
616 the chain and the independent chains should be indistinguishable from one another. Trace plots of each
617 model’s overall hyper-parameters are shown in Supplemental Figures S1-S4. None of the parameters show
618 extreme deviations between chains, supporting the assertion that the model’s posterior distributions properly
619 describe the data and prior beliefs.

620 4.1 Simulated Assemblages: testing model accuracy

621 Bayesian models work by updating prior information with new data to produce posterior distributions of
622 parameters of interest (Otárola-Castillo et al. 2022). Thus, the difference between a model parameter’s
623 prior and posterior distribution shows the amount that the model “learns” from the data. If the data do not
624 provide relevant information on a parameter’s potential values, then the posterior distribution will resemble
625 the prior distribution. Figure 5 compares the prior and posterior distributions of the main model hyper-
626 parameters for the single assemblage simulation (for prior-posterior comparisons of the other models, see
627 Supplemental Figures S5-S7). The results show that the data provides much more information about the
628 likely values of the two demographic parameters (the proportion of immature animals, π_1 , and the adult sex
629 ratio, $\frac{\pi_2}{\pi_2 + \pi_3}$) and the average 594 body size for females (μ_2). This is largely to be expected, as the prior
630 distribution definitions were weakly-informative priors (Gelman et al. 2008), but it also shows how these

631 choices did not appear to severely influence the resulting posterior distributions.

632 The prior distribution definitions for the size offsets (δ_1 and δ_2) and the size variability estimates (σ_1 , σ_2 ,
633 and σ_3) have a lot more overlap between the prior distributions and their respective posterior distributions.
634 This overlap stresses the importance of using a Bayesian framework, particularly one relying on informative
635 prior distributions, to produce meaningful parameter estimates from zooarchaeological data. But it also
636 highlights the interpretive weight given to the reference population. However, the overlap is not necessarily
637 a drawback of the model, as again the prior distribution definitions were designed as informative priors,
638 specifically to ensure that the resulting parameter estimates would be biologically feasible. Further, the
639 simulated population also has the same underlying biological population (the Shetland sheep population)
640 that was used to develop the prior distributions, so it is possible that this overlap reflects that fact.

641 The parametric and compositional accuracy of both simulation tests are summarized in Table 8. The
642 single assemblage model is well conditioned when examining parametric accuracy, though the multisite
643 model overfits in this respect; this is driven by poor performance on size variability (σ) parameters—the
644 model estimates average body size (μ) parameters well. The multisite model also has a tendency to underfit
645 when examining site-specific compositional accuracy. In both models, though, the compositional accuracy
646 improves (in the sense of no longer underfitting) by using the modeled assemblages rather than the measured
647 assemblages. This makes intuitive sense, as the measured assemblage is itself theoretically a sample from
648 the modeled assemblage (based on the assumption that “measurability” is random). In these simulations, of
649 course, this theory is held to be explicitly true, though the relationship between the measured and modeled
650 assemblage is generally held to be true implicitly in zooarchaeology and can be explicitly tested (see Section
651 2.4).

652 The parametric and compositional accuracy of both simulation tests are summarized in Table 8. The
653 single assemblage model is well conditioned when examining parametric accuracy, though the multisite
654 model overfits in this respect; this is driven by poor performance on size variability (σ) parameters—the
655 model estimates average body size (μ) parameters well. The multisite model also has a tendency to underfit
656 when examining site-specific compositional accuracy. In both models, though, the compositional accuracy
657 improves (in the sense of no longer underfitting) by using the modeled assemblages rather than the measured
658 assemblages. This makes intuitive sense, as the measured assemblage is itself theoretically a sample from
659 the modeled assemblage (based on the assumption that “measurability” is random). In these simulations, of
660 course, this theory is held to be explicitly true, though the relationship between the measured and modeled
661 assemblage is generally held to be true implicitly in zooarchaeology and can be explicitly tested (see Section
662 2.4).

663 Figures 6 and 7 show the posterior distributions of the site-level parameters for the single-assemblage

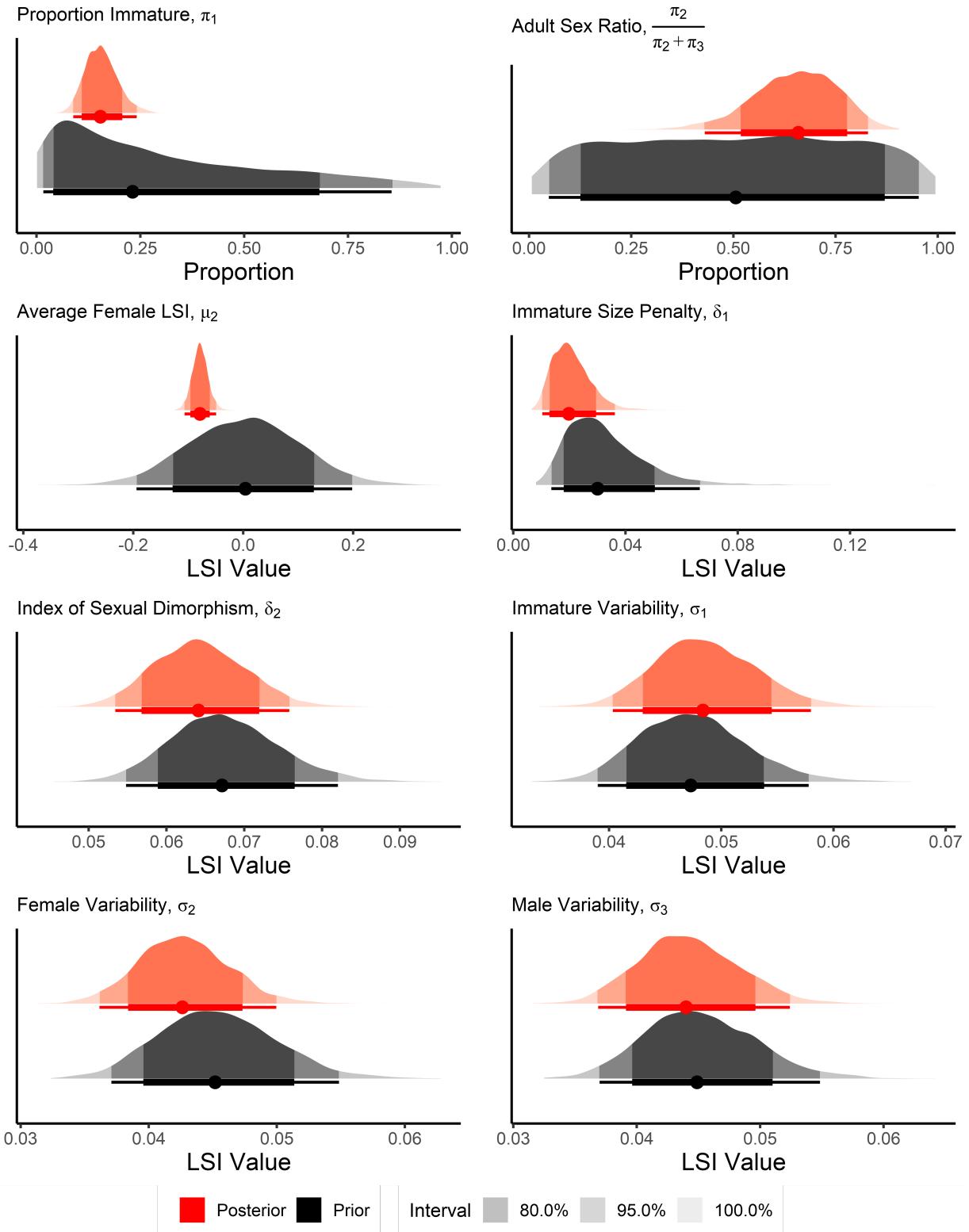


Figure 5: Comparison of prior and posterior distributions for mixture model hyper-parameters of the simulated single assemblage. The model hyper-parameters serve as assemblage-wide estimates accounting for size and composition variation across element portions.

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)	330 / 339 (97%)	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)	335 / 339 (99%)	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	22 / 24 (92%)	20-24	
Modeled Composition (95% CI)	Multisite (Overall)	22 / 24 (92%)	20-24	
Modeled Composition (95% CI)	Multisite (By Site)	350 / 360 (97%)	334-350	
Modeled Composition (95% CI)	Multisite (New Elements)	20 / 21 (95%)	18-21	

and multisite simulations. Figures 6A and 7A show the posterior distributions for the mixing proportions (π parameters), Figures 6B and 7B show the posterior distributions for the average body size (μ parameters), and Figures 6C and 7C show the posterior distributions for the size variability (σ parameters) for immature (black), adult female (blue), and adult male (red) specimens in the assemblage. Vertical bars summarize the 80% and 95% credible intervals for the parameter, while the solid horizontal lines denote the “true” parameter value for the Shetland sheep population from which the assemblage was sampled, including any deviations made for the multisite simulation (see Table 5). The model excels at estimating the average body sizes of the animal groups, even when those body sizes are manipulated (Figures 6B and 7B). The multisite model’s estimates of the overall proportions tend to be conservative (Figure 7A); that is, underestimating the proportions for animal groups with extremely low “true” proportions (e.g., adult males in Sites 6-10, immature animals in Sites 11-15) while overestimating the proportions for animal groups with extremely high “true” proportions (e.g., adult females in Sites 6-10). This is likely due to the multilevel modeling structure “shrinking” estimates towards a global mean, though it may also reflect overfitting from small sample sizes. This is particularly the case when examining variability parameters (σ : Figure 7C), which the multilevel model tends to underestimate.

Figure 8 shows posterior distributions of simulated group-specific compositions for both the single-assemblage and total composition of the multisite models (i.e., all sites combined) alongside true counts

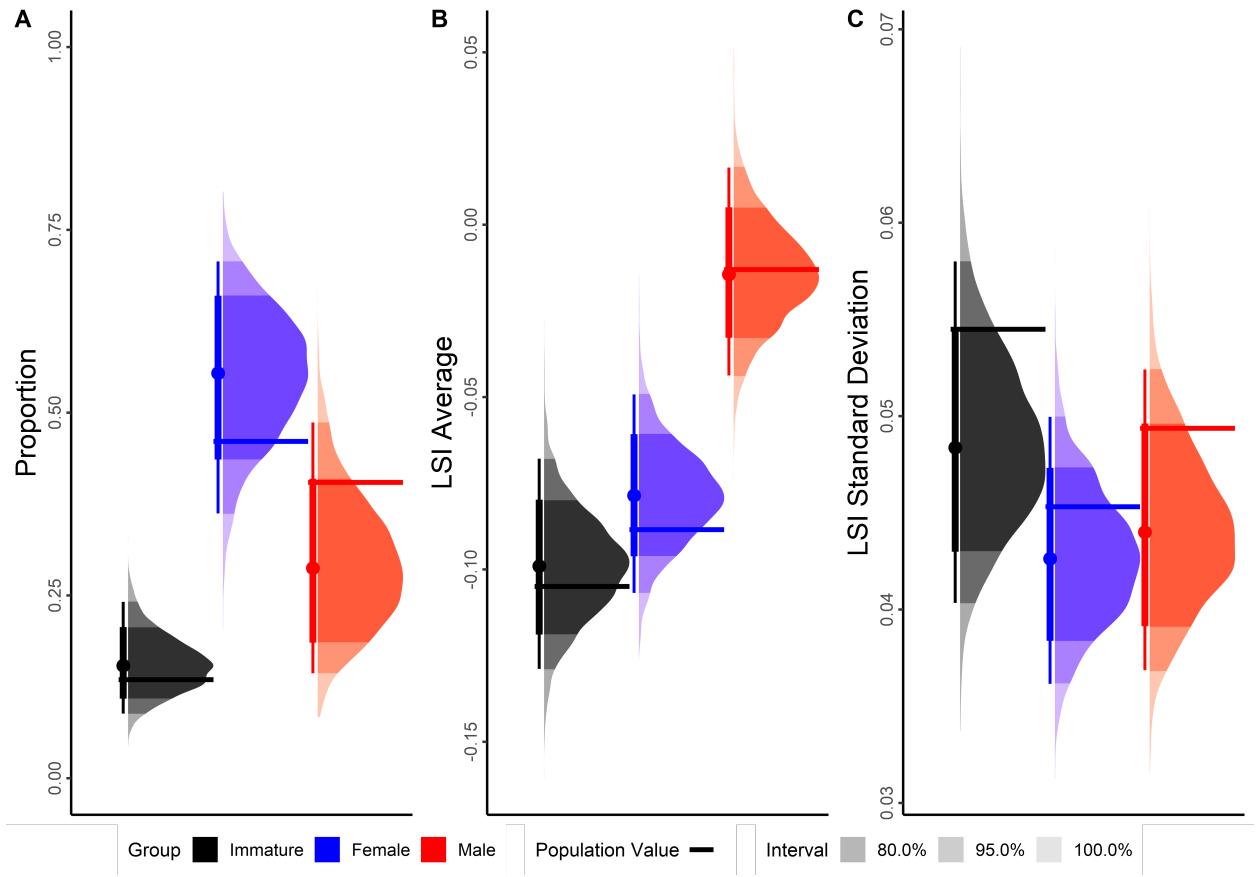


Figure 6: Posterior distributions of the mixture model hyper-parameters (assemblage-level estimates) for the simulated single assemblage. Horizontal lines denote the relevant parameter values from the Shetland sheep population from which the assemblage was sampled.

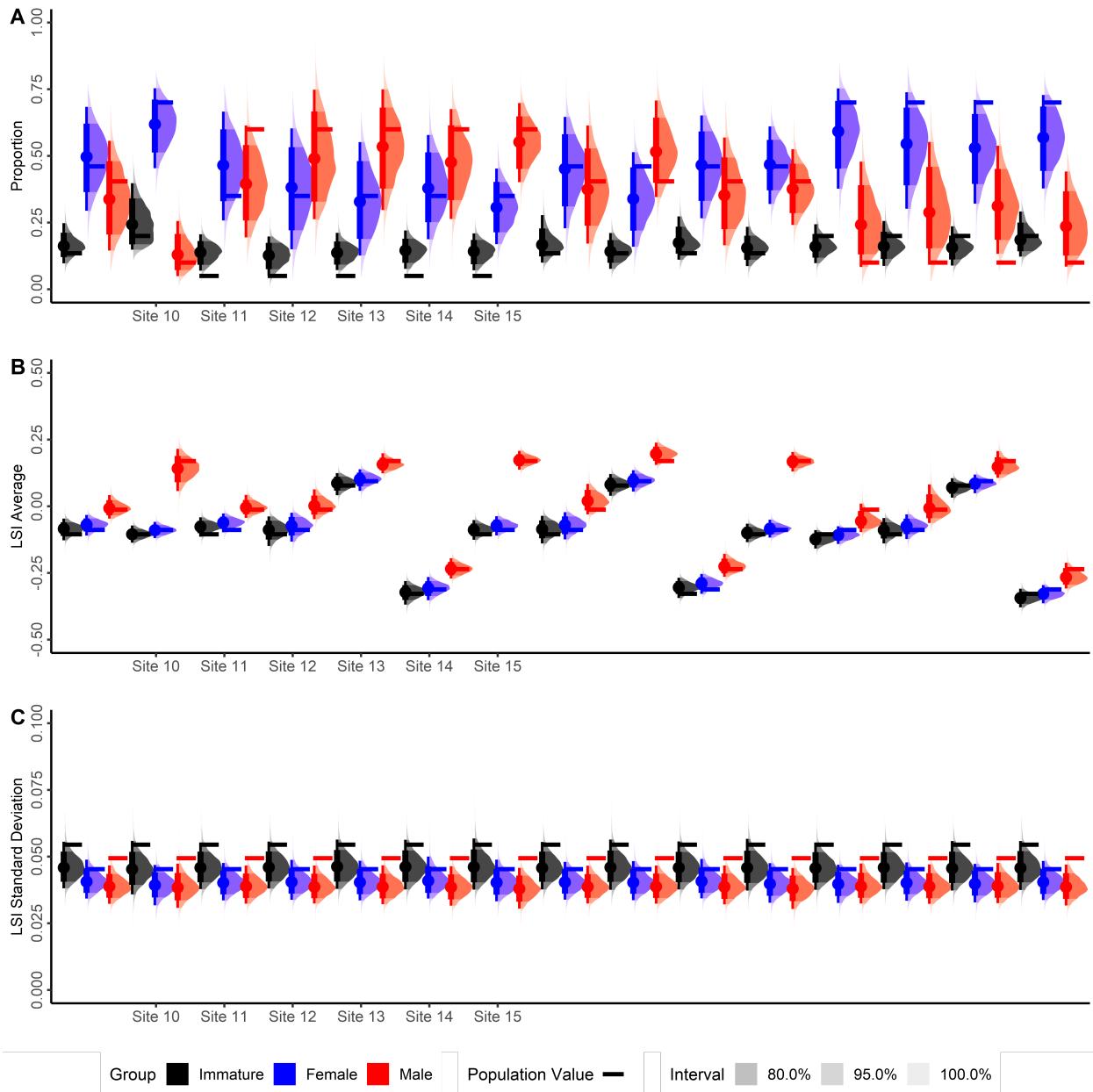


Figure 7: Posterior distributions of assemblage-level mixture model parameters for the simulated multisite assemblages. Horizontal lines denote the relevant parameter values from the Shetland sheep population from which the assemblage was sampled, including modifications to the composition and body size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

for each group. The underfitting performance of the multisite simulation, particularly at the level of individual sites, may be a consequence of low statistical power due to small sample sizes. The median number of specimens per element portion is 4 for the measured assemblage and 17 for the modeled assemblage. By contrast, the single-assemblage simulation has a median number of specimens per element portion of 19.5 for the measured assemblage and 92 for the modeled assemblage. More important than overall sample sizes, though, is the group-specific element counts, where the contrast between the median number of specimens in the measured and modeled assemblages from the multisite simulation (measurement assemblage: 1, modeled assemblage: 4) and the measured and modeled assemblages from the single-assemblage simulation (measurement assemblage: 5, modeled assemblage: 26) is starker. The larger sample sizes in the modeled assemblages also partially explains the reduced underfitting relative to the measured assemblages.

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

4.2 Pinarbaşı B Sheep: The impact of immature specimens

The first three distributions in Figure 9 show the posterior estimates of the assemblage-level proportion of immature, female, and male sheep at Pinarbaşı B. In general, the Pinarbaşı B sheep assemblage is overwhelmingly composed of immature animals (posterior μ_{π_1} median = 87%; 95% posterior credible interval for μ_{π_1} = 77-94%), somewhat lower than the observed fusion rate of proximal and middle phalanges (59 / 62 = 95%). Even though the overall proportions of female and male animals in the assemblage is low relative to immature specimens, the Bayesian multilevel mixture model can produce a posterior distribution of the overall adult sex ratio that suggests that adult females are more common than adult males (median θ_{Female} = 64%; 95% posterior credible interval for θ_{Female} = 10-96%); the wide credible interval of this distribution reflects the small proportions of mature specimens in the assemblage overall. Despite this uncertainty, we can use this distribution to estimate the probability that adult females are more common than adult males (i.e., that θ_{Female} is greater than 0.5); 68% of the posterior θ_{Female} samples are above 0.5, providing some confidence in the interpretation that the adult sex ratio is skewed towards females. Note that any threshold

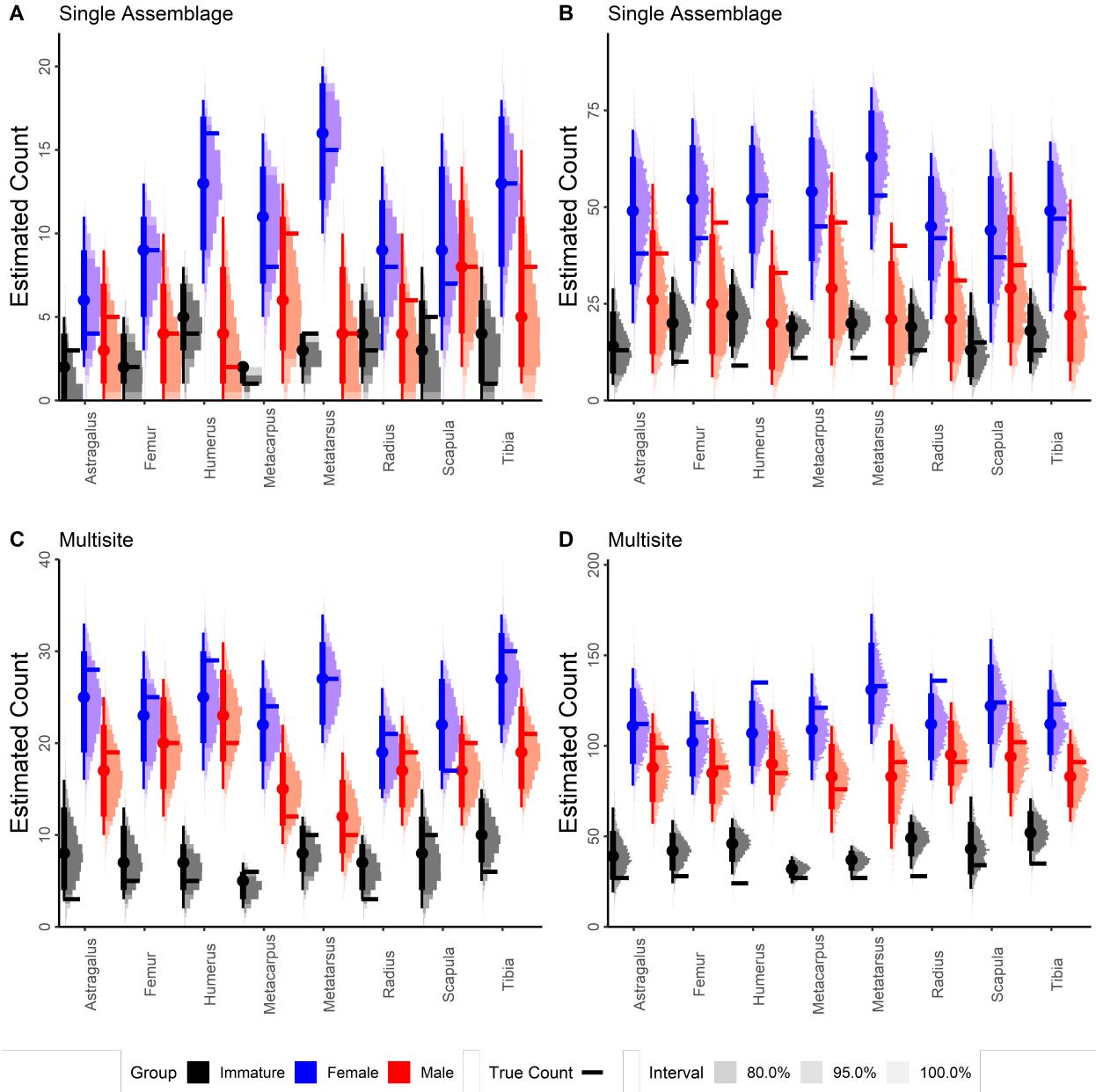


Figure 8: Posterior distributions of simulated group-specific composition for the simulated assemblages. Top row: single assemblage model (A) measured assemblage, (B) modeled assemblage. Bottom row: multisite assemblage model (combined counts) (C) measured assemblages, (D) modeled assemblages. Horizontal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

713 value could be chosen to test a hypothesis about the adult sex ratio—for instance, only 45% of the posterior
 714 θ_{Female} samples are above 0.67 (i.e., a 2:1 female:male ratio), providing poor support that the mature portion
 715 of the assemblage is heavily skewed towards females.

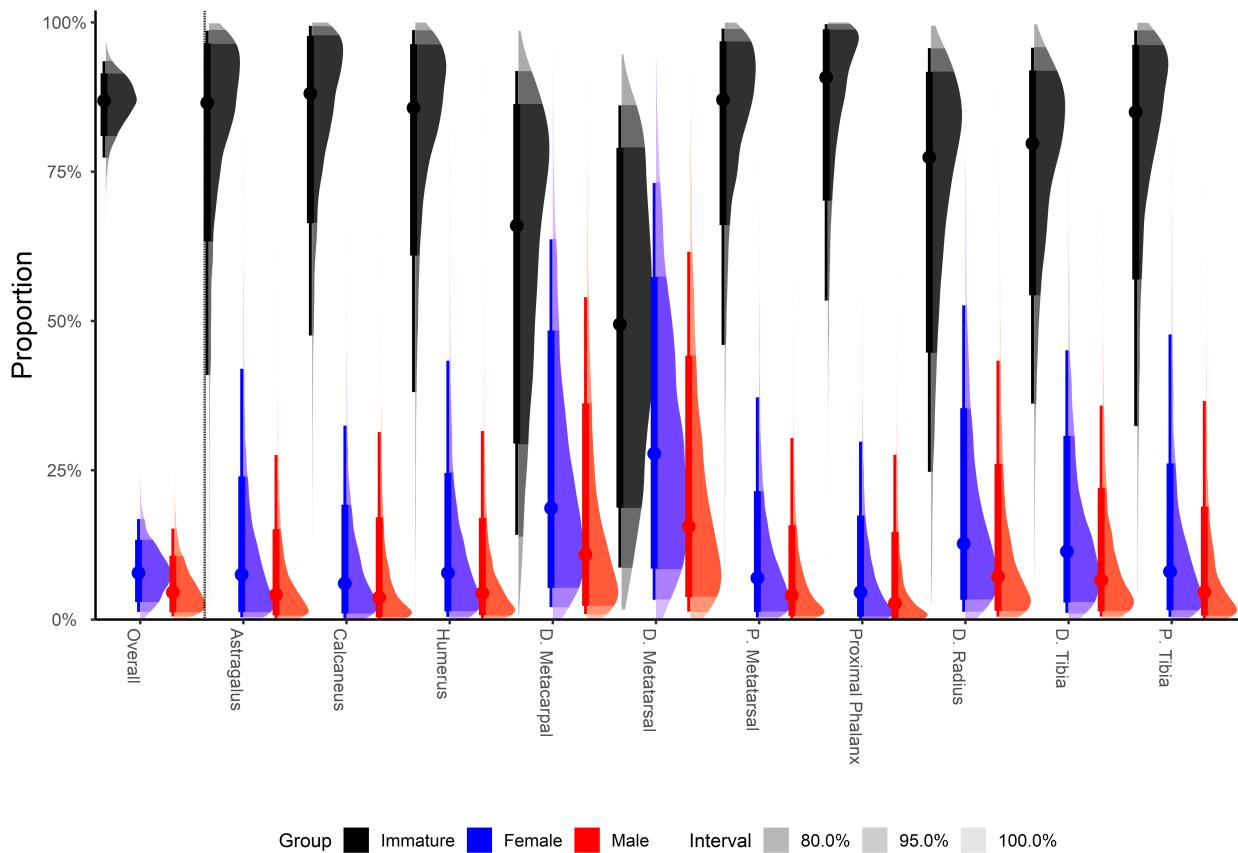


Figure 9: Posterior distributions of composition estimates (π) for the Pinarbaşı B sheep assemblage. The three distributions to the left of the vertical line are the assemblage-level estimates, while the distributions to the right of the line show element-specific composition estimates.

716 The other distributions show element-specific composition estimates for the Pinarbaşı B assemblage.
 717 Most of the element-specific proportions of immature animals (π_1) broadly match the overall estimates, with
 718 posterior medians over 75%. However, these element-specific distributions also have long tails extending into
 719 lower π_1 values, conveying less certainty about element-specific π_1 estimates relative to the assemblage-wide
 720 estimate. This likely owes to small element-specific sample sizes (the astragalus, calcaneus, and proximal
 721 phalanx have 9–10 specimens, all other element-specific samples sizes are 1–4, see Table 6) and to the presence
 722 of some element portions with lower modeled proportions of immature animals. Two element portions—the
 723 distal metacarpal and distal metatarsal—have posterior π_1 median values below 75%, though again have long
 724 tails that extend in both directions (Distal metacarpal: π_1 posterior median = 66%, 95% posterior credible
 725 interval: 46–92%; distal metatarsal: π_1 posterior median = 49%, 95% posterior credible interval: 32–86%).

726 Notably, all measured specimens from these two element portions have fused distal epiphyses, meaning that
 727 the model considers it impossible for the specimens to be immature.

728 Figure 10 shows the same comparison (overall and element-specific distributions) for the average size of
 729 female animals (μ_2) in the Pinarbaşı B assemblage. Average body sizes vary across elements, highlighting
 730 some allometric variation between Pinarbaşı B sheep and the standard sheep and the importance of using
 731 a multilevel model to account for allometric variation. The multilevel structure of the model provides a
 732 parameter ($\sigma_{\text{Element}}[3]$) that directly estimates this variation: the posterior mean of this distribution is 0.05,
 733 with an upper 95% quantile of 0.08 on the LSI_e scale. Combining all the measurements into a single LSI
 734 analysis would confound this element-level variation with differences in composition, muddling the ability to
 735 compare the biometry of Pinarbaşı B sheep with sheep from contemporary assemblages.

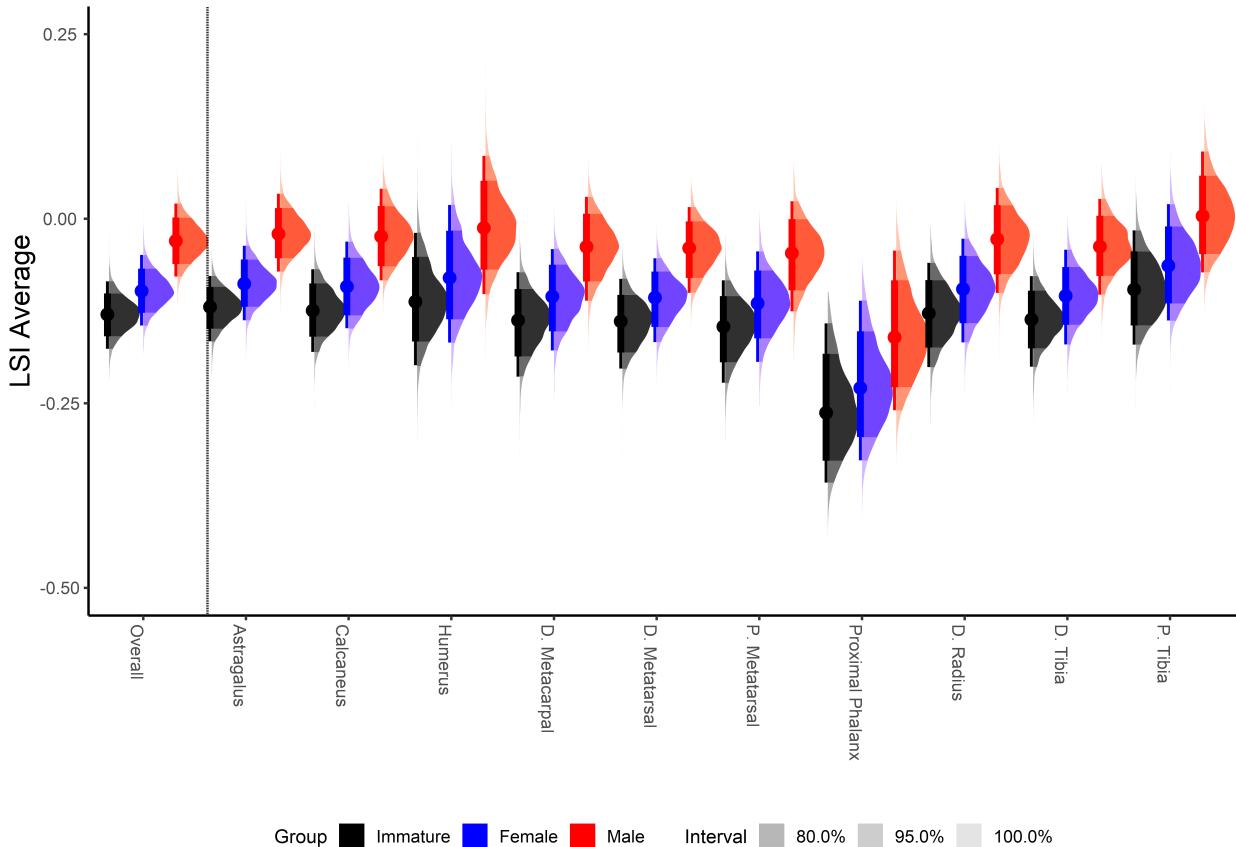


Figure 10: Posterior distributions of average LSI value for female animals (μ_2) for the Pinarbaşı B sheep assemblage. The distribution to the left of the vertical line is the assemblage-level estimate, while the distributions to the right of the line show element-specific size estimates.

736 Figure 11 shows the distribution of simulated compositions for immature, female, and male specimens
 737 in three Pinarbaşı B assemblages: the measured assemblage ($N = 44$), the assemblage of modeled element
 738 portions ($N = 277$), and the full sheep assemblage including five element portions that were not modeled

due to lack of measurements (additional elements: proximal radius, ulna, proximal metacarpal, pelvis, and middle phalanx; total N = 428). Beyond visualizing the group-specific composition of the assemblage, the simulated compositions provide more insight into the assemblage's formation. Five element portions in the full assemblage (astragalus, calcaneus, proximal radius, proximal fused metacarpal bones 3 and 4, pelvis) contain no mature specimens in 18% of the simulations; by contrast, no element portion has more than 2% of the simulations without immature specimens (see Table 9). Additionally, element portions vary in the probability that they contain male specimens: distal metapodials (no males in 8% of simulations) and phalanges (no males in 16% of simulations) contain males much more frequently than other element portions (no males in 39% of simulations; Table 9). While this analysis focuses on appendicular elements, these results suggest that mature animals—particularly mature males—did not enter the assemblage as complete animals but potentially as raw materials for bone tools (e.g., distal metapodials, distal tibia: Russell and Griffitts 2013) or on skins (e.g., proximal and middle phalanges). Immature animals, by contrast, appear to have entered the assemblage as complete animals: the element portions with the highest probability of missing immature specimens may be affected more by density-mediated attrition (e.g., proximal tibia: Symmons 2005) and identifiability of neonatal specimens (e.g., distal metapodials; see Discussion).

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. Metacarpal	0%	0%	3%	10%
D. Metatarsal	1%	0%	3%	6%
Proximal Phalanx	0%	0%	6%	16%
Middle Phalanx	0%	0%	7%	15%
D. Radius	0%	0%	8%	32%
D. Tibia	0%	0%	9%	21%
P. Metatarsal	0%	7%	15%	28%
Ulna	0%	0%	17%	31%
Humerus	0%	0%	19%	35%
P. Metacarpal	0%	11%	25%	37%
P. Radius	0%	13%	27%	40%
P. Tibia	1%	0%	32%	48%
Calcaneus	0%	18%	35%	41%
Pelvis	0%	20%	36%	50%
Astragalus	0%	28%	41%	63%

These results strongly reinforce the interpretation that Pinarbaşı B was used by herders as a camp where sheep gave birth, with on-site culling largely reflecting either the first seasonal cull of animals before winter

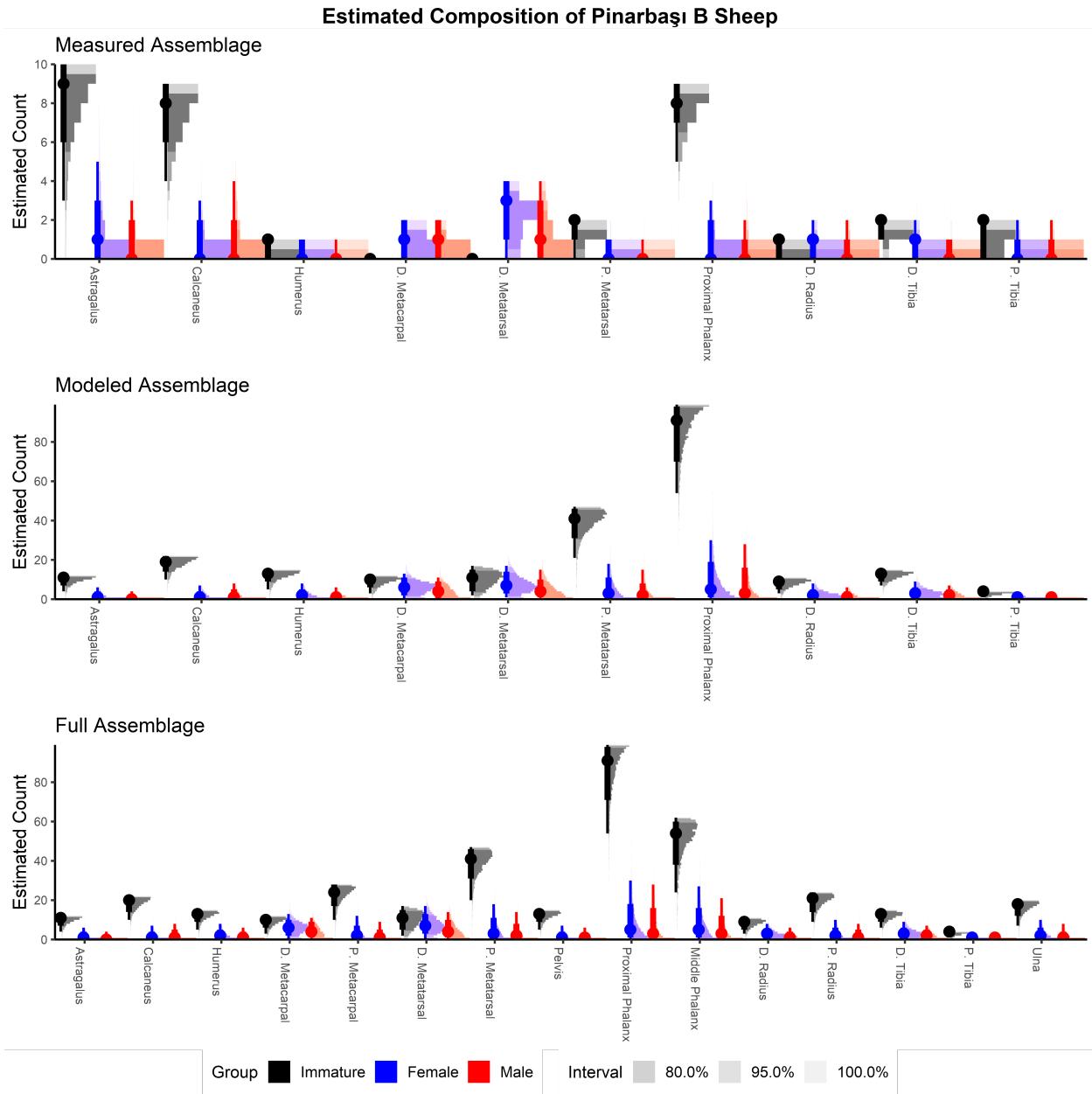


Figure 11: Posterior distributions of simulated group-specific composition for the Pınarbaşı B sheep (top) measured, (middle) modeled, and (bottom) full assemblages. The full assemblage includes five additional element portions compared to the measured and modeled assemblages (proximal metacarpal, pelvis,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

756 or animals that died naturally in their first year of life (Carruthers 2005; Martín et al. 2015). This method
757 reinforces previous analyses, which is not surprising; however, the mixture modeling results go beyond these
758 earlier interpretations by creating group-specific biometric estimates and providing a probabilistic framework
759 to estimate a specimen's membership into the three groups. The biometric estimates not only account for the
760 presence of immature specimens in the assemblage but also for allometric variation across element portions.
761 The probabilistic identifications allow analysts to simulate assemblage compositions, highlighting potential
762 differences in the ways that bones from immature, adult female, and adult male entered the Pinarbaşı B
763 assemblage.

764 **4.3 7th-6th Millennium BCE Northwest Anatolian Cattle: Examining differences between
765 assemblages**

766 Figure 12 shows the posterior distributions of average body sizes for female cattle (μ_2) from the four analyzed
767 assemblages. These distributions are produced from posterior samples; assemblage-specific estimates from a
768 single posterior sample share the same relevant hyper-parameters (μ_{μ_2} and $\sigma_{\text{Site}}[3]$), meaning that they covary
769 with one another to an extent. To compare these distributions, then, a contrast is necessary to account for this
770 potential covariation. This is done by simply evaluating the difference between two parameters (e.g., between
771 the average female LSI_e value μ_2 for Barçın Höyük and μ_2 for Neolithic İlipinar) in each posterior sample,
772 shown in the right-hand panel of Figure 12. These contrasts show that the female cattle from Chalcolithic
773 İlipinar are likely smaller, on average, than female cattle from the other sites. These cattle measurements are
774 3-7% smaller, on average, than those from the other northwestern Anatolian sites relative to the standard
775 animal's measurements. Thus, the mixture modeling approach not only identifies a size difference that was
776 unrecognized when using standard LSI analysis (e.g., Buitenhuis 2008), it also reveals a diachronic trend
777 in cattle body size that can be studied to evaluate factors like climate or herding practices that affected
778 animal body size over time (e.g., Munro, Lebzon, and Sapir-Hen 2022; Wright and Viner-Daniels 2015).
779 Accounting for differences in the elemental and demographic composition of different assemblages provides
780 more accurate reconstructions of body size that allow us to critically examine patterns of animal size change
781 in the past (e.g., Arbuckle and Kassebaum 2021; Manning et al. 2015).

782 In addition to identifying size differences among the assemblages, the modeling results also reveal varia-
783 tion in the age and sex composition of the four assemblages. Figure 13 shows the distributions of assemblage-
784 level demographic variables—the proportion of immature animals and the adult sex ratio (the proportion
785 of adults that are female)—for the four Northwest Anatolian assemblages. The assemblages have broadly
786 similar estimates for the adult sex ratio (right-hand panel of Figure 13), with strong evidence that they
787 contain more females than males. However, modeling shows significant variation in the proportion of imma-

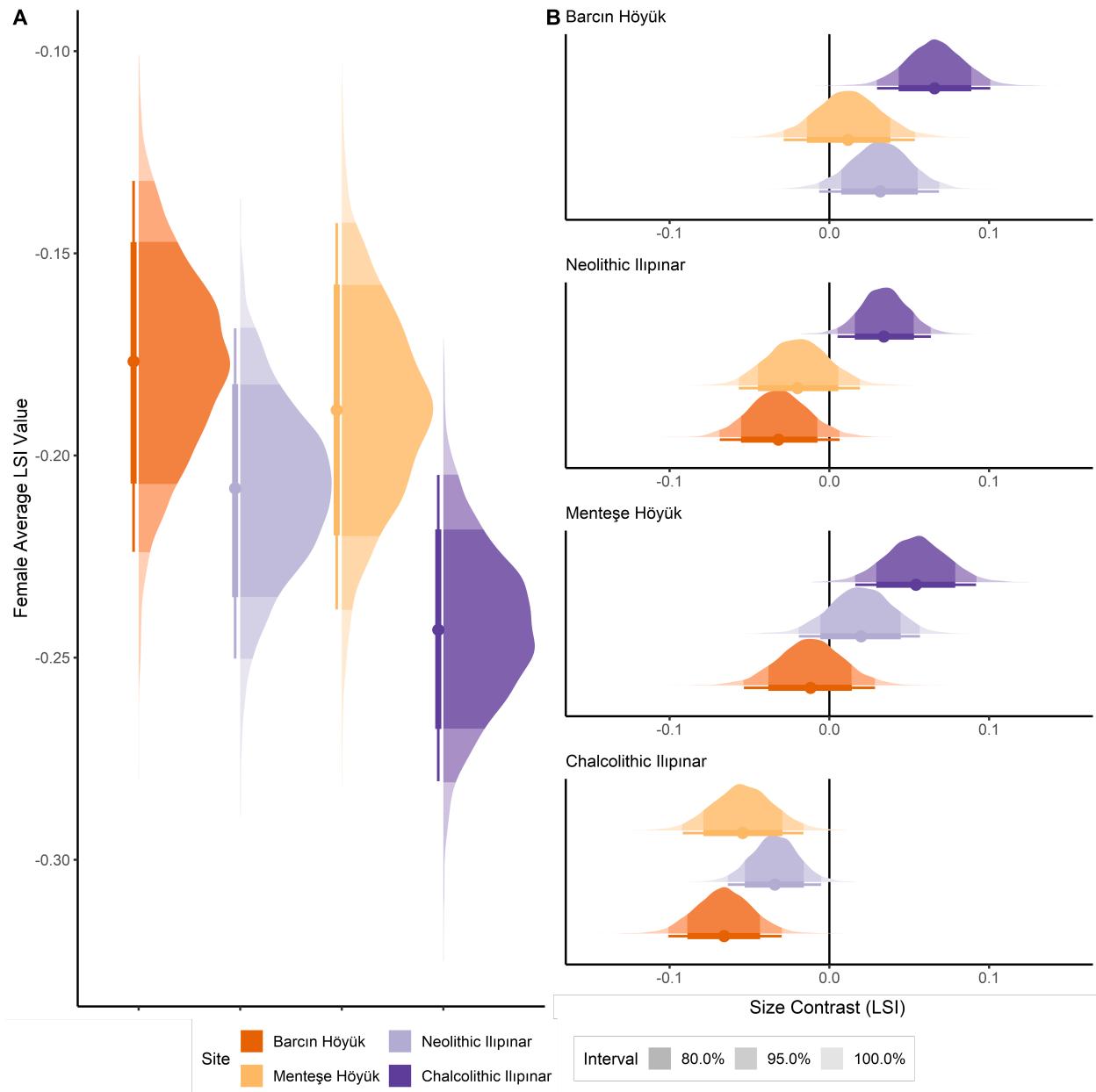


Figure 12: Posterior distributions of site-specific average LSI value for female animals (μ_2) for the Northwest Anatolian cattle assemblages. The left-side panel (A) shows the distributions in broadly temporal order from left to right, while the right-side panel (B) shows site-specific contrasts for 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 Barcin Höyük is significantly larger than Chalcolithic İlipinar).

ture animals in the assemblages (top panel of Figure 13): the proportion of immature cattle in the Menteş Höyük and Barçın Höyük assemblages is significantly higher than the proportions of immature cattle in either İlipinar Höyük assemblage. This difference between the sites could be investigated further to understand whether there are taphonomic factors that have selectively removed immature specimens from the İlipinar Höyük assemblage post-depositionally or whether herd management strategies differed among communities at these sites (e.g., Gillis et al. 2014, 2015).

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

5. DISCUSSION

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

The performance of the Bayesian multilevel mixture models relies on the prior distributions, which provide constraints against overfitting and ensure that the model produces biologically reasonable parameter

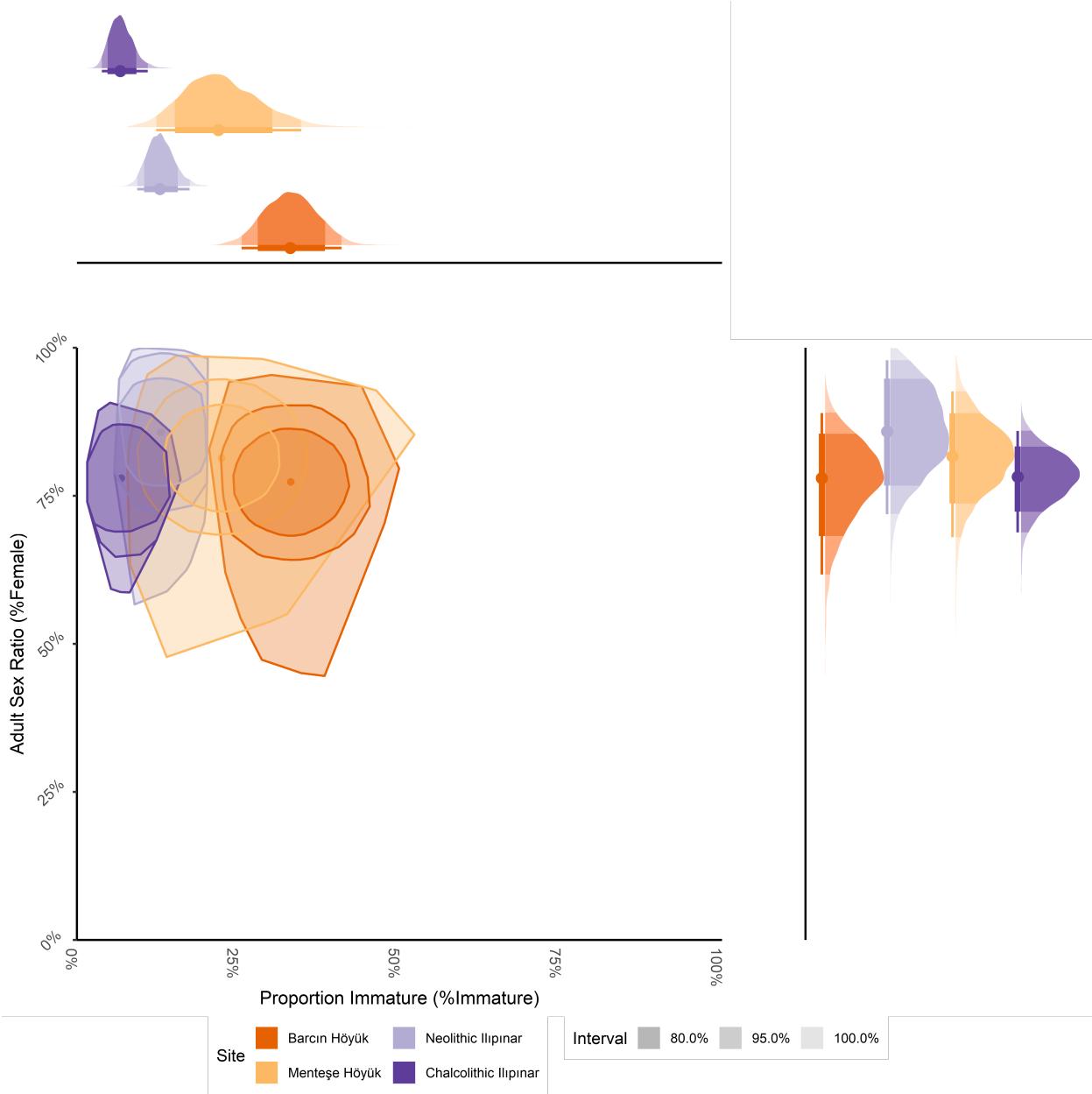


Figure 13: Comparison of posterior distributions of site-specific demographic parameters (proportion of immature π_{Immature} and adult sex ratio θ_{Female}) for the Northwest Anatolian cattle assemblages. Top panel shows the marginal plot for the proportion of immature specimens, while the right-hand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

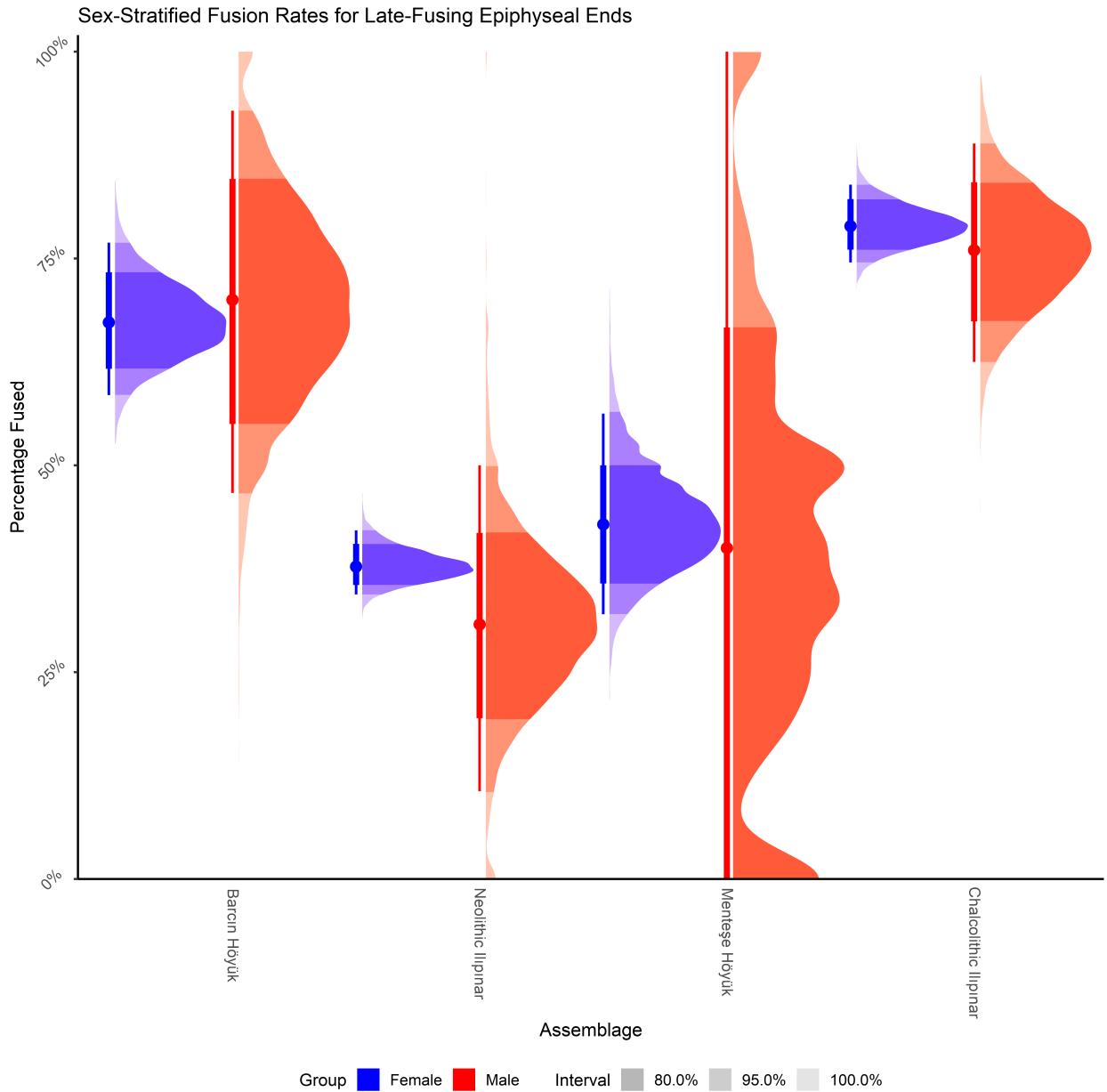


Figure 14: Posterior distributions of simulated sex-specific fusion rates for late-fusing elements among Northwest Anatolian cattle full assemblages. Included element portions are distal tibia, distal metapodials, calcaneus, femur, proximal ulna, distal radius, proximal tibia, and proximal humerus.

estimates. The prior distribution definitions in this paper were derived largely from the measurements of a herd of known-age, known-sex population of Shetland sheep (Popkin et al. 2012), though for the multisite cattle model some of the definitions were changed based on data on European aurochsen (Degerbøl 1970). It is important to note that prior distribution definitions can be derived from many different sources—including quantification based on one's judgment (e.g., Gelman et al. 2008; McCarthy and Masters 2005). This could allow researchers to use different kinds of demographic priors to adjust for their expectations about the ways that assemblages are created, for instance if one may expect catastrophic profiles (e.g., Stiner 1990). More important than the source of one's prior distribution definitions, is investigating the expectations of those prior distribution definitions by performing prior predictive checks as in the Model Supplement (Gabry et al. 2019; Gelman, Carlin, et al. 2020). Further, emphasis should be paid to increasing the diversity of known-age, known-sex animal populations with individual measurement data (e.g., Lebennon and Munro 2022; Zeder and Lemoine 2020), which could help develop prior distributions relevant to different taxa and to understand how variable different parameters, especially size variability (σ) parameters, are across populations.

One of the central tenets of the mixture model's extension to modeled assemblages is the idea that "measurability" (adequate preservation to maintain a biometric dimension) is unrelated to a specimen's status as immature, female, or male. Variation in the mixture proportions among elements, especially the proportion of immature specimens π_1 may highlight group-specific biases in the deposition of specimens but could also indicate issues with the assumption that "measurability" is random. The Pinarbaşı B sheep assemblage potentially demonstrates this issue, as the distal metapodials have much lower element-specific π_1 estimates than other element portions. While metapodial bones from mature sheep could have been selectively over-represented in the assemblage, it is also likely that distal metapodials from immature animals—particularly very young animals—are less likely to be measurable compared to adult animals. Because the distal breadth measurement requires both distal condyles to be present, distal metapodial specimens from neonatal or extremely juvenile individuals may be missed while those from other element portions (e.g., proximal metapodial, distal humerus) would still be theoretically measurable (Martín and García-González 2015). The inclusion of condyle-specific measurements could address this issue, though would require identifying whether the isolated condyle is medial or lateral (e.g., width of condyle: Payne 1969).

The ability to create accurate simulated estimates of age and sex composition provides many opportunities for further analyses. For instance, comparison of the composition of animals in different depositional contexts could support contextual taphonomic analyses (e.g., Meier 2020). Access to certain kinds of animals could highlight systems of provisioning or status-related restrictions (Arbuckle 2012; Twiss 2019: pp. 73-97). Differences in the ontogenetic age and sex composition of different body parts could also highlight ritual

behaviors reflected in the use of certain contexts or sites (e.g., Madgwick and Mulville 2015). In a similar vein, tracking adult sex ratios could identify the use of castrates to take advantage of secondary products like wool or labor; the models used here would identify castrates as males, though alternative prior distributions and measured dimensions could distinguish these groups (e.g., Popkin et al. 2012: Figure 7). On a more practical level, providing specimen-specific probabilities of being immature, female, or male can provide a useful baseline for sampling strategies focused on ancient DNA or stable isotopes, allowing researchers to explore potential sex differences in diets (e.g., Post et al. 2001) or more easily identify male specimens to isolate Y-chromosomal DNA to explore sex-specific selection (e.g., Daly et al. 2021; McGrory et al. 2012).

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

Examining LSI_e (or LSI_{10}) values of Pinarbaşı B sheep without accounting for the impact of immature animals would mislead a researcher into believing that sheep were smaller, on average, than sheep from contemporaneous sites (e.g., late Çatalhöyük: Baird et al. 2011). While stable isotopic analyses have suggested that the Pinarbaşı B and Çatalhöyük sheep had similar diets (e.g., Baird et al. 2011, 2018; Middleton 2018), biometric analyses had not previously supported the idea that these assemblages derived from the same animal population (e.g., Arbuckle et al. 2014: Figure 4). Only eight measurements from the Pinarbaşı B sheep are excluded from being immature due to fusion status, four of which are distal metatarsals; the median LSI_e of these specimens is -0.12, compared to a median of -0.06 for the 362 Late Çatalhöyük sheep that must be mature based on fusion status (Wolfhagen et al. 2021). Even though the mixture model's estimate for the average LSI_e for female sheep (μ_2) at Pinarbaşı B is relatively uncertain (95% credible interval: -0.14 - -0.05 LSI_e), it still provides a useful reference point for comparison that supports the idea that these assemblages have similar biometries (95% credible interval of μ_2 for late Çatalhöyük sheep: -0.11 - -0.06 LSI_e : Wolfhagen et al. 2021).

885 Examining LSI_e (or LSI_{10}) values of Pinarbaşı B sheep without accounting for the impact of immature
886 animals would mislead a researcher into believing that sheep were smaller, on average, than sheep from
887 contemporaneous sites (e.g., late Çatalhöyük: Baird et al. 2011). While stable isotopic analyses have
888 suggested that the Pinarbaşı B and Çatalhöyük sheep had similar diets (e.g., Baird et al. 2011, 2018;
889 Middleton 2018), biometric analyses had not previously supported the idea that these assemblages derived
890 from the same animal population (e.g., Arbuckle et al. 2014: Figure 4). Only eight measurements from the
891 Pinarbaşı B sheep are excluded from being immature due to fusion status, four of which are distal metatarsals;
892 the median LSI_e of these specimens is -0.12, compared to a median of -0.06 for the 362 Late Çatalhöyük
893 sheep that must be mature based on fusion status (Wolfhagen et al. 2021). Even though the mixture
894 model's estimate for the average LSI_e for female sheep (μ_2) at Pinarbaşı B is relatively uncertain (95%
895 credible interval: -0.14 - -0.05 LSI_e), it still provides a useful reference point for comparison that supports
896 the idea that these assemblages have similar biometries (95% credible interval of μ_2 for late Çatalhöyük
897 sheep: -0.11 - -0.06 LSI_e : Wolfhagen et al. 2021).

898 The increased ability to specify the ontogenetic age and sex composition of faunal assemblages with
899 Bayesian multi-level mixture models also highlights the limitations of our current language used to describe
900 and interpret these compositions. Results like the sex-specific fusion biases in Northwest Anatolian cattle
901 complicate straightforward expectations of clear sex-specific patterns in archaeological assemblages, which
902 is perhaps unsurprising. Many discussions that examine changes in the composition of faunal assemblages
903 to identify shifts in exploitation patterns use terms like “prime-dominated age structure” (e.g., Stiner 1990),
904 “dominance of females,” (e.g., Peters, Driesch, and Helmer 2005) or “young male slaughter/kill-off” (e.g.,
905 Zeder and Hesse 2000; Arbuckle and Atici 2013). These terms are deceptive in their utility—they describe
906 some empirical pattern but are ordinal-scale, thus it is up to the individual researcher to define the cut-off
907 between a “dominant” and “non-dominant” assemblage (Wolverton et al. 2016). In the case of the adult
908 sex ratio for Barçın Höyük, 98% of the posterior samples are above 60% (1.5 females:1 male), but only 68%
909 of the posterior samples are over 75% (3 females:1 male). Meanwhile, 79% of the posterior samples for the
910 adult sex ratio for Neolithic İlpınar are over 80% (4 females:1 male). Are both assemblages “dominated
911 by females”? More formalized language in our hypotheses—or, rather, the adoption of statistical modeling
912 frameworks (McElreath 2020: 4-17)—is necessary to clarify what changes in assemblage-level estimates of
913 biometry and composition mean for past human-animal interactions.

914 6. CONCLUSIONS

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

924 Furthermore, the estimates of the age and sex composition of the assemblage can be used to simulate
925 assemblages of specimens with known group assignment (immature, female, and male). These simulations
926 are the baseline for comparing differences in the composition of sub-assemblages. Using these simulations
927 allows researchers to make testable statements about the composition of the assemblage and to directly
928 test hypotheses about differences in the age and sex composition of animal bones from different parts of a
929 site, different fusion groups, or other categories. The Bayesian structure of the model allows researchers the
930 flexibility to create hypotheses that can be tested directly, rather than relying on null hypothesis testing for
931 inference (Otárola-Castillo and Torquato 2018; Otárola-Castillo et al. 2022). Thus, the mixture modeling
932 framework described here provides a foundation for biometric and compositional analyses that operate at
933 multiple scales and presents a new avenue for summarizing and comparing zooarchaeological assemblages.

934 ##ACKNOWLEDGEMENTS

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

944 References

- 945 Arbuckle, Benjamin S. 2012. "Animals and Inequality in Chalcolithic Central Anatolia." Journal Article.
946 *Journal of Anthropological Archaeology* 31 (3): 302–13. <https://doi.org/10.1016/j.jaa.2012.01.008>.
- 947 Arbuckle, Benjamin S., and Levent Atici. 2013. "Initial Diversity in Sheep and Goat Management in
948 Neolithic South-Western Asia." Journal Article. *Levant* 45 (2): 219–35.
- 949 Arbuckle, Benjamin S., Sarah Whitcher Kansa, Eric C. Kansa, David C. Orton, Canan Cakirlar, Lionel
950 Gourichon, A. Levent Atici, et al. 2014. "Data Sharing Reveals Complexity in the Westward Spread of
951 Domestic Animals Across Neolithic Turkey." Journal Article. *PLoS One* 9 (6): 1–11. <https://doi.org/10.1371/journal.pone.0099845>.
- 952
- 953 Arbuckle, Benjamin S., and Theo M. Kassebaum. 2021. "Management and Domestication of Cattle (Bos
954 Taurus) in Neolithic Southwest Asia." Journal Article. *Animal Frontiers* 11 (3): 10–19. <https://doi.org/10.1093/af/vfab015>.
- 955
- 956 Arbuckle, Benjamin S., and V. Özkaya. 2006. "Animal Exploitation at Körtik Tepe: An Early Aceramic
957 Neolithic Site in Southeastern Turkey." Journal Article. *Paléorient* 32 (2): 113–36.
- 958
- 959 Arbuckle, Benjamin S., Max D. Price, Hitomi Hongo, and Banu Öksüz. 2016. "Documenting the Initial
960 Appearance of Domestic Cattle in the Eastern Fertile Crescent (Northern Iraq and Western Iran)." Journal Article. *Journal of Archaeological Science* 72: 1–9. <https://doi.org/10.1016/j.jas.2016.05.008>.
- 961
- 962 Baird, Douglas, Denise Carruthers, Andrew Fairbairn, and Jessica A. Pearson. 2011. "Ritual in the Land-
963 scape: Evidence from Pınarbaşı in the Seventh-Millennium Cal BC Konya Plain." Journal Article. *Antiquity* 85: 380–94.
- 964
- 965 Baird, Douglas, Andrew Fairbairn, Emma Jenkins, Louise Martin, Caroline Middleton, Jessica A. Pearson,
966 Eleni Asouti, et al. 2018. "Agricultural Origins on the Anatolian Plateau." Journal Article. *Proc Natl Acad Sci U S A*. <https://doi.org/10.1073/pnas.1800163115>.
- 967
- 968 Benaglia, Tatiana, Didier Chauveau, David R. Hunter, and Derek Young. 2009. "mixtools: An R Package
969 for Analyzing Finite Mixture Models." *Journal of Statistical Software* 32 (6): 1–29. <http://www.jstatsoft.org/v32/i06/>.
- 970
- 971 Betancourt, Michael. 2017. "A Conceptual Introduction to Hamiltonian Monte Carlo." Journal Article. *arXiv* 1701.02434 [Preprint]: Available from <https://arxiv.org/abs/1701.02434v1>.
- 972
- 973 Breslawski, Ryan P. 2023. "Minimum Animal Units and the Standardized Count Problem." Journal Article. *Journal of Archaeological Method and Theory*. <https://doi.org/10.1007/s10816-022-09563-9>.
- 974
- 975 Breslawski, Ryan P., and David A. Byers. 2015. "Assessing Measurement Error in Paleozoological Osteometrics with Bison Remains." Journal Article. *Journal of Archaeological Science* 53: 235–42. <https://doi.org/10.1016/j.jas.2015.07.011>.

- 976 //doi.org/10.1016/j.jas.2014.10.001.
- 977 Buitenhuis, Hijlke. 2008. "Ilipinar: The Faunal Remains from the Late Neolithic and Early Chalcolithic
978 Levels." Book Section. In *Archaeozoology of the Near East VIII*.
- 979 ———. 2013. "Ilipinar Zooarchaeology." Dataset. <http://opencontext.org/projects/D297CD29-50CA->
980 4B2C-4A07-498ADF3AF487. <https://doi.org/https://doi.org/10.6078/M76H4FB>.
- 981 Çakırlar, Canan. 2013. "Rethinking Neolithic Subsistence at the Gateway to Europe With New Archaeo-
982 zoological Evidence from Istanbul." Book Section. In *Barely Surviving or More Than Enough?: The*
983 *Environmental Archaeology of Subsistence, Specialisation and Surplus Food Production*, edited by Maaike
984 Groot, Daphne Lentjes, and Jørn Zeiler, 59–80. Leiden: Sidestone Press.
- 985 Canty, Angelo, and B. D. Ripley. 2021. *Boot: Bootstrap r (s-Plus) Functions*.
- 986 Carruthers, Denise. 2005. "Hunting and Herding in Central Anatolian Prehistory: The Sites at Pinarbasi." Book Section. In *Archaeozoology of the Near East VI*.
- 987 ———. 2006. "Pınarbaşı 1994: Animal Bones." Dataset. <https://opencontext.org/projects/1677643a-15d0->
988 1b2c-3a35-37f04c765387.
- 989 Corporation, Microsoft, and Steve Weston. 2022. *doParallel: Foreach Parallel Adaptor for the 'Parallel'*
990 *Package*. <https://CRAN.R-project.org/package=doParallel>.
- 991 Corti, Paulo, and David M. Shackleton. 2002. "Relationship Between Predation-Risk Factors and Sexual
992 Segregation in Dall's Sheep (*Ovis Dalli Dalli*). Journal Article. *Canadian Journal of Zoology* 80 (12):
993 2108–17. <https://doi.org/10.1139/z02-207>.
- 994 Crabtree, Pam J. 1990. *West Stow: Early Anglo-Saxon Animal Husbandry*. Book. East Anglian Archaeology.
995 Suffolk: Suffolk County Planning Dept.
- 996 Crema, Enrico R. 2012. "Modelling Temporal Uncertainty in Archaeological Analysis." Journal Article. *Journal of Archaeological Method and Theory* 19 (3): 440–61. <https://doi.org/10.1007/s10816-011-9122->
997 3.
- 998 Dahl, Gudrun, and Anders Hjort. 1976. *Having Herds: Pastoral Herd Growth and Household Economy*.
1000 Book. Stockholm Studies in Social Anthropology. Stockholm: University of Stockholm.
- 1001 Daly, K. G., V. Mattiangeli, A. J. Hare, H. Davoudi, H. Fathi, S. B. Doost, S. Amiri, et al. 2021. "Herded
1002 and Hunted Goat Genomes from the Dawn of Domestication in the Zagros Mountains." Journal Article.
1003 *Proc Natl Acad Sci U S A* 118 (25). <https://doi.org/10.1073/pnas.2100901118>.
- 1004 Davis, Simon J. M. 1981. "The Effects of Temperature Change and Domestication on the Body Size of Late
1005 Pleistocene to Holocene Mammals of Israel." Journal Article. *Paleobiology* 7 (1): 101–14.
- 1006 ———. 1982. "Climatic Change and the Advent of Domestication: The Succession of Ruminant Artiodactyls
1007 in the Late Pleistocene-Holocene in the Israel Region." Journal Article. *Paléorient* 8 (2): 5–15.

- 1009 ———. 1996. “Measurements of a Group of Adult Female Shetland Sheep Skeletons from a Single Flock:
1010 A Baseline for Zooarchaeologists.” Journal Article. *Journal of Archaeological Science* 23: 593–612.
- 1011 ———. 2000. “The Effect of Castration and Age on the Development of the Shetland Sheep Skeleton and
1012 a Metric Comparison Between Bones of Males, Females and Castrates.” Journal Article. *Journal of*
1013 *Archaeological Science* 27 (5): 373–90. <https://doi.org/10.1006/jasc.1999.0452>.
- 1014 Davison, A. C., and D. V. Hinkley. 1997. *Bootstrap Methods and Their Applications*. Cambridge: Cambridge
1015 University Press. <http://statwww.epfl.ch/davison/BMA/>.
- 1016 De Cupere, Bea, Wim Van Neer, Hervé Monchot, Elina Rijmenants, Mircea Udrescu, and Marc Waelkens.
1017 2005. “Ancient Breeds of Domestic Fowl (*Gallus Gallus f. Domestica*) Distinguished on the Basis of
1018 Traditional Observations Combined with Mixture Analysis.” Journal Article. *Journal of Archaeological*
1019 *Science* 32 (11): 1587–97. <https://doi.org/10.1016/j.jas.2005.04.015>.
- 1020 Degerbøl, Magnus. 1970. “The Urus (*Bos Primigenius Bojanus*) and Neolithic Domesticated Cattle (*Bos*
1021 *Taurus Domesticus Linné*) in Denmark.” Journal Article. *Det Kongelige Danske Videnskabernes Selskab*
1022 *Biologiske Skrifter* 17 (1): 5–177.
- 1023 Dong, Zhuan. 1997. “Mixture Analysis and Its Preliminary Application in Archaeology.” Journal Article.
1024 *Journal of Archaeological Science* 24: 141–61.
- 1025 Dowle, Matt, and Arun Srinivasan. 2021. *Data.table: Extension of ‘Data.frame’*. <https://CRAN.R-project.org/package=data.table>.
- 1027 Driesch, Angela von den. 1976. *A Guide to the Measurement of Animal Bones from Archaeological Sites*.
1028 Book. Peabody Museum Bulletins. Cambridge, MA: Harvard University.
- 1029 Evershed, Richard P., Sebastian Payne, Andrew G. Sherratt, Mark S. Copley, Jennifer Coolidge, Duska
1030 Urem-Kotsu, Kostas Kotsakis, et al. 2008. “Earliest Date for Milk Use in the Near East and Southeastern
1031 Europe Linked to Cattle Herding.” Journal Article. *Nature* 455 (7212): 528–31. <https://doi.org/10.1038/nature07180>.
- 1033 Fernández, Hélène, and Hervé Monchot. 2007. “Sexual Dimorphism in Limb Bones of Ibex (*Capra Ibex*
1034 l.): Mixture Analysis Applied to Modern and Fossil Data.” Journal Article. *International Journal of*
1035 *Osteoarchaeology* 17: 479–91. <https://doi.org/10.1002/oa.876>.
- 1036 Fernée, Christianne L., and Konstantinos P. Trimmis. 2021. “Detecting Variability: A Study on the
1037 Application of Bayesian Multilevel Modelling to Archaeological Data. Evidence from the Neolithic Adriatic
1038 and the Bronze Age Aegean.” Journal Article. *Journal of Archaeological Science* 128. <https://doi.org/10.1016/j.jas.2021.105346>.
- 1040 Gabry, Jonah, and Rok Cešnovar. 2022. *Cmdstanr: R Interface to ‘CmdStan’*.
- 1041 Gabry, Jonah, Daniel Simpson, Aki Vehtari, Michael Betancourt, and Andrew Gelman. 2019. “Visualization

- 1042 in Bayesian Workflow.” Journal Article. *Journal of the Royal Statistical Society, Series A* 182 (2): 389–
1043 402.
- 1044 Galik, Alfred. 2013. “Barcın Höyük Zooarchaeology.” Dataset. <http://opencontext.org/projects/74749949-4FD4-4C3E-C830-5AA75703E08E>. [https://doi.org/https://doi.org/10.6078/M78G8HM0](https://doi.org/10.6078/M78G8HM0).
- 1045 Gelman, Andrew. 2006a. “Multilevel (Hierarchical) Modeling: What It Can and Cannot Do.” Journal Article. *Technometrics* 48 (3): 432–35. <https://doi.org/10.1198/004017005000000661>.
- 1046 ———. 2006b. “Prior Distributions for Variance Parameters in Hierarchical Models (Comment on Article by Browne and Draper).” Journal Article. *Bayesian Analysis* 3: 515–34.
- 1047 Gelman, Andrew, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin. 2020.
1048 ———. *Bayesian Data Analysis*. Book. Third Edition.
- 1049 Gelman, Andrew, Aleks Jakulin, Maria Grazia Pittau, and Yu-Sung Su. 2008. “A Weakly Informative
1050 Default Prior Distribution for Logistic and Other Regression Models.” Journal Article. *The Annals of Applied Statistics* 2 (4): 1360–83. <https://doi.org/10.1214/08-aos191>.
- 1051 Gelman, Andrew, Aki Vehtari, Daniel Simpson, Charles C. Margossian, Bob Carpenter, Yuling Yao, Lauren
1052 Kennedy, Jonah Gabry, Paul-Christian Bürkner, and Martin Modrák. 2020. “Bayesian Workflow.”
1053 Journal Article. *ArXiv*.
- 1054 Gerritsen, Fokke, and Rana Özbal. 2019. “Barcın Höyük, a Seventh Millennium Settlement in the Eastern
1055 Marmara Region of Turkey.” Journal Article. *Documenta Praehistorica* 46: 58–67. <https://doi.org/10.4312/dp.46.4>.
- 1056 Gillis, R., R. M. Arbogast, J. F. Piningre, K. Debue, and J. D. Vigne. 2015. “Prediction Models for Age-
1057 at-Death Estimates for Calves, Using Unfused Epiphyses and Diaphyses.” Journal Article. *International
1058 Journal of Osteoarchaeology* 25 (6): 912–22. <https://doi.org/10.1002/oa.2377>.
- 1059 Gillis, R., I. Carrère, M. Saña Seguí, G. Radi, and J. D. Vigne. 2014. “Neonatal Mortality, Young Calf
1060 Slaughter and Milk Production During the Early Neolithic of North Western Mediterranean.” Journal Article.
1061 ———. *International Journal of Osteoarchaeology* 26 (2): 303–13. <https://doi.org/10.1002/oa.2422>.
- 1062 Gourichon, Lionel, and Daniel Helmer. 2008. “Etude de La Faune Neolithique de Menteşe.” Book Section. In
1063 ———. *Life and Death in a Prehistoric Settlement in Northwest Anatolia. The İlipinar Excavations, Volume III*,
1064 edited by Jacob Roodenberg and Songül Alpaslan Roodenberg, 435–48. Leiden: Nederlands Instituut
1065 voor he Nabije Oosten.
- 1066 ———. 2013. “Faunal Data from Neolithic Menteşe.” Dataset. <https://opencontext.org/projects/42ead4db-4a58-9b9b-7ec85266d2a9>. [https://doi.org/https://doi.org/10.6078/M7S46PVN](https://doi.org/10.6078/M7S46PVN).
- 1067 Grau-Sologesto, Idoia, and Umberto Albarella. 2019. “The ‘Long’ Sixteenth Century: A Key Period of
1068 Animal Husbandry Change in England.” Journal Article. *Archaeological and Anthropological Sciences* 11
1069 ———. *Archaeological and Anthropological Sciences* 11 (1): 1–12. <https://doi.org/10.1007/s12523-019-01020-w>.
- 1070 ———. 2020. “The Long Sixteenth Century: A Key Period of Animal Husbandry Change in England.”
1071 Journal Article. *Archaeological and Anthropological Sciences* 11 (1): 1–12. <https://doi.org/10.1007/s12523-019-01020-w>.
- 1072 ———. 2021. “The Long Sixteenth Century: A Key Period of Animal Husbandry Change in England.”
1073 Journal Article. *Archaeological and Anthropological Sciences* 11 (1): 1–12. <https://doi.org/10.1007/s12523-019-01020-w>.
- 1074 ———. 2022. “The Long Sixteenth Century: A Key Period of Animal Husbandry Change in England.”
1075 Journal Article. *Archaeological and Anthropological Sciences* 11 (1): 1–12. <https://doi.org/10.1007/s12523-019-01020-w>.

- 1075 (6): 2781–2803. <https://doi.org/10.1007/s12520-018-0723-6>.
- 1076 Grigson, Caroline. 1982. “Sex and Age Determination of Some Bones and Teeth of Domestic Cattle: A
1077 Review of the Literature.” Book Section. In *Ageing and Sexing Animal Bones from Archaeological Sites*,
1078 edited by Bob Wilson, Caroline Grigson, and Sebastian Payne, 7–23. BAR British Series. Oxford: British
1079 Archaeological Reports.
- 1080 ———. 1989. “Size and Sex: Evidence for the Domestication of Cattle in the Near East.” Book Section.
1081 In *People and Culture in Change: Preceedings of the 2nd Symposium on Upper Paleolithic, Mesolithic,
1082 and Mesolithic Populations of Europe and the Mediterranean Basin*, edited by Israel Hershkovitz, 77–109.
1083 BAR International Series. Oxford: Archaeopress.
- 1084 Hammer, Oyvind. 2013. *PAST PAleontological STatistics Reference Manual*. Book.
- 1085 Helmer, Daniel, Lionel Gourichon, Hervé Monchot, Joris Peters, and Maria Saña Segui. 2005. “Identifying
1086 Early Domestic Cattle from Pre-Pottery Neolithic Sites on the Middle Euphrates Using Sexual Dimor-
1087 phism.” Book Section. In *First Steps of Animal Domestication: New Archaeozoological Approaches*,
1088 edited by Jean-Denis Vigne, Joris Peters, and Daniel Helmer, 86–95. Proceedings of the 9th Conference
1089 of the International Council of Archaeozoology, Durham, August 2002. Oxford: Oxbow Books.
- 1090 Hill, Matthew E., Matthew G. Hill, and Christopher C. Widga. 2008. “Late Quaternary Bison Diminution
1091 on the Great Plains of North America: Evaluating the Role of Human Hunting Versus Climate Change.”
1092 Journal Article. *Quaternary Science Reviews* 27 (17-18): 1752–71. <https://doi.org/10.1016/j.quascirev.2008.07.002>.
- 1093
- 1094 Hongo, Hitomi, Jessica A. Pearson, Banu Öksüz, and Gülcin Ilgezdi. 2009. “The Process of Ungulate
1095 Domestication at Çayönü, Southeastern Turkey: A Multidisciplinary Approach Focusing on Bos Sp.
1096 And Cervus Elaphus.” Journal Article. *Anthropozoologica* 44 (1): 63–78.
- 1097 Jasra, A., C. C. Holmes, and D. A. Stephens. 2005. “Markov Chain Monte Carlo Methods and the Label
1098 Switching Problem in Bayesian Mixture Modeling.” Journal Article. *Statistical Science* 20 (1): 50–67.
1099 <https://doi.org/10.1214/088342305000000016>.
- 1100 Kabukcu, Ceren. 2017. “Woodland Vegetation History and Human Impacts in South-Central Anatolia
1101 16,000–6500 Cal BP: Anthracological Results from Five Prehistoric Sites in the Konya Plain.” Journal
1102 Article. *Quaternary Science Reviews* 176: 85–100. <https://doi.org/10.1016/j.quascirev.2017.10.001>.
- 1103 Karul, Necmi. 2019. “Early Farmers in Northwestern Anatolia in the Seventh Millennium.” Book Section.
1104 In *Concluding the Neolithic: The Near East in the Second Half of the Seventh Millennium BC*, edited by
1105 Arkadiusz Marciniak, 269–86. Atlanta, Georgia: Lockwood Press.
- 1106 Kassambara, Alboukadel. 2020. *Ggpubr: 'Ggplot2' Based Publication Ready Plots*. <https://CRAN.R-project.org/package=ggpubr>.

- 1108 Kay, Matthew. 2022. *ggdist: Visualizations of Distributions and Uncertainty*. <https://doi.org/10.5281/zenodo.3879620>.
- 1109
- 1110 Koch, Paul L. 1986. “Clinal Geographic Variation in Mammals: Implications for the Study of Chronoclines.” Journal Article. *Paleobiology* 12 (3): 269–81.
- 1111
- 1112 Lebenson, Roxanne, and Natalie D. Munro. 2022. “Body Size Variation in a Modern Collection of Mountain Gazelle (*Gazella Gazella*) Skeletons.” Journal Article. *Journal of Archaeological Science: Reports* 41. <https://doi.org/10.1016/j.jasrep.2021.103285>.
- 1113
- 1114
- 1115 Lyman, R. Lee. 2008. *Quantitative Paleoanthropology*. Book. Cambridge Manuals in Archaeology. Cambridge: Cambridge University Press.
- 1116
- 1117 Madgwick, Richard, and Jacqui Mulville. 2015. “Reconstructing Depositional Histories Through Bone Taphonomy: Extending the Potential of Faunal Data.” Journal Article. *Journal of Archaeological Science* 53: 255–63. <https://doi.org/10.1016/j.jas.2014.10.015>.
- 1118
- 1119
- 1120 Manning, Katie, Adrian Timpson, Stephen Shennan, and Enrico Crema. 2015. “Size Reduction in Early European Domestic Cattle Relates to Intensification of Neolithic Herding Strategies.” Journal Article. *PLoS One* 10 (12): e0141873. <https://doi.org/10.1371/journal.pone.0141873>.
- 1121
- 1122
- 1123 Marin, Jean-Michel, Kerrie Mengerson, and Christian P. Robert. 2005. “Bayesian Modelling and Inference on Mixtures of Distributions.” Journal Article. *Handbook of Statistics* 25 (Bayesian Thinking, Modeling and Computation): 459–509.
- 1124
- 1125
- 1126 Marom, Nimrod, and Guy Bar-Oz. 2013. “The Prey Pathway: A Regional History of Cattle (*Bos Taurus*) and Pig (*Sus Scrofa*) Domestication in the Northern Jordan Valley, Israel.” Journal Article. *PLoS One* 8 (2): e55958. <https://doi.org/10.1371/journal.pone.0055958.g001>.
- 1127
- 1128
- 1129 Martín, Patricia, and Ricardo García-González. 2015. “Identifying Sheep (*Ovis Aries*) Fetal Remains in Archaeological Contexts.” Journal Article. *Journal of Archaeological Science* 64: 77–87. <https://doi.org/10.1016/j.jas.2015.10.003>.
- 1130
- 1131
- 1132 Martín, Patricia, Ricardo García-González, Jordi Nadal, and Josep Maria Vergès. 2015. “Perinatal Ovicaprine Remains and Evidence of Shepherding Activities in Early Holocene Enclosure Caves: El Mirador (Sierra de Atapuerca, Spain).” Journal Article. *Quaternary International*. <https://doi.org/10.1016/j.quaint.2015.08.024>.
- 1133
- 1134
- 1135
- 1136 Marwick, Ben. 2017. “Computational Reproducibility in Archaeological Research: Basic Principles and a Case Study of Their Implementation.” Journal Article. *Journal of Archaeological Method and Theory* 24 (2): 424–50. <https://doi.org/10.1007/s10816-015-9272-9>.
- 1137
- 1138
- 1139 Marwick, Ben, and Suzanne E. Pilaar Birch. 2018. “A Standard for the Scholarly Citation of Archaeological Data as an Incentive to Data Sharing.” Journal Article. *Advances in Archaeological Practice* 6 (02):
- 1140

- 1141 125–43. <https://doi.org/10.1017/aap.2018.3>.
- 1142 McCarthy, Michael A., and P. I. P. Masters. 2005. “Profiting from Prior Information in Bayesian Analyses
1143 of Ecological Data.” Journal Article. *Journal of Applied Ecology* 42 (6): 1012–19. <https://doi.org/10.1111/j.1365-2664.2005.01101.x>.
- 1144
- 1145 McElreath, Richard. 2020. *Statistical Rethinking: A Bayesian Course with Examples in r and Stan*. Book.
1146 Texts in Statistical Science. Boca Raton, FL: CRC Press.
- 1147 McGrory, S., E. M. Svensson, A. Götherström, J. Mulville, A. J. Powell, M. J. Collins, and T. P. O’Connor.
1148 2012. “A Novel Method for Integrated Age and Sex Determination from Archaeological Cattle Mandibles.”
1149 Journal Article. *Journal of Archaeological Science* 39 (10): 3324–30. <https://doi.org/10.1016/j.jas.2012.05.021>.
- 1150
- 1151 Meadow, Richard. 1999. “The Use of Size Index Scaling Techniques for Research on Archaeozoological
1152 Collections from the Middle East.” Book Section. In *Historia Animalium Ex Ossibus: Beiträge Zur
1153 Paläoanatomie, Archäologie, Ägyptologie, Ethnologie Und Geschichte Der Tiermedizin*, edited by Cornelia
1154 Becker, Henriette Manhart, Joris Peters, and Jörg Schibler, 285–300. Leidorf: Verlag Marie.
- 1155 Meier, Jacqueline. 2020. “The Contextual Taphonomy of Middens at Neolithic Kfar HaHoresh.” Journal
1156 Article. *Journal of Archaeological Science: Reports* 33. <https://doi.org/10.1016/j.jasrep.2020.102531>.
- 1157 Middleton, Caroline. 2018. “The Beginning of Herding and Animal Management: The Early Development
1158 of Caprine Herding on the Konya Plain, Central Anatolia.” Journal Article. *Anatolian Studies* 68: 1–31.
1159 <https://doi.org/10.1017/s0066154618000017>.
- 1160 Millard, A. R. 2006. “A Bayesian Approach to Ageing Sheep/Goats from Toothwear.” Book Section. In
1161 *Recent Advances in Ageing and Sexing Animal Bones*, edited by Deborah Rusillo, 145–54. Oxford:
1162 Oxbow Books.
- 1163 Milner, J. M., E. B. Nilsen, and H. P. Andreassen. 2007. “Demographic Side Effects of Selective Hunting in
1164 Ungulates and Carnivores.” Journal Article. *Conserv Biol* 21 (1): 36–47. <https://doi.org/10.1111/j.1523-1739.2006.00591.x>.
- 1165
- 1166 Monchot, Hervé, and Jacques Léchelle. 2002. “Statistical Nonparametric Methods for the Study of Fossil
1167 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).
- 1168
- 1169 Monchot, Hervé, Marjan Mashkour, and Jean-Denis Vigne. 2005. “Kernel Smoothing and Mixture Analyses
1170 for the Determination of the Sex Ratios at Death, at the Beginning of the Domestication of Ungulates.”
1171 Book Section. In *First Steps of Animal Domestication: New Archaeozoological Approaches*, edited by
1172 Jean-Denis Vigne, Joris Peters, and Daniel Helmer, 55–60. Oxford: Oxbow Books.
- 1173 Munro, Natalie D., Guy Bar-Oz, and Austin C. Hill. 2011. “An Exploration of Character Traits and Linear

- 1174 Measurements for Sexing Mountain Gazelle (*Gazella Gazella*) Skeletons.” Journal Article. *Journal of*
1175 *Archaeological Science* 38 (6): 1253–65. <https://doi.org/10.1016/j.jas.2011.01.001>.
- 1176 Munro, Natalie D., Roxanne Lebennon, and Lisa Sapir-Hen. 2022. “Revisiting Late Pleistocene-Early
1177 Holocene Mountain Gazelle (*Gazella Gazella*) Body Size Change in the Southern Levant: A Case for
1178 Anthropogenic Impact.” Journal Article. *PLoS One* 17 (8): e0273024. <https://doi.org/10.1371/journal.>
1179 *pone.0273024*.
- 1180 Otárola-Castillo, Erik, and Melissa G. Torquato. 2018. “Bayesian Statistics in Archaeology.” Journal Article.
1181 *Annual Review of Anthropology* 47 (1). <https://doi.org/10.1146/annurev-anthro-102317-045834>.
- 1182 Otárola-Castillo, Erik, Melissa G. Torquato, Jesse Wolfhagen, Matthew E. Hill, and Caitlin E. Buck. 2022.
1183 “Beyond Chronology, Using Bayesian Inference to Evaluate Hypotheses in Archaeology.” Journal Article.
1184 *Advances in Archaeological Practice*.
- 1185 Özdoğan, Mehmet. 2011. “Archaeological Evidence on the Westward Expansion of Farming Communities
1186 from Eastern Anatolia to the Aegean and the Balkans.” Journal Article. *Current Anthropology* 52 (S4):
1187 S415–30. <https://doi.org/10.1086/658895>.
- 1188 ———. 2019. “Early Farmers in Northwestern Turkey: What Is New?” Book Section. In *Concluding
1189 the Neolithic: The Near East in the Second Half of the Seventh Millennium BC*, edited by Arkadiusz
1190 Marciniak, 307–28. Atlanta, Georgia: Lockwood Press.
- 1191 Payne, Sebastian. 1969. “A Metrical Distinction Between Sheep and Goat Metacarpals.” Book Section. In
1192 *The Domestication and Exploitation of Plants and Animals*, edited by Peter J. Ucko and Geoffrey W.
1193 Dimbleby, 295–305. London: Duckworth.
- 1194 ———. 1973. “Kill-Off Patterns in Sheep and Goats: The Mandibles from Aşvan Kale.” Journal Article.
1195 *Anatolian Studies* 23: 281–303.
- 1196 Payne, Sebastian, and Gail Bull. 1988. “Components of Variation in Measurements of Pig Bones and Teeth,
1197 and the Use of Measurements to Distinguish Wild from Domestic Pigs.” Journal Article. *Archaeozoologia*
1198 II (1,2): 27–66.
- 1199 Pebesma, Edzer. 2018. “Simple Features for R: Standardized Support for Spatial Vector Data.” *The R
1200 Journal* 10 (1): 439–46. <https://doi.org/10.32614/RJ-2018-009>.
- 1201 Pérez-Barbería, F. Javier, Iain J. Gordon, and M. Pagel. 2002. “The Origins of Sexual Dimorphism in Body
1202 Size in Ungulates.” Journal Article. *Evolution* 56 (6): 1276–85.
- 1203 Peters, Joris, Angela von den Driesch, and Daniel Helmer. 2005. “The Upper Euphrates-Tigris Basin:
1204 Cradle of Agro-Pastoralism?” Book Section. In *The First Steps of Animal Domestication*, edited by
1205 Jean-Denis Vigne, Joris Peters, and Daniel Helmer, 96–124. Proceedings of the 9th Conference of the
1206 International Council of Archaeozoology, Durham, August 2002. Oxford: Oxbow Books.

- 1207 Popkin, Peter R. W., Polydora Baker, Fay Worley, Sebastian Payne, and Andy Hammon. 2012. "The Sheep
1208 Project (1): Determining Skeletal Growth, Timing of Epiphyseal Fusion and Morphometric Variation
1209 in Unimproved Shetland Sheep of Known Age, Sex, Castration Status and Nutrition." Journal Article.
1210 *Journal of Archaeological Science* 39 (6): 1775–92. <https://doi.org/10.1016/j.jas.2012.01.018>.
- 1211 Post, Diane M., Trent S. Armbrust, Eva A. Horne, and Jacob R. Goheen. 2001. "Sexual Segregation
1212 Results in Differences in Content and Quality of Bison (*Bos Bison*) Diets." Journal Article. *Journal of*
1213 *Mammalogy* 82 (2): 407–13.
- 1214 Pozo, Jose M, Valenzuela-Lamas, Silvia, Trentacoste, Angela, Nieto-Espinet, Ariadna, Guimarães Chiarelli,
1215 and Silvia. 2021. "Zoolog: Zooarchaeological Analysis with Log-Ratios." *Journal of Statistical Software*
1216 In preparation. <https://josempozo.github.io/zoolog/>.
- 1217 Proaktor, G., T. Coulson, and E. J. Milner-Gulland. 2007. "Evolutionary Responses to Harvesting in
1218 Ungulates." Journal Article. *J Anim Ecol* 76 (4): 669–78. <https://doi.org/10.1111/j.1365-2656.2007.01244.x>.
- 1220 R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R
1221 Foundation for Statistical Computing. <https://www.R-project.org/>.
- 1222 Ram, Karthik, and Ben Marwick. 2018. "Building Toward a Future Where Reproducible, Open Science Is
1223 the Norm." Book Section. In *The Practice of Reproducible Research: Case Studies and Lessons from the*
1224 *Data-Intensive Sciences*, edited by Justin Kitzes, Daniel Turek, and Fatma Deniz, 69–78. Oakland, CA:
1225 University of California Press.
- 1226 Redding, Richard W. 1984. "Theoretical Determinants of a Herder's Decisions: Modeling Variation in the
1227 Sheep/Goat Ratio." Book Section. In *Animals and Archaeology*, edited by Juliet Clutton-Brock and
1228 Caroline Grigson, 3. Early Herders and their Flocks:223–41. BAR International Series. Oxford: British
1229 Archaeological Reports.
- 1230 Roodenberg, Jacob. 2012a. "Change in Food Production and Its Impact on an Early 6th Millennium
1231 Community in Northwest Anatolia. The Example of İlipinar." Journal Article. *Praehistorische Zeitschrift*
1232 87 (2). <https://doi.org/10.1515/pz-2012-0015>.
- 1233 ———. 2012b. "İlipinar: A Neolithic Settlement in the Eastern Marmara Region." Book Section. In
1234 *The Oxford Handbook of Ancient Anatolia (10,000-323 BCE)*, edited by Gregory McMahon and Sharon
1235 Steadman. <https://doi.org/10.1093/oxfordhb/9780195376142.013.0044>.
- 1236 Roodenberg, Jacob, A. van As, L. Jacobs, and M. H. Wijnen. 2003. "Early Settlement in the Plain of
1237 Yenisehir (NW Anatolia): The Basal Occupation Layers at Mentese." Journal Article. *Anatolica* XXIX
1238 (1): 18–58.
- 1239 Rosenthal, Robert. 1979. "The "File Drawer Problem" and Tolerance for Null Results." Journal Article.

- 1240 *Psychological Bulletin* 86 (3): 638–41.
- 1241 RStudio Team. 2022. *RStudio: Integrated Development Environment for r*. Boston, MA: RStudio, PBC.
1242 <http://www.rstudio.com/>.
- 1243 Ruckstuhl, K. E. 2007. “Sexual Segregation in Vertebrates: Proximate and Ultimate Causes.” Journal
1244 Article. *Integr Comp Biol* 47 (2): 245–57. <https://doi.org/10.1093/icb/icm030>.
- 1245 Ruckstuhl, K. E., and P. Neuhaus. 2002. “Sexual Segregation in Ungulates: A Comparative Test of
1246 Three Hypotheses.” Journal Article. *Biological Reviews* 77 (1): 77–96. <https://doi.org/10.1017/s1464793101005814>.
- 1248 Russell, Nerissa, and Janet L. Griffitts. 2013. “Catalhoyuk Worked Bones: South and 4040 Areas.” Book
1249 Section. In *Substantive Technologies at Catalhoyuk: Reports from the 2000-2008 Seasons*, edited by Ian
1250 Hodder, 277–306. Los Angeles.
- 1251 Said, Sonia, Vincent Tolon, Serge Brandt, and Eric Baubet. 2011. “Sex Effect on Habitat Selection in
1252 Response to Hunting Disturbance: The Study of Wild Boar.” Journal Article. *European Journal of
1253 Wildlife Research* 58 (1): 107–15. <https://doi.org/10.1007/s10344-011-0548-4>.
- 1254 Sasson, Aharon, and Susan Arter. 2020. “Earliest Utilization of Chicken in Upper California: The Zooar-
1255 chaeology of Avian Remains from the San Diego Royal Presidio.” Journal Article. *American Antiquity*
1256 85 (3): 516–34. <https://doi.org/10.1017/aaq.2020.27>.
- 1257 Scrucca, Luca, Michael Fop, T. Brendan Murphy, and Adrian E. Raftery. 2016. “mclust 5: Clustering,
1258 Classification and Density Estimation Using Gaussian Finite Mixture Models.” *The R Journal* 8 (1):
1259 289–317. <https://doi.org/10.32614/RJ-2016-021>.
- 1260 Slowikowski, Kamil. 2021. *Ggrepel: Automatically Position Non-Overlapping Text Labels with 'Ggplot2'*.
1261 <https://CRAN.R-project.org/package=ggrepel>.
- 1262 Smith, A. F. M., and Alan E. Gelfand. 1992. “Bayesian Statistics Without Tears: A Sampling-Resampling
1263 Perspective.” Journal Article. *The American Statistician* 46 (2): 84–88.
- 1264 South, Andy. 2017a. *Rnaturalearth: World Map Data from Natural Earth*. <https://CRAN.R-project.org/package=rnaturalearth>.
- 1266 ———. 2017b. *Rnaturalearthdata: World Vector Map Data from Natural Earth Used in 'Rnaturalearth'*.
1267 <https://CRAN.R-project.org/package=rnaturalearthdata>.
- 1268 Speth, John D. 1983. *Bison Kills and Bone Counts: Decision Making by Ancient Hunters*. Book. Chicago:
1269 University of Chicago Press.
- 1270 ———. 2013. “Thoughts about Hunting: Some Things We Know and Some Things We Don’t Know.”
1271 Journal Article. *Quaternary International* 297: 176–85. <https://doi.org/10.1016/j.quaint.2012.12.005>.
- 1272 Stan Development Team. 2021. “RStan: The R Interface to Stan.” <https://mc-stan.org/>.

- 1273 Stiner, Mary C. 1990. "The Use of Mortality Patterns in Archaeological Studies of Hominid Predatory
1274 Adaptations." Journal Article. *Journal of Anthropological Archaeology* 9: 305–51.
- 1275 Stiner, Mary C., Natalie D. Munro, Hilke Buitenhuis, G. Duru, and M. Ozbasaran. 2022. "An Endemic
1276 Pathway to Sheep and Goat Domestication at Asikli Hoyuk (Central Anatolia, Turkey)." Journal Article.
1277 *Proc Natl Acad Sci U S A* 119 (4). <https://doi.org/10.1073/pnas.2110930119>.
- 1278 Symmons, Robert. 2005. "New Density Data for Unfused and Fused Sheep Bones, and a Preliminary
1279 Discussion on the Modelling of Taphonomic Bias in Archaeofaunal Age Profiles." Journal Article. *Journal
1280 of Archaeological Science* 32 (11): 1691–98. <https://doi.org/10.1016/j.jas.2005.05.011>.
- 1281 Tchernov, Eitan, and Liora Kolska Horwitz. 1991. "Body Size Diminution Under Domestication: Uncon-
1282 scious Selection in Primeval Domesticates." Journal Article. *Journal of Anthropological Archaeology* 10:
1283 54–75.
- 1284 Team, Stan Development. 2022. "Stan Modeling Language Users Guide and Reference Manual, Version
1285 2.29." Generic. <https://mc-stan.org>.
- 1286 Thissen, Laurens, Hadi Özbal, Ayla Türkkul Büyik, Fokke Gerritsen, and Rana Özbal. 2010. "The Land
1287 of Milk? Approaching Dietary Preferences of Late Neolithic Communities in NW Anatolia." Journal
1288 Article. *Leiden Journal of Pottery Studies* 26: 157–72.
- 1289 Todd, Lawrence C. 1983. "The Horner Site: Taphonomy of an Early Holocene Bison Bonebed." Thesis.
- 1290 Trentacoste, A., A. Nieto-Espinet, S. Guimaraes, B. Wilkens, G. Petrucci, and S. Valenzuela-Lamas. 2021.
1291 "New Trajectories or Accelerating Change? Zooarchaeological Evidence for Roman Transformation of
1292 Animal Husbandry in Northern Italy." Journal Article. *Archaeol Anthropol Sci* 13 (1): 25. <https://doi.org/10.1007/s12520-020-01251-7>.
- 1294 Twiss, Katheryn C. 2019. *The Archaeology of Food*. Book. <https://doi.org/10.1017/9781108670159>.
- 1295 Twiss, Katheryn C., and Nerissa Russell. 2009. "Taking the Bull by the Horns: Ideology, Masculinity, and
1296 Cattle Horns at atalhöyük (Turkey)." Journal Article. *Paléorient* 35 (2): 19–32.
- 1297 Uerpmann, Margarethe, and Hans-Peter Uerpmann. 1994. "Animal Bones." Book Section. In *Qala'at Al-
1298 Bahrain*, edited by Flemming Hojlund and H. Hellmuth Andersen, 1: The Northern City Wall and the
1299 Islamic Fortress:417–54. Jutland Archaeological Society Publications. Aarhus: Aarhus University Press.
- 1300 Urbanek, Simon, and Jeffrey Horner. 2022. *Cairo: R Graphics Device Using Cairo Graphics Library for
1301 Creating High-Quality Bitmap (PNG, JPEG, TIFF), Vector (PDF, SVG, PostScript) and Display (X11
1302 and Win32) Output*. <https://CRAN.R-project.org/package=Cairo>.
- 1303 Vigne, Jean-Denis. 2011. "Le Mouton (Ovis Aries)." Book Section. In *Shillourokambos: Un Établissement
1304 Néolithique Pré-Céramique à Chypre: Les Fouilles Du Secteur i*, edited by Jean Guilaine, François Briois,
1305 and Jean-Denis Vigne, 1021–38. Paris: Editions Errance.

- 1306 Watkins, Trevor. 1996. "Excavations at Pinarbaşı: The Early Stages." Book Section. In *On the Surface: Çatalhöyük 1993-1995*, edited by Ian Hodder, 47–57. Cambridge: McDonald Institute for Archaeological Research.
- 1309 Weinstock, Jacobo. 2006. "Environment, Body Size and Sexual Dimorphism in Late Glacial Reindeer." Book Section. In *Recent Advances in Ageing and Sexing Animal Bones*, edited by Deborah Ruscillo, 247–53. Oxford: Oxbow Books.
- 1312 Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>.
- 1314 Wickham, Hadley, and Jennifer Bryan. 2022. *Readxl: Read Excel Files*. <https://CRAN.R-project.org/package=readxl>.
- 1316 Widga, Chris. 2006. "Niche Variability in Late Holocene Bison: A Perspective from Big Bone Lick, KY." Journal Article. *Journal of Archaeological Science* 33 (9): 1237–55. <https://doi.org/10.1016/j.jas.2005.12.011>.
- 1319 Wolfhagen, Jesse. 2020. "Re-Examining the Use of the LSI Technique in Zooarchaeology." Journal Article. *Journal of Archaeological Science* 123: 105254. <https://doi.org/10.1016/j.jas.2020.105254>.
- 1321 Wolfhagen, Jesse, and Max D. Price. 2017. "A Probabilistic Model for Distinguishing Between Sheep and Goat Postcranial Remains." Journal Article. *Journal of Archaeological Science: Reports* 12: 625–31. <https://doi.org/10.1016/j.jasrep.2017.02.022>.
- 1324 Wolfhagen, Jesse, Katheryn C. Twiss, Jacqui A. Mulville, and G. Arzu Demirergi. 2021. "Examining Caprine Management and Cattle Domestication Through Biometric Analyses at Çatalhöyük East (North and South Areas)." Book Section. In *Peopling the Landscape of Çatalhöyük: Reports from the 2009-2017 Seasons*, edited by Ian Hodder, 181–98. Los Angeles: Cotsen Institute of Archaeology Press.
- 1328 Wolverton, Steve. 2008. "Harvest Pressure and Environmental Carrying Capacity: An Ordinal-Scale Model of Effects on Ungulate Prey." Journal Article. *American Antiquity* 73 (2): 179–99.
- 1330 Wright, Elizabeth, and Sarah Viner-Daniels. 2015. "Geographical Variation in the Size and Shape of the European Aurochs (*Bos Primigenius*).". Journal Article. *Journal of Archaeological Science* 54: 8–22. <https://doi.org/10.1016/j.jas.2014.11.021>.
- 1333 Xie, Yihui. 2015. *Dynamic Documents with R and Knitr*. 2nd ed. Boca Raton, Florida: Chapman; Hall/CRC. <https://yihui.org/knitr/>.
- 1335 ———. 2022. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. <https://yihui.org/knitr/>.
- 1337 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,

- 1339 227–59. Cambridge: Cambridge University Press.
- 1340 Zeder, Melinda A., and Brian Hesse. 2000. “The Initial Domestication of Goats (*Capra Hircus*) in the Zagros
1341 Mountains 10,000 Years Ago.” Journal Article. *Science* 287 (5461): 2254–57.
- 1342 Zeder, Melinda A., and Ximena Lemoine. 2020. “A Method for Constructing Demographic Profiles in Sus
1343 Scrofa Using Logarithm Size Index Scaling.” Journal Article. *Journal of Archaeological Science* 116.
1344 <https://doi.org/10.1016/j.jas.2020.105115>.
- 1345 Zhu, Hao. 2021. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. <https://CRAN.R-project.org/package=kableExtra>.
- 1346 Zohary, D., Eitan Tchernov, and Liora Kolska Horwitz. 1998. “The Role of Unconscious Selection in the
1347 Domestication of Sheep and Goats.” Journal Article. *Journal of Zoology, London* 245: 129–35.

1349 **Appendix 1 (Supplemental Table 6)**

1350 ***Posterior Summary Tables for Overall and Site-Level Model Parameters: Simulated Assem-***

1351 ***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 θ_1	0.16	0.15	0.04	0.10	0.23	1	6049	3074
Single Assemblage Model θ_2	0.55	0.55	0.09	0.40	0.69	1	5087	2939
Single Assemblage Model θ_3	0.29	0.29	0.09	0.16	0.45	1	4702	2688
Single Assemblage Model μ_1	-0.10	-0.10	0.02	-0.12	-0.07	1	2580	2348
Single Assemblage Model μ_2	-0.08	-0.08	0.01	-0.10	-0.05	1	2367	2360
Single Assemblage Model μ_3	-0.01	-0.01	0.02	-0.04	0.01	1	2550	2714
Single Assemblage Model σ_1	0.05	0.05	0.00	0.04	0.06	1	6153	3082
Single Assemblage Model σ_2	0.04	0.04	0.00	0.04	0.05	1	5620	3156
Single Assemblage Model σ_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 θ_1	0.16	0.16	0.02	0.12	0.19	1	4144	3323
Multisite Model θ_2	0.47	0.47	0.05	0.38	0.56	1	3529	3180
Multisite Model θ_3	0.37	0.37	0.06	0.28	0.47	1	3246	3407
Multisite Model μ_1	-0.10	-0.10	0.04	-0.16	-0.04	1	1539	2146
Multisite Model μ_2	-0.09	-0.09	0.04	-0.15	-0.02	1	1523	2236
Multisite Model μ_3	-0.01	-0.01	0.04	-0.07	0.06	1	1574	2062
Multisite Model σ_1	0.05	0.05	0.00	0.04	0.05	1	7355	3379
Multisite Model σ_2	0.04	0.04	0.00	0.04	0.05	1	5039	3218
Multisite Model σ_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 θ_1	0.16	0.16	0.04	0.11	0.23	1	5304	3451
Site 2 θ_1	0.49	0.50	0.10	0.33	0.66	1	5358	3448
Site 3 θ_1	0.34	0.34	0.11	0.17	0.52	1	5653	3503
Site 4 θ_1	0.17	0.17	0.04	0.11	0.25	1	5950	3302
Site 5 θ_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 θ_1	0.38	0.37	0.11	0.20	0.58	1	5416	3485
Site 7 θ_1	0.14	0.14	0.03	0.09	0.20	1	2975	3311
Site 8 θ_1	0.34	0.34	0.09	0.18	0.49	1	4815	3327
Site 9 θ_1	0.52	0.52	0.09	0.37	0.68	1	4927	3496
Site 10 θ_1	0.18	0.17	0.04	0.12	0.26	1	3525	3322
Site 11 θ_1	0.46	0.47	0.10	0.29	0.62	1	4115	2988
Site 12 θ_1	0.36	0.35	0.10	0.19	0.54	1	5188	3581
Site 13 θ_1	0.16	0.16	0.04	0.10	0.22	1	3174	3437
Site 14 θ_1	0.47	0.47	0.07	0.34	0.59	1	4010	3347
Site 15 θ_1	0.38	0.38	0.07	0.26	0.50	1	4407	3581
Site 1 θ_2	0.16	0.16	0.04	0.11	0.22	1	5425	3617
Site 2 θ_2	0.59	0.59	0.10	0.41	0.73	1	4464	2976
Site 3 θ_2	0.25	0.24	0.10	0.10	0.43	1	4712	2977
Site 4 θ_2	0.16	0.16	0.04	0.10	0.24	1	6403	3254
Site 5 θ_2	0.54	0.55	0.11	0.35	0.71	1	4958	3252
Site 6 θ_2	0.30	0.29	0.12	0.12	0.51	1	5537	3137
Site 7 θ_2	0.16	0.16	0.04	0.10	0.22	1	4024	3302
Site 8 θ_2	0.53	0.53	0.10	0.35	0.69	1	3380	3491
Site 9 θ_2	0.32	0.31	0.10	0.16	0.50	1	3935	3376
Site 10 θ_2	0.19	0.19	0.04	0.13	0.27	1	2665	3226
Site 11 θ_2	0.57	0.57	0.09	0.41	0.71	1	4313	3469
Site 12 θ_2	0.24	0.24	0.09	0.10	0.41	1	5093	3552
Site 13 θ_2	0.25	0.24	0.07	0.16	0.37	1	1556	2133
Site 14 θ_2	0.61	0.62	0.08	0.48	0.73	1	2398	3166
Site 15 θ_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 θ_3	0.14	0.14	0.04	0.08	0.19	1	3812	3688
Site 2 θ_3	0.46	0.47	0.11	0.29	0.63	1	5744	3098
Site 3 θ_3	0.40	0.39	0.11	0.22	0.58	1	6001	3199
Site 4 θ_3	0.13	0.13	0.04	0.06	0.19	1	2707	3291
Site 5 θ_3	0.38	0.38	0.12	0.18	0.57	1	5785	3678
Site 6 θ_3	0.49	0.49	0.13	0.30	0.71	1	5195	3418
Site 7 θ_3	0.14	0.14	0.03	0.08	0.19	1	2906	3264
Site 8 θ_3	0.33	0.33	0.11	0.15	0.52	1	4235	3443
Site 9 θ_3	0.53	0.53	0.12	0.34	0.72	1	4324	3168
Site 10 θ_3	0.15	0.15	0.04	0.09	0.21	1	3329	3058
Site 11 θ_3	0.38	0.38	0.10	0.21	0.55	1	4363	3451
Site 12 θ_3	0.47	0.48	0.11	0.30	0.65	1	4804	2672
Site 13 θ_3	0.14	0.14	0.04	0.08	0.20	1	2801	3183
Site 14 θ_3	0.31	0.31	0.07	0.19	0.43	1	3858	3683
Site 15 θ_3	0.55	0.55	0.08	0.43	0.68	1	4182	3722
Site 1 μ_1	-0.09	-0.08	0.02	-0.12	-0.05	1	3147	3290
Site 2 μ_1	-0.09	-0.09	0.03	-0.13	-0.04	1	4013	3495
Site 3 μ_1	0.08	0.08	0.02	0.05	0.11	1	3571	3474
Site 4 μ_1	-0.30	-0.30	0.02	-0.34	-0.27	1	3258	3185
Site 5 μ_1	-0.10	-0.10	0.02	-0.13	-0.07	1	3723	3459
Site 6 μ_1	-0.12	-0.12	0.02	-0.15	-0.10	1	3439	3181
Site 7 μ_1	-0.09	-0.09	0.02	-0.13	-0.05	1	3748	3231
Site 8 μ_1	0.07	0.07	0.02	0.04	0.10	1	3605	3549
Site 9 μ_1	-0.34	-0.34	0.02	-0.37	-0.32	1	2618	3402
Site 10 μ_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 μ_1	-0.08	-0.08	0.02	-0.11	-0.05	1	3756	3689
Site 12 μ_1	-0.09	-0.09	0.03	-0.14	-0.05	1	4229	3256
Site 13 μ_1	0.09	0.09	0.02	0.05	0.12	1	3577	3764
Site 14 μ_1	-0.32	-0.32	0.02	-0.36	-0.29	1	3862	3616
Site 15 μ_1	-0.09	-0.09	0.02	-0.12	-0.06	1	3308	3433
Site 1 μ_2	-0.07	-0.07	0.02	-0.10	-0.04	1	3139	3517
Site 2 μ_2	-0.07	-0.07	0.03	-0.12	-0.03	1	4132	3698
Site 3 μ_2	0.10	0.10	0.02	0.06	0.13	1	3566	3580
Site 4 μ_2	-0.29	-0.29	0.02	-0.32	-0.26	1	3144	3192
Site 5 μ_2	-0.08	-0.08	0.02	-0.11	-0.06	1	3635	3549
Site 6 μ_2	-0.11	-0.11	0.02	-0.14	-0.08	1	3292	3486
Site 7 μ_2	-0.07	-0.07	0.02	-0.11	-0.04	1	3633	3061
Site 8 μ_2	0.08	0.09	0.02	0.06	0.11	1	3554	3601
Site 9 μ_2	-0.33	-0.33	0.02	-0.36	-0.30	1	2394	3112
Site 10 μ_2	-0.09	-0.09	0.02	-0.11	-0.06	1	2887	3457
Site 11 μ_2	-0.06	-0.06	0.02	-0.09	-0.03	1	3668	3467
Site 12 μ_2	-0.07	-0.07	0.03	-0.12	-0.03	1	4224	3519
Site 13 μ_2	0.10	0.10	0.02	0.07	0.13	1	3588	3845
Site 14 μ_2	-0.31	-0.31	0.02	-0.34	-0.27	1	3825	3615
Site 15 μ_2	-0.07	-0.07	0.02	-0.10	-0.04	1	3054	3549
Site 1 μ_3	-0.01	-0.01	0.02	-0.04	0.03	1	3005	3538
Site 2 μ_3	0.02	0.02	0.03	-0.02	0.07	1	3982	3477
Site 3 μ_3	0.20	0.20	0.02	0.16	0.23	1	3683	3643
Site 4 μ_3	-0.22	-0.23	0.02	-0.26	-0.19	1	3666	3573
Site 5 μ_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 μ_3	-0.05	-0.06	0.03	-0.09	0.00	1	3589	3281
Site 7 μ_3	0.00	-0.01	0.04	-0.05	0.06	1	4161	3492
Site 8 μ_3	0.15	0.15	0.03	0.11	0.20	1	3669	3493
Site 9 μ_3	-0.26	-0.27	0.03	-0.30	-0.22	1	3611	3648
Site 10 μ_3	0.14	0.14	0.04	0.07	0.20	1	3534	2869
Site 11 μ_3	0.00	0.00	0.02	-0.04	0.03	1	4004	3837
Site 12 μ_3	0.00	0.00	0.03	-0.04	0.05	1	4345	3491
Site 13 μ_3	0.16	0.16	0.02	0.13	0.19	1	3543	3362
Site 14 μ_3	-0.23	-0.23	0.02	-0.27	-0.20	1	3913	3876
Site 15 μ_3	0.17	0.17	0.02	0.14	0.20	1	3493	3140
Site 1 σ_1	0.05	0.05	0.00	0.04	0.05	1	6814	3498
Site 2 σ_1	0.05	0.05	0.00	0.04	0.05	1	6055	3324
Site 3 σ_1	0.05	0.05	0.00	0.04	0.05	1	6527	3465
Site 4 σ_1	0.05	0.05	0.00	0.04	0.05	1	6019	3341
Site 5 σ_1	0.05	0.05	0.00	0.04	0.05	1	6164	3238
Site 6 σ_1	0.05	0.05	0.00	0.04	0.05	1	5797	3189
Site 7 σ_1	0.05	0.05	0.00	0.04	0.05	1	6146	3283
Site 8 σ_1	0.05	0.05	0.00	0.04	0.05	1	5800	3265
Site 9 σ_1	0.05	0.05	0.00	0.04	0.05	1	6374	3280
Site 10 σ_1	0.05	0.05	0.01	0.04	0.05	1	6445	3225
Site 11 σ_1	0.05	0.05	0.00	0.04	0.05	1	6186	3486
Site 12 σ_1	0.05	0.05	0.00	0.04	0.05	1	6500	3482
Site 13 σ_1	0.05	0.05	0.00	0.04	0.05	1	5610	3059
Site 14 σ_1	0.05	0.05	0.00	0.04	0.05	1	6082	3516
Site 15 σ_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 σ_2	0.04	0.04	0.00	0.04	0.05	1	4285	3288
Site 2 σ_2	0.04	0.04	0.00	0.03	0.05	1	4846	3415
Site 3 σ_2	0.04	0.04	0.00	0.03	0.05	1	4497	3379
Site 4 σ_2	0.04	0.04	0.00	0.04	0.05	1	4278	3329
Site 5 σ_2	0.04	0.04	0.00	0.03	0.05	1	4497	3378
Site 6 σ_2	0.04	0.04	0.00	0.03	0.05	1	4255	3939
Site 7 σ_2	0.04	0.04	0.00	0.03	0.05	1	3878	3332
Site 8 σ_2	0.04	0.04	0.00	0.03	0.05	1	4095	3456
Site 9 σ_2	0.04	0.04	0.00	0.03	0.05	1	4661	3448
Site 10 σ_2	0.04	0.04	0.00	0.03	0.05	1	4130	3547
Site 11 σ_2	0.04	0.04	0.00	0.03	0.05	1	4440	3362
Site 12 σ_2	0.04	0.04	0.00	0.03	0.05	1	4216	3590
Site 13 σ_2	0.04	0.04	0.00	0.03	0.05	1	4757	3407
Site 14 σ_2	0.04	0.04	0.00	0.04	0.05	1	4436	3141
Site 15 σ_2	0.04	0.04	0.00	0.03	0.05	1	4796	3449
Site 1 σ_3	0.04	0.04	0.00	0.03	0.05	1	4378	3465
Site 2 σ_3	0.04	0.04	0.00	0.03	0.05	1	4342	3351
Site 3 σ_3	0.04	0.04	0.00	0.03	0.05	1	4301	3492
Site 4 σ_3	0.04	0.04	0.00	0.03	0.05	1	4638	3445
Site 5 σ_3	0.04	0.04	0.00	0.03	0.04	1	4687	3613
Site 6 σ_3	0.04	0.04	0.00	0.03	0.05	1	4418	3216
Site 7 σ_3	0.04	0.04	0.00	0.03	0.05	1	4111	3079
Site 8 σ_3	0.04	0.04	0.00	0.03	0.05	1	4403	3249
Site 9 σ_3	0.04	0.04	0.00	0.03	0.05	1	4651	3346
Site 10 σ_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 σ_3	0.04	0.04	0.00	0.03	0.05	1	4625	3360
Site 12 σ_3	0.04	0.04	0.00	0.03	0.05	1	4316	3605
Site 13 σ_3	0.04	0.04	0.00	0.03	0.05	1	4486	3579
Site 14 σ_3	0.04	0.04	0.00	0.03	0.04	1	4498	3565
Site 15 σ_3	0.04	0.04	0.00	0.03	0.04	1	4484	3238

¹³⁵² **Appendix 2 (Supplemental Table 7)**

¹³⁵³ ***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 θ_1	0.87	0.87	0.04	0.79	0.93	1	6174	2962
Pinarbaşı B Sheep θ_2	0.08	0.08	0.04	0.02	0.15	1	5319	3352
Pinarbaşı B Sheep θ_3	0.05	0.05	0.04	0.01	0.13	1	6247	3246
Pinarbaşı B Sheep μ_1	-0.13	-0.13	0.02	-0.17	-0.09	1	3559	3130
Pinarbaşı B Sheep μ_2	-0.10	-0.10	0.02	-0.14	-0.06	1	3389	3209
Pinarbaşı B Sheep μ_3	-0.03	-0.03	0.03	-0.07	0.01	1	3569	2908
Pinarbaşı B Sheep σ_1	0.06	0.06	0.01	0.05	0.07	1	5239	3121
Pinarbaşı B Sheep σ_2	0.05	0.04	0.00	0.04	0.05	1	10190	2813
Pinarbaşı B Sheep σ_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 θ_1	0.17	0.16	0.05	0.09	0.26	1	1746	2189
NW Anatolian Cattle θ_2	0.67	0.68	0.06	0.56	0.77	1	1739	2254
NW Anatolian Cattle θ_3	0.16	0.16	0.05	0.09	0.24	1	1685	2126
NW Anatolian Cattle μ_1	-0.26	-0.26	0.03	-0.32	-0.21	1	1901	2118
NW Anatolian Cattle μ_2	-0.20	-0.20	0.03	-0.25	-0.15	1	1767	2050
NW Anatolian Cattle μ_3	-0.05	-0.05	0.03	-0.10	0.00	1	1841	2292
NW Anatolian Cattle σ_1	0.06	0.05	0.01	0.05	0.06	1	4127	3408
NW Anatolian Cattle σ_2	0.05	0.05	0.00	0.04	0.05	1	2513	2582
NW Anatolian Cattle σ_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 θ_1	0.33	0.33	0.04	0.27	0.40	1	3932	3536
Neolithic İlpınar θ_1	0.52	0.52	0.06	0.42	0.60	1	2322	2896
Menteşe θ_1	0.15	0.15	0.05	0.08	0.24	1	2190	3336
Chalcolithic İlpınar θ_1	0.13	0.13	0.02	0.10	0.17	1	4136	3519
Barcın θ_2	0.75	0.75	0.06	0.64	0.84	1	1115	2651
Neolithic İlpınar θ_2	0.12	0.12	0.06	0.03	0.22	1	1027	2511
Menteşe θ_2	0.22	0.22	0.06	0.14	0.33	1	4965	3208
Chalcolithic İlpınar θ_2	0.63	0.63	0.07	0.52	0.74	1	3136	3198
Barcın θ_3	0.14	0.14	0.05	0.07	0.23	1	2259	2766
Neolithic İlpınar θ_3	0.07	0.07	0.02	0.04	0.10	1	3781	3215
Menteşe θ_3	0.73	0.73	0.04	0.65	0.79	1	2483	3228
Chalcolithic İlpınar θ_3	0.21	0.20	0.04	0.14	0.28	1	2087	2936
Barcın μ_1	-0.40	-0.40	0.03	-0.44	-0.36	1	2568	3173
Neolithic İlpınar μ_1	-0.24	-0.24	0.02	-0.28	-0.21	1	2022	2658
Menteşe μ_1	-0.25	-0.25	0.03	-0.30	-0.20	1	2377	2891
Chalcolithic İlpınar μ_1	-0.29	-0.29	0.03	-0.33	-0.25	1	2270	3036
Barcın μ_2	-0.18	-0.18	0.02	-0.22	-0.14	1	2045	2507
Neolithic İlpınar μ_2	-0.21	-0.21	0.02	-0.24	-0.17	1	1803	2453
Menteşe μ_2	-0.19	-0.19	0.02	-0.23	-0.15	1	2074	2711
Chalcolithic İlpınar μ_2	-0.24	-0.24	0.02	-0.27	-0.21	1	1916	2416
Barcın μ_3	-0.02	-0.02	0.03	-0.07	0.03	1	2032	2723
Neolithic İlpınar μ_3	-0.07	-0.07	0.03	-0.12	-0.03	1	1581	2871
Menteşe μ_3	-0.04	-0.04	0.03	-0.09	0.01	1	2239	2545
Chalcolithic İlpınar μ_3	-0.09	-0.09	0.02	-0.13	-0.06	1	2041	2233
Barcın σ_1	0.06	0.06	0.01	0.05	0.07	1	4302	3717
Neolithic İlpınar σ_1	0.06	0.06	0.01	0.05	0.07	1	2964	3419
Menteşe σ_1	0.06	0.06	0.01	0.05	0.07	1	3950	3607
Chalcolithic İlpınar σ_1	0.06	0.05	0.01	0.05	0.07	1	3839	3731
Barcın σ_2	0.04	0.04	0.01	0.03	0.06	1	2503	2941
Neolithic İlpınar σ_2	0.07	0.07	0.01	0.05	0.08	1	1429	1427
Menteşe σ_2	0.05	0.05	0.01	0.04	0.06	1	3077	3188
Chalcolithic İlpınar σ_2	0.04	0.04	0.01	0.03	0.05	1	1552	2146
Barcın σ_3	0.05	0.05	0.01	0.04	0.06	1	2702	3106
Neolithic İlpınar σ_3	0.05	0.05	0.01	0.04	0.05	1	3433	3506
Menteşe σ_3	0.05	0.05	0.00	0.04	0.05	1	2786	3165
Chalcolithic İlpınar σ_3	0.05	0.05	0.00	0.04	0.05	1	2704	3247

1355 Appendix 3 (Supplemental Figures 1-4)

1356 *Traceplots of Posterior Distributions of Overall Model Parameters*

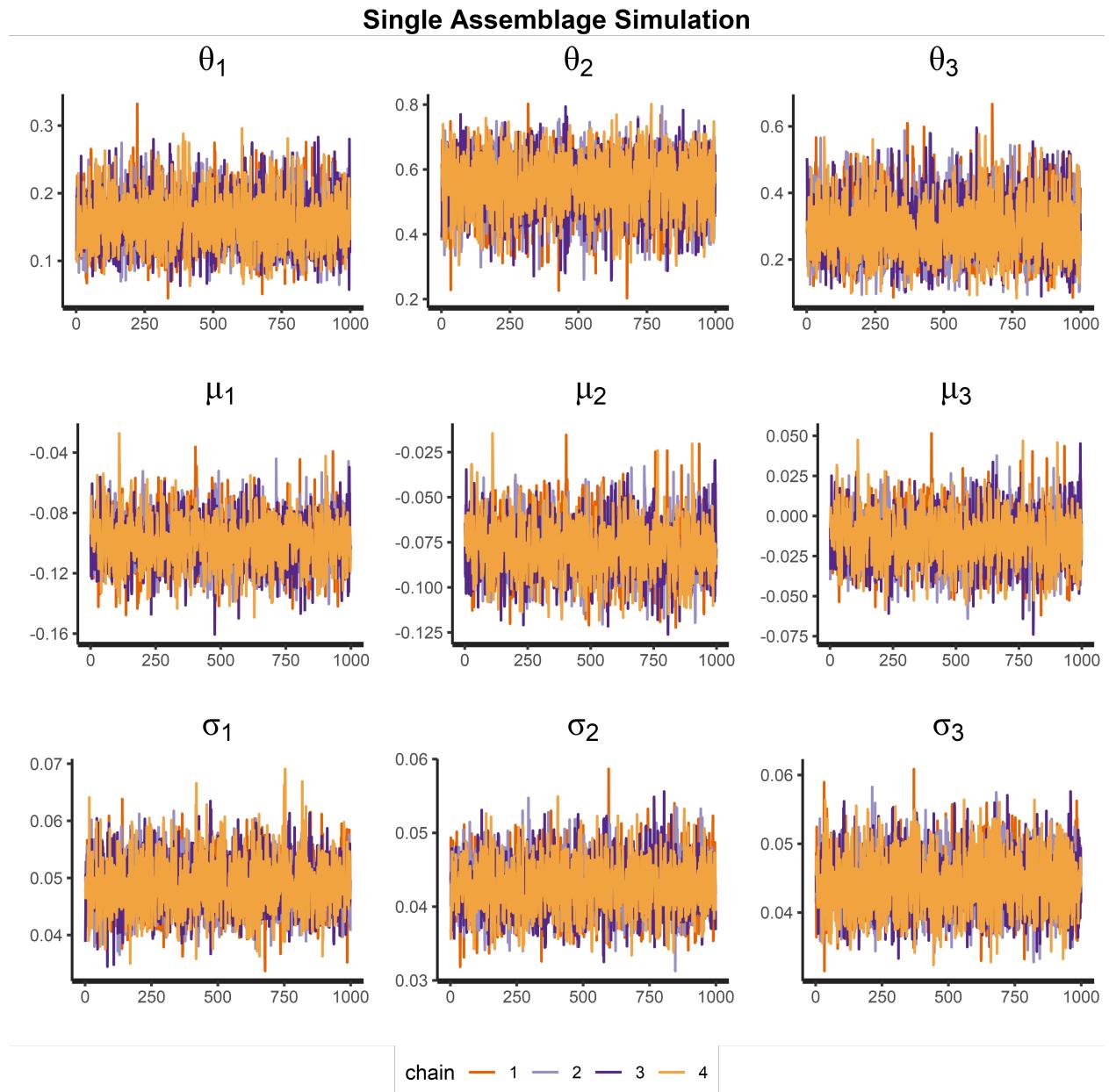


Figure 15: Traceplots of Model Parameters (Single Assemblage Simulation)

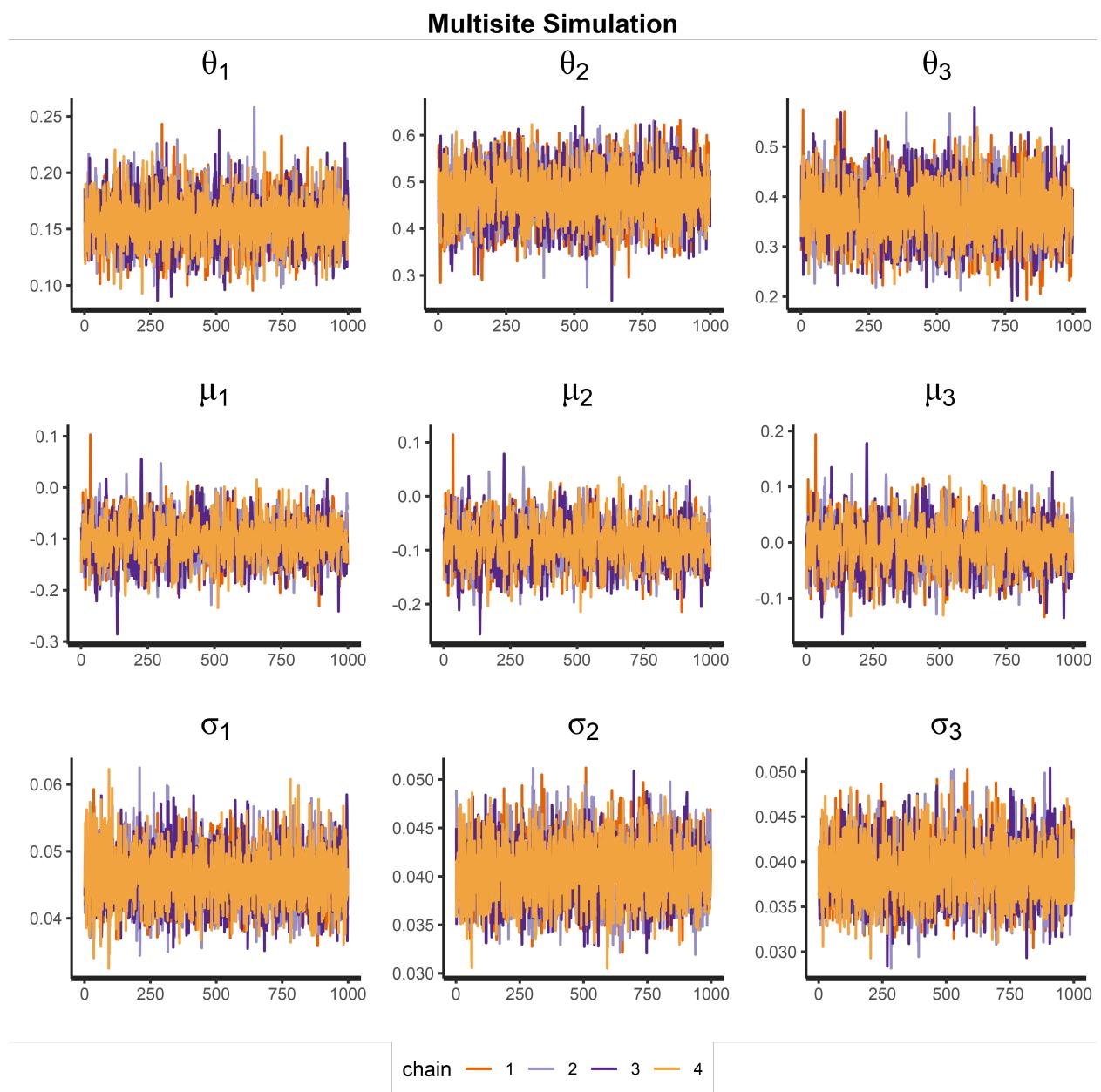


Figure 16: Traceplots of Model Parameters (Multisite Simulation)

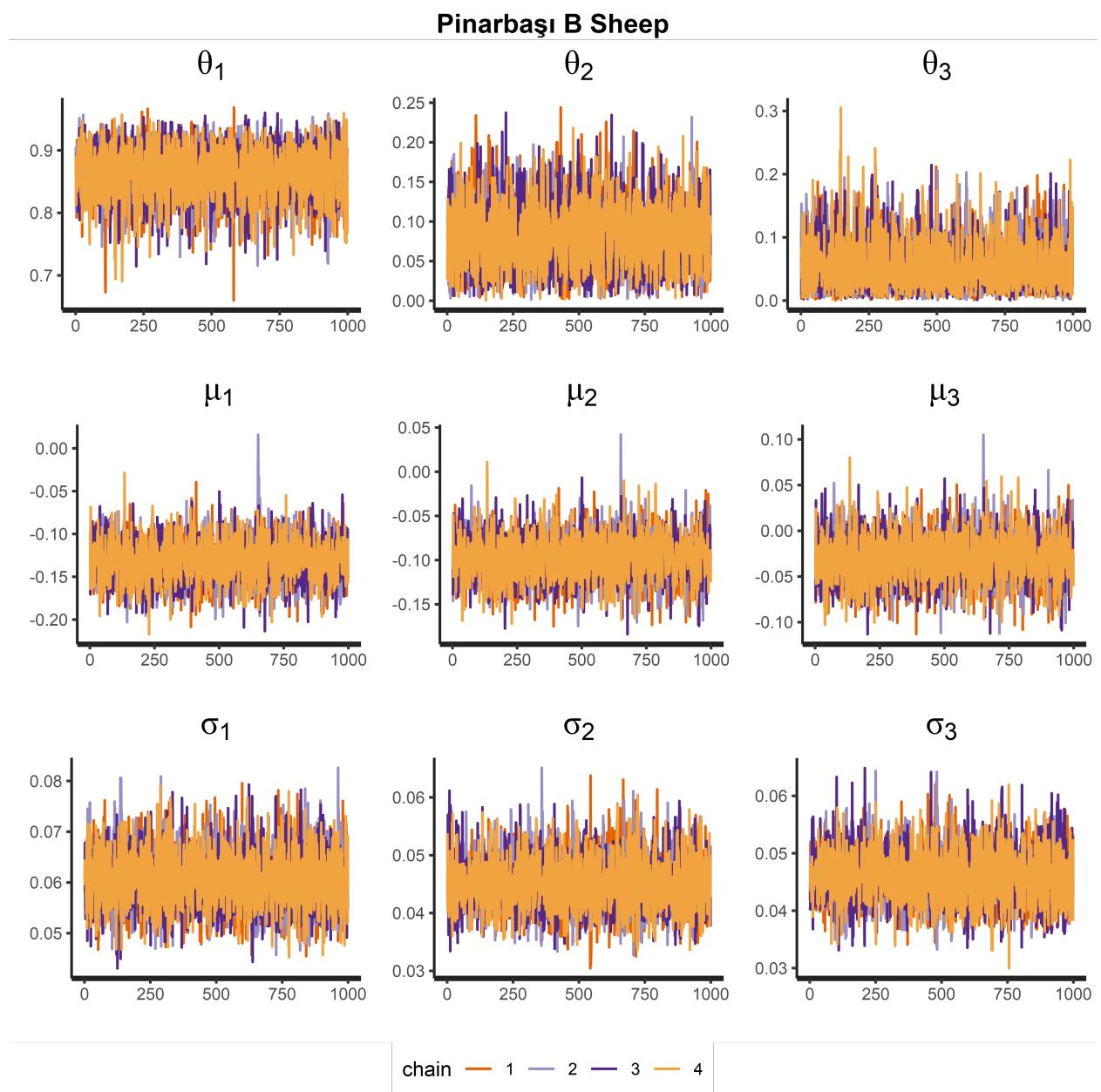


Figure 17: Traceplots of Model Parameters (Pınarbaşı B Sheep)

Northwest Anatolian Cattle

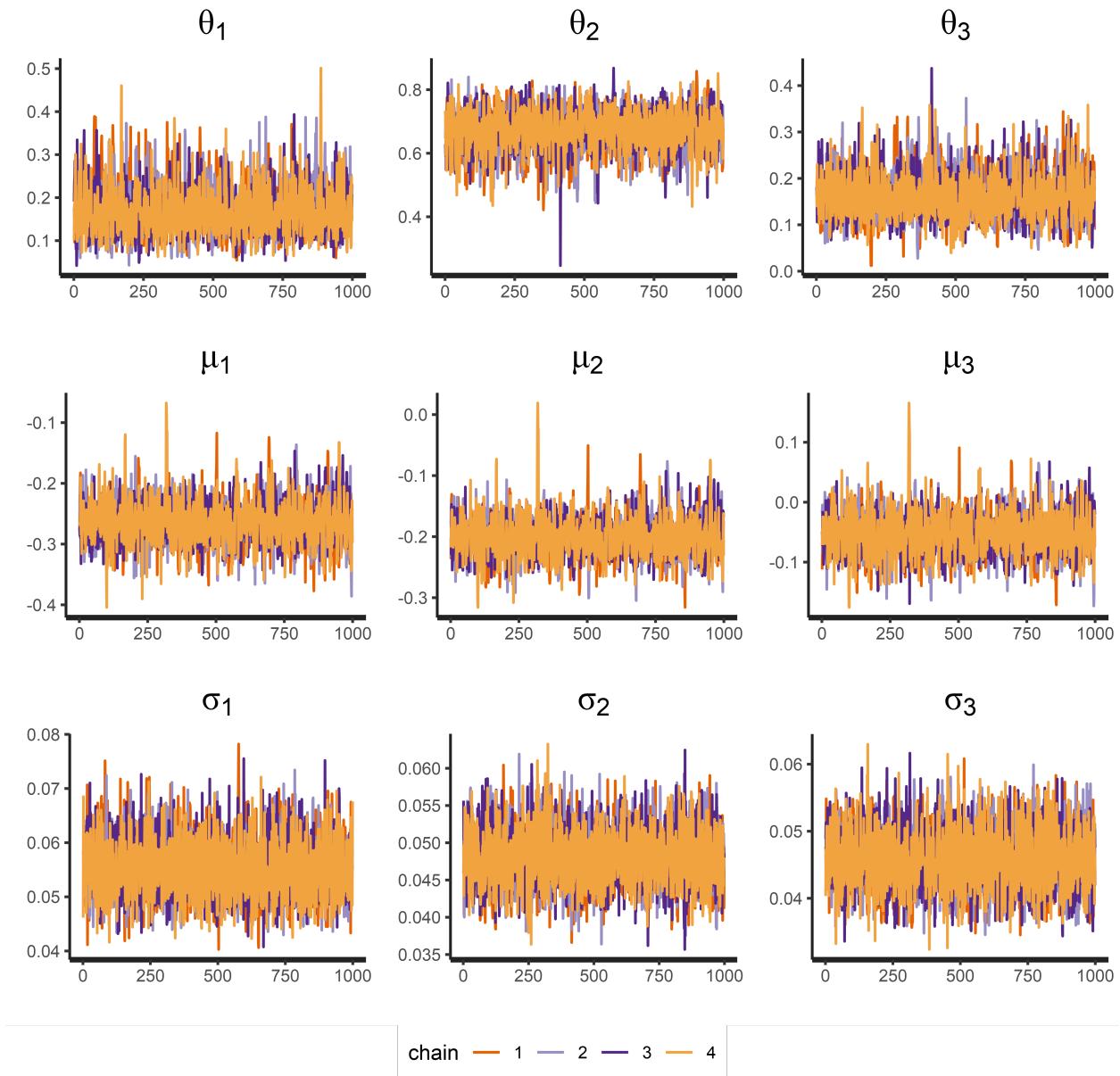


Figure 18: Traceplots of Model Parameters (Northwest Anatolian Cattle)

¹³⁵⁷ Appendix 4 (Supplemental Figures 5-7)

¹³⁵⁸ *Prior-Posterior Comparisons for Multisite, Pinarbaşı B, and NW Anatolian Model Hyper-*

¹³⁵⁹ *Parameters*

Multisite Simulation

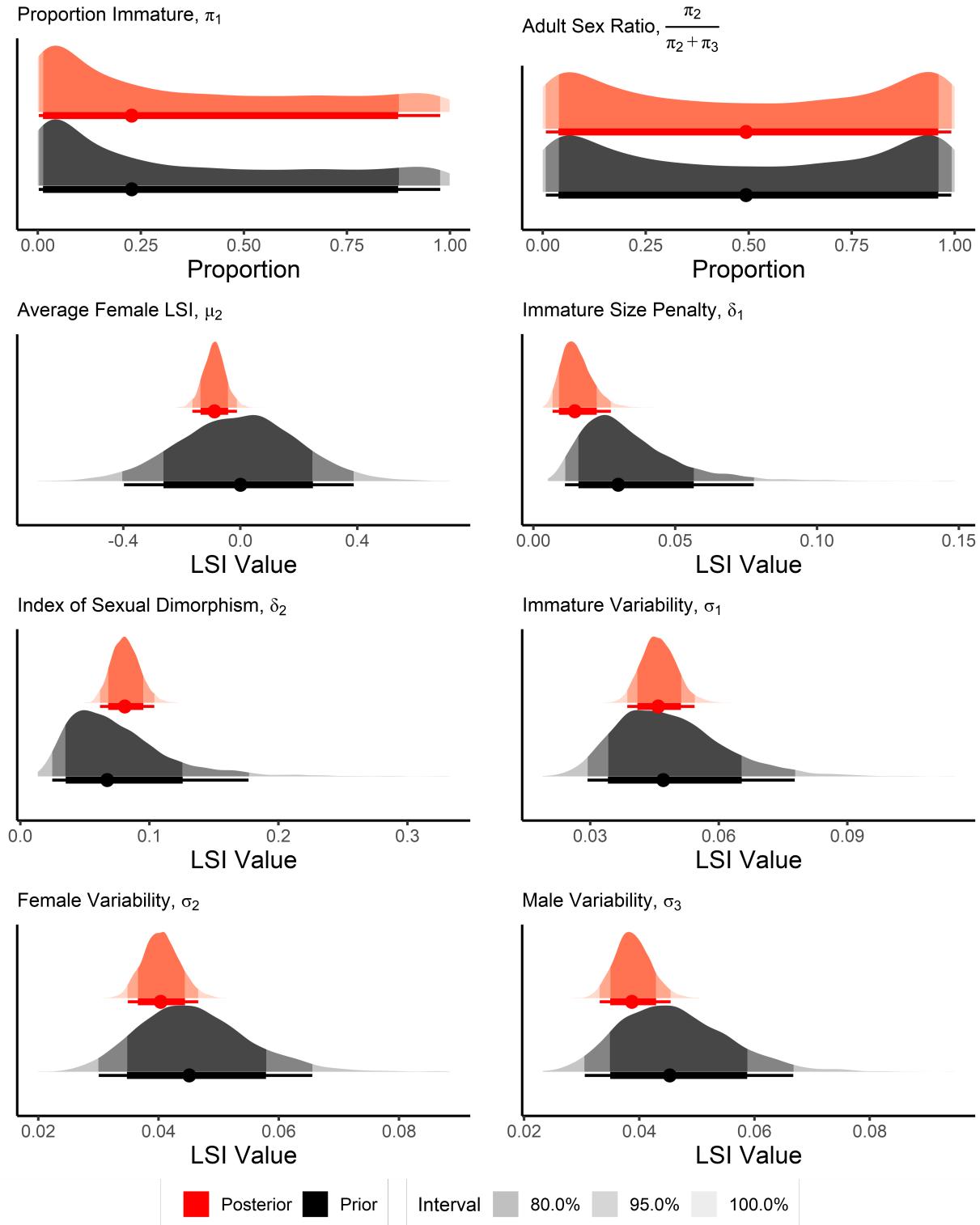


Figure 19: Prior-Posterior Comparison of Multisite Simulation Model Hyper-Parameters

Pinarbaşı B Sheep

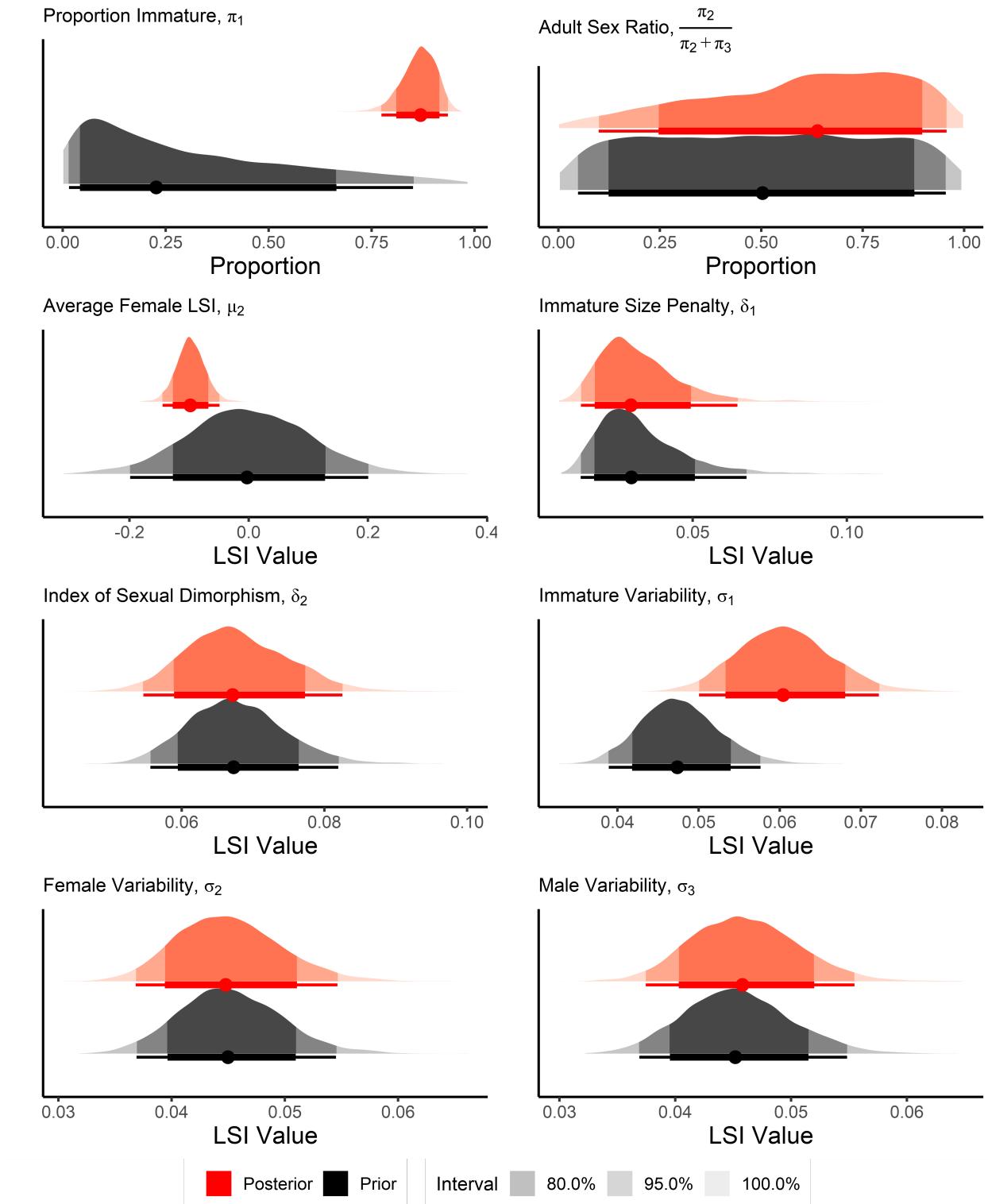


Figure 20: Prior-Posterior Comparison of Pinarbaşı B Sheep Model Hyper-Parameters

Northwest Anatolian Cattle

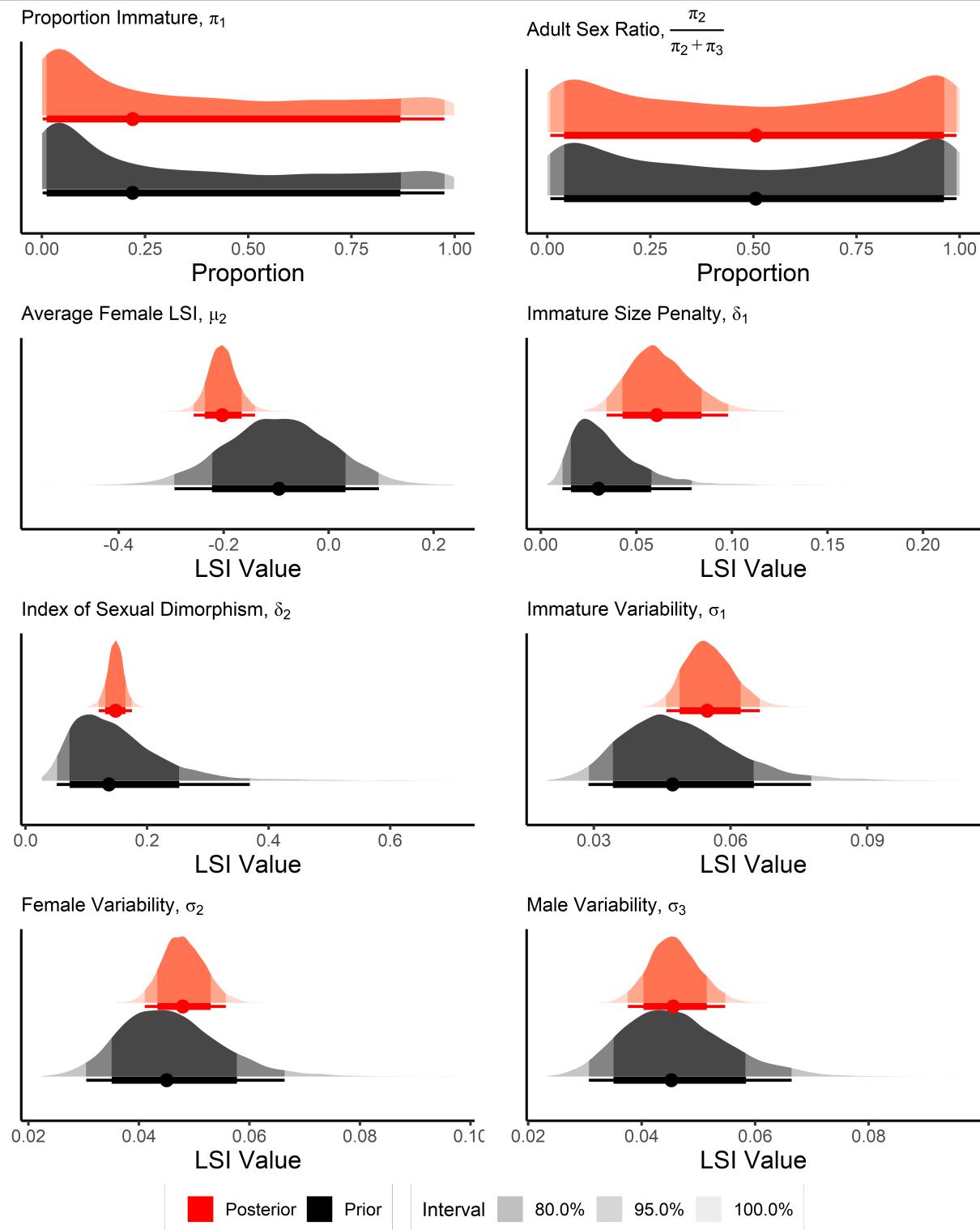


Figure 21: Prior-Posterior Comparison of NW Anatolian Cattle Model Hyper-Parameters

1360 ***Model Supplement: A Bayesian Multilevel Mixture Model for***
1361 ***Zooarchaeological Measurements***

1362 The Bayesian model developed for this paper describes assemblages of faunal measurements as a mixture of
1363 immature animals, (adult-sized) females, and (adult-sized) males that have distinct average body sizes and
1364 expected variation around that average size. The model uses multiple measured dimensions (e.g., humerus
1365 distal breadth “humerus Bd,” radius proximal breadth “radius Bp,” abbreviations following Driesch 1976),
1366 which are first converted into a logarithmic size index, or LSI, values with a natural logarithm base (Meadow
1367 1999; Wolfhagen 2020). LSI observations from measurement sets are grouped together within a specimen
1368 to create individuals grouped into defined “element portions” that serve as the basis for the mixture model
1369 analysis. Element portions relate to categories used for element fusion (e.g., distal humerus) to relate biome-
1370 try and mortality profiles; specimens that contain multiple element portions—like complete limb bones—are
1371 grouped into the latest-fusing element portion (compare to “skeletal part type” in **breslawski_2022?**).

1372 The multilevel structure of the model uses partial pooling to allow the parameters of the mixture model
1373 to vary between element portions while resisting overfitting. These element portion-specific parameters are
1374 related to each other through hyper-parameters, which describe the average value of the model parameters
1375 and the variability of model parameters across element portions (Wolfhagen 2020). The following sections
1376 describe the details of the multilevel mixture model, including the model’s likelihood, the ways that the direct
1377 observations of measurements and demographic data are used by the model to account for measurement error,
1378 and the development of prior distributions for the model’s hyper-parameters and for parameters that govern
1379 the model’s multilevel structure. Finally, this supplement provides the full sets of equations for the single-site
1380 and multisite Bayesian multilevel mixture models and the prior distribution definitions used in the model
1381 applications described in the main text.

1382 ***1. Definition of the Bayesian Multilevel Model***

1383 The central likelihood of the mixture model uses parameters that are specific to each element portion. These
1384 parameters include the relative proportions for the different animal groups: immature animals, females, and
1385 males (π_1, π_2, π_3), the average size for each group (μ_1, μ_2, μ_3), and the standard deviation for each group
1386 ($\sigma_1, \sigma_2, \sigma_3$). For each element portion, immature animals are described with the first set of parameters ($\pi_1,$
1387 μ_1, σ_1), adult-sized females with the second set of parameters (π_2, μ_2, σ_2), and adult-sized males with the
1388 third set of parameters (π_3, μ_3, σ_3). This results in both a set of parameters that describe the composition
1389 of the assemblage (of measurements from that element portion) and an equation to estimate the probability

1390 that a particular specimen comes from an immature, adult female, or adult male individual.

1391 *Mixture Model Likelihood Equation:*

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

1392 In addition to a specimen's LSI value, the model needs two additional observed variables to address the
1393 potential presence of immature animals in the assemblage. First, an indicator variable Immature[specimen]
1394 describes whether the specimen could be from an immature animal based on the body part and the fusion
1395 characteristics (1 = potentially immature, 0 = cannot be immature). Data from known-age Shetland sheep
1396 show that specimens killed at younger than one year of age are significantly smaller than those killed at
1397 older ages, regardless of fusion status (Popkin et al. 2012). Thus, any measurement from an element with
1398 an unfused epiphysis or from an element that does not fuse or could fuse before one year of age is considered
1399 potentially immature. Measurements from specimens with fused epiphyses that fuse after one year of age
1400 are considered ineligible to be immature, so the model does not consider that probability (it considers $\pi_1 =$
1401 0 for fitting that specimen).

1402 Second, the proportion of specimens from an element portion that could be immature ($\text{proportion}_{\text{immature}}$)
1403 determines how to re-weight the mixture components (π_1 , π_2 , and π_3) for potentially-immature specimens
1404 from that element portion. The mixture components describe the entire assemblage for an element portion
1405 (a combination of potentially-immature and non-immature specimens), meaning that if $\pi_1 = 0.25$ we should
1406 expect 25% of the specimens to be from immature animals. If every specimen could be immature—say, for
1407 specimens from an early-fusing element—then the mixture components do not need to be re-weighted. If,
1408 however, only half of the specimens could be immature, then the mixture components must be re-weighted
1409 to ensure that the whole-assemblage proportions are correct. In such a case, we would expect half of the
1410 potentially-immature animals to be from immature animals if $\pi_1 = 0.25$ for the whole assemblage and
1411 we know that $\text{proportion}_{\text{immature}} = (\frac{0.25}{0.50} = 0.50)$; this same re-weighting would make it less likely that
1412 potentially-immature animals are from adult-sized female or adult-sized male animals. The code includes
1413 checks to ensure that π_1 can never exceed 1.00 after accounting for the proportion of immature specimens in
1414 cases where there are very few potentially-immature specimens and/or a relatively high expected proportion
1415 of immature animals in the assemblage.

1416 **2. Measurement Error and Observations**

1417 The model estimates measurement error for different observed quantities that are input into the model.
1418 Measurements on both the archaeological specimens and the standard values used to calculate LSI values
1419 are assumed to have a 1% measurement error (Breslawski and Byers 2015; Popkin et al. 2012: Figure
1420 6). This 1% value comes from an evaluation of the Breslawski and Byers (2015) measurement data, where
1421 the average standard deviation of repeated measurements on bison radius proximal breadth measurements
1422 was 1.1% the average value of the measurement. This means that each measurement is given a standard
1423 deviation based on the observed value, which is used to estimate the “modeled” measurement value based on
1424 the observation. These modeled measurements are then used to calculate the LSI value for that observation
1425 ($LSI_{Measurement}$).

1426 Because specimens can have multiple measured dimensions that are included in the mixture model on
1427 them (e.g., a distal humerus with both Bd and BT observations or a complete radius with Bp and Bd
1428 observations), the mixture model calculates specimen-specific LSI values ($LSI_{Specimen}$) that are related to
1429 the observed measurement-specific LSI values ($LSI_{Measurement}$). $LSI_{Measurement}$ values are the “observations”
1430 with a standard deviation of 0.01 (in LSI_e scale) based on intra-individual variation of LSI_e values for the
1431 Popkin et al. (2012) sheep using the *Ovis orientalis* female standard animal (FMC 57951) from Uerpmann
1432 and Uerpmann (1994, Table 12).

1433 *Observation Error Equations for Measurements:*

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

1434 The model also uses observations of sex ratios and fusion rates to estimate assemblage-level demographic
1435 proportions. This allows relevant data to inform the model about the expected relative proportions of
1436 different animal groups while still allowing these proportions to vary across different element portions. These
1437 observations are interpreted as binomial data: counts of some quantity (e.g., immature specimens) out of a
1438 total count of relevant specimens (e.g., total ageable specimens); this approach lets the model incorporate the
1439 uncertainty caused by small sample sizes. The observation of the average proportion of immature specimens

1440 (μ_{π_1}) is based on the fusion rate of proximal and middle phalanges (the number of unfused phalanges
 1441 N_{Unfused} out of the total number of phalanges with fusion data N_{Ageable}), which fuse at around the same
 1442 time as the estimated time that animals reach adult body size in the Shetland sheep population (Popkin et
 1443 al. 2012). The observation of the average adult sex ratio—the proportion of females among mature animals
 1444 $\left(\frac{\mu_{\pi_2}}{\mu_{\pi_2} + \mu_{\pi_3}}\right)$ —is based on the sex ratio of fused pelvises (the number of female pelvises N_{Female} out of the
 1445 total number of pelvises with a sex assignment N_{Sexable}). In each case, the number of observable specimens
 1446 determines the measurement error using the binomial distribution. While this paper uses these quantities
 1447 to estimate the relevant hyper-parameters, relevant observations from other elements can be incorporated
 1448 into the model in the same fashion if there is a clear sense of the total number of specimens that could have
 1449 potentially been immature or female.

1450 *Observation Error Equations for Demographic Estimates:*

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

1451 ***3. Prior Distributions***

1452 The prior distributions in this model are focused on describing previous beliefs about the value of the mix-
 1453 ture model hyper-parameters, as the element portion-specific parameters are derived from these distributions.
 1454 Mixture modeling performs well with unconstrained parameters values because it is more straightforward to
 1455 estimate variation across element portions, meaning that constrained parameters—those where the range of
 1456 possible values depends on the values of other parameters—must first be transformed into related uncon-
 1457 strained parameters (Betancourt 2017) . The following sub-sections describe the necessary transformations
 1458 for different sets of the mixture model parameters, describing the unconstrained parameters that can be
 1459 modeled and the transformations that result in the mixture model parameters. While these sections use the
 1460 mixture model parameter notations, prior distributions are for the ‘central tendency’ hyper-parameter for
 1461 the described unconstrained parameter.

1462 It is important to remember that for all of these prior distributions are arbitrary choices made by the
 1463 researcher, regardless of whether the distributions are based on specific animal populations or on reference
 1464 priors. Other researchers could and should use different prior distributions to best reflect their intuition
 1465 about likely parameter values for particular case studies. This also highlight the importance of reporting
 1466 the prior distributions used in a Bayesian analysis to ensure replicability. Examining the implications of
 1467 different prior distributions is an important step in the development of Bayesian models, one that should be

1468 regularly tested even before models are fit to datasets (Gelman, Vehtari, et al. 2020).

1469 **3.1 Mixture Proportion Priors**

1470 Prior distributions for the mixture proportions reflect our prior beliefs about the relative proportions of the
1471 three animal groups in the assemblage (immature, adult-sized females, and adult-sized males). The three
1472 mixture proportions (π_1, π_2, π_3) are a three-value unit simplex, meaning that the values are constrained as
1473 a group to sum up to one. Thus, the simplex can be described by only two variables because the third value
1474 cannot vary once those two values are known. The model uses two unconstrained variables (θ_1 and θ_2) to
1475 describe the π values. These θ values are related back to π values using a ‘stick-breaking’ transformation that
1476 iteratively estimates the relative proportions of the simplex taken up by each θ value (Team 2022: Section
1477 10.7).

1478 *Stick-Breaking Transformations:*

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

1479 The θ_1 value can be directly related to the π_1 value using the first line of the stick-breaking transformation,
1480 meaning that one can examine the associated π_1 estimate for a given θ_1 value. Within the stick-breaking
1481 transformation, θ_2 relates to the relative proportions of π_2 and π_3 after π_1 has been estimated, which is
1482 effectively the adult sex ratio. Just as we could examine the expected π_1 estimates from a distribution of
1483 θ_1 values, we can thus use expected adult sex ratios ($\frac{\pi_2}{\pi_2 + \pi_3}$) estimates from a particular prior distribution
1484 for θ_2 . Relating these θ values back to observable phenomena makes it easier to define reasonable prior
1485 distribution definitions for the parameters from domain expertise (see Section 2.2 and Figure 2 of the main
1486 text).

1487 **3.2 Average Body Size and Size Variability Priors**

1488 While the average body sizes of the different components (μ_1, μ_2, μ_3) are not intrinsically linked in the
1489 same way that π values are, the model still requires some structure to aid interpretability. Bayesian mixture
1490 models that are fit using Markov Chain Monte Carlo (MCMC) methods, like the model in this paper, can
1491 suffer from an issue called “label switching” if μ values are not ordered in some way (Jasra, Holmes, and
1492 Stephens 2005). MCMC methods rely on running multiple “chains”—separate iterations of the model that
1493 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 μ_1 in one chain and to μ_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 μ_2 directly (average LSI_e value for females) and estimating the average LSI_e value for immature and male animals with offsets (δ_1, δ_2) from the female average. The δ values must be positive to maintain the ordering of the μ values, so each δ is modeled in a log-transformed space.

Offset Equations for μ 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 (δ_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 δ_1 and δ_2 values across different sites could conceivably highlight variation in the timing of the killing of immature animals (δ_1) and the degree of adult sexual dimorphism (δ_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 σ values should be relatively stable across elements. While σ parameters from different

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

3.3 Developing Priors from a Simulated Prior Assemblage

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

The prior distributions used in this Bayesian multilevel mixture model on the reference population are more straightforward. While the same transformations to create unconstrained parameters are necessary (e.g., modeling average size as μ_2 with offsets for immature and male animals), the definition of these prior distributions can be broadly described as weakly-informative priors (Gelman et al. 2008). These weakly-informative prior distributions are reasonable in this case—and not in the archaeological case—because all the mixture model parameters have direct observations rather than relying on latent state estimations. That is, parameters like the size difference between males and females (δ_2) and the size variability in female animals (σ_2) can be directly observed because the group identities of every specimen are known. With these direct observations, the prior distributions have a more muted influence on the resulting posterior distributions. This does not mean that the prior distributions have no effect, however, which is why objective priors can have undesirable impacts on modeling results (Gabry et al. 2019).

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

1553 The average LSI_e value for females (μ_2) is likely to vary across contexts in reaction to different selective
 1554 pressures, both anthropogenic and ecological (e.g., Davis 1981; Manning et al. 2015; Wright and Viner-
 1555 Daniels 2015). While the posterior distribution is extremely focused on this specific population, there is no
 1556 reason to think that this value should be centered at any particular value since that relates to the standard
 1557 animal used (Meadow 1999; Wolfhagen 2020). Therefore, the prior distribution used in archaeological models
 1558 for μ_2 uses a larger standard deviation, $\mu_2 \sim \text{Normal}(0, 0.1)$, to encompass likely LSI_e values (Figure 3A
 1559 in the main text). Under this definition, there is a 95% probability that the μ_2 value lies within the range
 1560 of -0.20 and 0.20 on the LSI_e scale, which translates to roughly 82-122% the size of the standard animal's
 1561 measurement.

1562 For the average size difference between immature and female animals (δ_1), the narrowness of the posterior
 1563 distribution likely reflects the fact that immature animals in the sample cover a narrow age range. Animals
 1564 killed under one year of age span only 36 days and the youngest animals are nearly half a year old (178-214
 1565 days: Popkin et al. 2012). Thus while the posterior results provide a useful starting point for estimating
 1566 this offset, there is a good potential for larger δ_1 values (i.e., greater size differences between immature and
 1567 adult-sized female animals) in other contexts that could include animals killed at a younger age (Figure 3B
 1568 in the main text). To capture this possibility, the archaeological model uses a prior distribution with a larger
 1569 standard deviation and a slightly higher center, $\log \delta_1 \sim \text{Normal}(-3.5, 0.4)$, which results in an average size
 1570 difference of 0.03 on the LSI_e scale and a 95% probability that the size difference is between 0.01 and 0.07.
 1571 This translates into expecting the average body size of immature animals in an assemblage being 3% smaller
 1572 than the average body size of adult-sized female animals, but also plausibly believing that this size difference
 1573 could range from 1-7% smaller.

1574 The average size difference between adult males and females (δ_2), also known as the index of sexual

1575 dimorphism (Fernández and Monchot 2007), is likely to be under stricter biological control than the other
1576 “average body size” parameters in the model. This does not mean that this difference could not vary between
1577 contexts, however. Some models of animal domestication argue that initial domestication removed sexual
1578 selective pressures on male body size, reducing sexual dimorphism (e.g., Tchernov and Horwitz 1991). In
1579 a similar fashion, specialized hunting strategies could also reduce sexual dimorphism by targeting large-
1580 bodied males, for example (Zeder 2012; Proaktor, Coulson, and Milner-Gulland 2007; Milner, Nilsen, and
1581 Andreassen 2007). Again, the posterior distribution of the extent of sexual dimorphism in the Shetland
1582 sheep population provides a useful starting point to describe a prior distribution for the model (Figure 3C
1583 in the main text). Increasing the standard deviation of the distribution slightly, $\log \delta_2 \sim \text{Normal}(-2.7, 0.1)$,
1584 produces a distribution centered at 0.07 LSI_e units with a 95% probability that the value is between 0.06-0.08,
1585 translating to the average male being 6-9% larger than the average female relative to a standard measurement.
1586 The smaller standard deviation in the prior distribution of δ_2 than for δ_1 reflects our understanding that
1587 the extent of sexual dimorphism, as a biological phenomenon, is less likely to have extreme values than
1588 the average size difference between immature and female animals, since δ_2 is unaffected by the specific age
1589 structure of the assemblage.

1590 As in the average body size parameters, prior distributions for the size variability model parameters
1591 are developed from the Bayesian model of known-identity Shetland sheep measurements. The resulting σ
1592 hyper-parameters provide a baseline for establishing hyper-parameter prior distributions in archaeological
1593 cases. Figure 3D-F of the main text shows the posterior distributions of these σ hyper-parameters in both
1594 the log-transformed values and associated LSI_e values. Average size variability within an element portion for
1595 immature animals (σ_1) is higher, on average, than for females (σ_2) and males (σ_3). The immature category
1596 includes both male and female animals, so larger size variability makes sense; again, it is possible that σ_1 is
1597 relatively low in this population relative to other contexts given the narrow age range of immature animals
1598 in the Shetland sheep population. Unlike the average body size parameters, there are not compelling reasons
1599 to believe that size variability parameters for females and males (σ_2 and σ_3) should vary widely in different
1600 contexts given the consistency of coefficients of variation in mammals broadly (Davis 1996). Thus, the results
1601 of this analysis are used for the prior distributions of $\log \sigma_2$ and $\log \sigma_3$, while the prior distribution of $\log \sigma_1$
1602 is given an increased standard deviation and slightly increased average value. Overall, however, these prior
1603 distributions suggest that the average size variability within an element portion is between 0.04-0.06 for
1604 females and males and is between 0.04-0.05 for immature animals. Note that even though σ_2 and σ_3 have
1605 the same prior distributions, these values can still vary from each other in different contexts.

1606 **4. Multilevel Structure of the Model**

1607 The previous section described prior distributions that describe the *average* value for different mixture model
 1608 parameters across all element portions. To create parameter estimates that are specific to different element
 1609 portions, it is necessary to estimate the *variation* around these average values that different parameters
 1610 can have among different element portions. The model uses a Multivariate Normal definition of the model
 1611 parameters to allow for correlations between different parameters; effectively, the possibility that multiple
 1612 model parameters will covary from element portion to element portion. To do this, each hyper-parameter
 1613 has an associated σ_{element} parameter that describes inter-element variation in parameter values. The model
 1614 uses a non-centered parameterization, wherein the Multivariate Normal distribution is centered at zero to
 1615 calculate offsets, ν_{element} , that are added to the average hyper-parameters to calculate model parameters for
 1616 each element portion. This definition provides computational stability and makes it more straightforward
 1617 to incorporate other levels of multilevel structure.

1618 *Equations for Defining Inter-Element Variation (Multilevel Modeling):*

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

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

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

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

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

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

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

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

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

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

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

1619 The multilevel structure used to allow variation in parameter estimates across element portions can

1620 also be expanded to create multisite models that can directly compare sex-specific biometric estimates
 1621 alongside the age/sex composition of different assemblages. Such comparisons can highlight variation in
 1622 herd management strategies or diachronic body size trends related to population turnover (e.g., Arbuckle
 1623 and Atici 2013; Arbuckle et al. 2016). To do this, an additional multilevel structure can be applied to the
 1624 same mixture model parameters, using σ_{Site} rather than $\sigma_{Element}$ parameters. However, an additional set of
 1625 multilevel structure parameters, $\sigma_{Interaction}$, are also necessary to ensure that elemental variation is different
 1626 at different sites (e.g., the difference between μ_2 for the distal humerus and μ_2 for the distal radius is not
 1627 necessarily the same at different sites). Again, weakly-informative priors are appropriate for both sets of
 1628 parameters. Each additional term is included in the sum to create specific mixture model parameter values.

1629 *Example of Parameter Definition for Inter-Site and Inter-Element Variation:*

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

1630 The inclusion of multiple sites changes the definition of the ‘grand mean’ variable (θ_1 in the example
 1631 equation) from a site-level estimate to an overall mean across the sites and elements. These parameter esti-
 1632 mates thus describe the average composition of the entire set of assemblages. The details of the assemblages
 1633 included in the analyses would affect how useful these estimates are for interpretation. Assemblage-specific
 1634 estimates can be calculated for each model parameter by adding the relevant ν_{site} estimate to the ‘grand
 1635 mean’ parameter, which would again act to describe the average composition of the assemblage regardless
 1636 of its elemental composition.

1637 Prior distributions for $\sigma_{Element}$ values (and σ_{Site} and $\sigma_{Interaction}$ values in multisite models) are weakly-
 1638 informative priors based on the scale of the parameter and the expectation for variation for the parameter.
 1639 For example, there is likely more variation in θ parameters—that govern the relative composition of imma-
 1640 ture, female, and male animals—among element portions than variation in σ parameters that govern size
 1641 variability within each group. Similarly, it is expected that average body sizes of females μ_2 will vary more
 1642 between sites $\sigma_{Site}[3]$ than between elements within a site $\sigma_{Element}[3]$. The impacts of these prior distri-
 1643 bution definitions were evaluated using prior predictive checking, suggesting that these prior distributions
 1644 allow enough variability to encompass reasonable size estimates without providing too much prior support
 1645 to implausible or impossible values (see Section 6 of the Model Supplement).

1646 *Prior Distributions for Element-level Variation (Multilevel Component):*

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

₁₆₄₇ *Prior Distributions for Site-level Variation (Multilevel Component):*

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

₁₆₄₈ *Prior Distributions for Interaction Effect (Multilevel Component):*

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

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

₁₆₅₀ *Prior Distribution Definitions for the Single Assemblage Simulation Model Hyper-Parameters*

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

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

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

1654

Prior Distribution Definitions for the Northwest Anatolian Cattle Model Hyper-Parameters

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

1655 6. Simulating Assemblages from the Prior Distributions (Prior Predictive 1656 Checking)

1657 Prior predictive checks are a critical component of Bayesian model development workflows, ensuring that
 1658 reasonable prior definitions are chosen (Gabry et al. 2019; Gelman, Vehtari, et al. 2020). This process
 1659 uses the model's prior distribution definitions to simulate data, which can then be evaluated against domain
 1660 knowledge and observed data. This is particularly important when dealing with model parameters that are
 1661 difficult to examine in isolation, like parameters that govern inter-element variation (σ_{Element}) in a multilevel
 1662 model structure. Prior predictive checking is an iterative process, informing researchers about the potential
 1663 consequences of their prior distribution definitions; in particular, it can highlight how excessively imprecise
 1664 definitions can provide considerable prior weight on implausible and even impossible values for data (Gabry
 1665 et al. 2019: Figure 4). Thus, prior predictive checking allows researchers to create more accurate summaries
 1666 of their domain knowledge but also produces more efficient MCMC performance because less time is spent
 1667 evaluating parameter values that are inconsistent with even cursory prior knowledge about the problem being
 1668 modeled.

1669 Prior predictive checks were developed for both the single-assemblage and multisite model fits, using the
 1670 prior distribution definitions used in the sheep simulations in the main text (see Section 5 of the Model
 1671 Supplement). In each simulation, 25 LSI_{Specimen} values were calculated for each of the 5 element portions
 1672 based on the relevant prior distributions and model structures; for the multisite simulation, this was done
 1673 for 3 assemblages. To evaluate the feasibility of these prior distribution definitions, the LSI_{Specimen} were
 1674 converted into simulated measurement values based on the reference values of the *Ovis orientalis* female
 1675 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

1676 five element portions were converted into measurements from specific dimensions, following the equations in
 1677 Section 2 of this model supplement. Each simulation was run 1000 times, producing 1000 assemblages of
 1678 relevant simulated measurement values.

1679 Figure 1 shows two results for the single assemblage (left) and multisite (right) prior prediction sim-
 1680 ulations. The top row shows the simulated proportions of immature, female, and male specimens in the
 1681 simulated assemblages (Figure 1A). These proportions were sampled directly from the element-specific mix-
 1682 ture proportion variables, showing the range of potential distributions the model is expecting before seeing
 1683 any data. The 95% quantiles of the single assemblage model's proportion of immature specimens ranges
 1684 from 2-87%, 2-83% for female specimens, and 2-82% for male specimens. It is noteworthy that the simulated
 1685 proportion of immature specimens for the single assemblage model does not reach the extreme value seen in
 1686 the Pınarbaşı B sheep data, which may explain some of the long tails in the element-specific compositional
 1687 estimates (Figure 9 of the main text). The multisite model has similar expected ranges (immature: 2-85%,
 1688 female: 2-82%, male: 3-82%).

1689 The bottom row shows the distribution of simulated humerus Bd measurements for the assemblages
 1690 (Figure 1B). These plots include all simulated specimens, showing the range of sheep humerus Bd values the
 1691 model expects before seeing any data. Several vertical lines on the plots give a sense of domain knowledge
 1692 about sheep humerus Bd values. First, the standard value (33.0 mm) from the *Ovis orientalis* female
 1693 standard animal (FMC 57951) from Uerpmann and Uerpmann (1994: Table 12) is shown in the red dashed
 1694 line. Second, vertical blue lines show ranges of observed archaeological sheep humerus Bd values: the largest
 1695 sheep (technically *Ovis orientalis*) humerus Bd (39.0 mm) from the 10th millennium BP site of Körtik Tepe,
 1696 in southeastern Anatolia (Arbuckle and Özkaya 2006: Table b) and the smallest sheep (*Ovis aries*) humerus
 1697 Bd (25.0 mm) from the fifth-sixth century CE site of West Stow, United Kingdom (Crabtree 1990: Table
 1698 29). These plots highlight how the multisite model's structure allows for much more variation in body

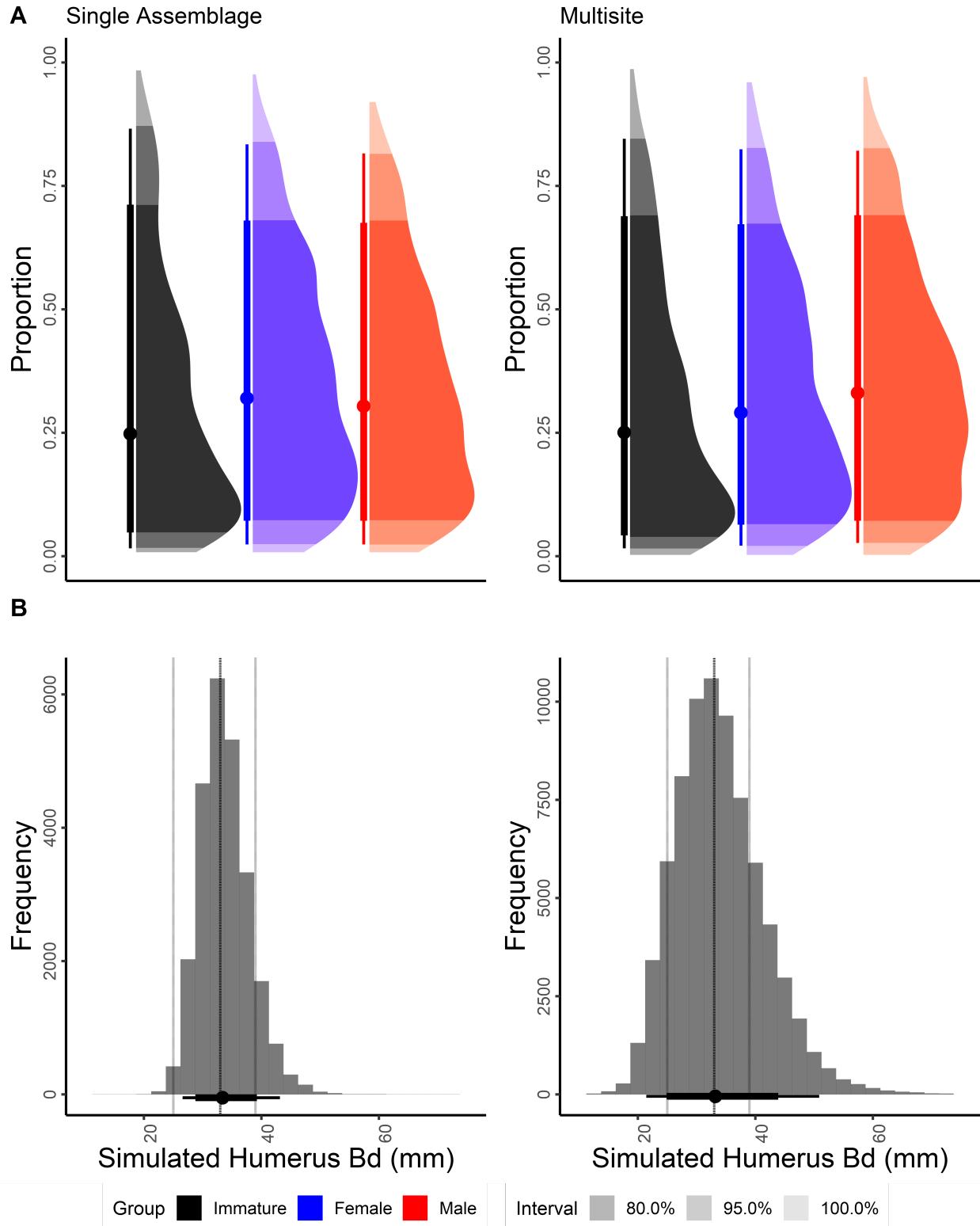


Figure 22: Single Assemblage and Multisite Prior Predictive Checks. Top row (A): Estimates of the proportion of immature, female, and male specimens. Bottom row (B): Histograms of simulated humerus Bd measurements from the simulations. Vertical lines show the value of the standard reference value (dashed)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

size, especially large measurements. While the 95% quantiles of the single assemblage model's simulated measurements (27-43 mm) do not exactly encapsulate the range of the observed extreme measurements, the multisite model's simulated measurements go well beyond the range (21-51 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

- 1730 Appearance of Domestic Cattle in the Eastern Fertile Crescent (Northern Iraq and Western Iran).”
1731 Journal Article. *Journal of Archaeological Science* 72: 1–9. <https://doi.org/10.1016/j.jas.2016.05.008>.
- 1732 Baird, Douglas, Denise Carruthers, Andrew Fairbairn, and Jessica A. Pearson. 2011. “Ritual in the Land-
1733 scape: Evidence from Pınarbaşı in the Seventh-Millennium Cal BC Konya Plain.” Journal Article.
1734 *Antiquity* 85: 380–94.
- 1735 Baird, Douglas, Andrew Fairbairn, Emma Jenkins, Louise Martin, Caroline Middleton, Jessica A. Pearson,
1736 Eleni Asouti, et al. 2018. “Agricultural Origins on the Anatolian Plateau.” Journal Article. *Proc Natl
1737 Acad Sci U S A*. <https://doi.org/10.1073/pnas.1800163115>.
- 1738 Benaglia, Tatiana, Didier Chauveau, David R. Hunter, and Derek Young. 2009. “mixtools: An R Package
1739 for Analyzing Finite Mixture Models.” *Journal of Statistical Software* 32 (6): 1–29. <http://www.jstatsoft.org/v32/i06/>.
- 1740 Betancourt, Michael. 2017. “A Conceptual Introduction to Hamiltonian Monte Carlo.” Journal Article.
1741 *arXiv* 1701.02434 [Preprint]: Available from <https://arxiv.org/abs/1701.02434v1>.
- 1742 Breslawski, Ryan P. 2023. “Minimum Animal Units and the Standardized Count Problem.” Journal Article.
1743 *Journal of Archaeological Method and Theory*. <https://doi.org/10.1007/s10816-022-09563-9>.
- 1744 Breslawski, Ryan P., and David A. Byers. 2015. “Assessing Measurement Error in Paleozoological Osteo-
1745 metrics with Bison Remains.” Journal Article. *Journal of Archaeological Science* 53: 235–42. <https://doi.org/10.1016/j.jas.2014.10.001>.
- 1746 Buitenhuis, Hjalke. 2008. “Ilipinar: The Faunal Remains from the Late Neolithic and Early Chalcolithic
1747 Levels.” Book Section. In *Archaeozoology of the Near East VIII*.
- 1748 ———. 2013. “Ilipinar Zooarchaeology.” Dataset. <http://opencontext.org/projects/D297CD29-50CA-4B2C-4A07-498ADF3AF487>. <https://doi.org/https://doi.org/10.6078/M76H4FB>.
- 1749 Çakırlar, Canan. 2013. “Rethinking Neolithic Subsistence at the Gateway to Europe With New Archaeo-
1750 zoological Evidence from Istanbul.” Book Section. In *Barely Surviving or More Than Enough?: The
1751 Environmental Archaeology of Subsistence, Specialisation and Surplus Food Production*, edited by Maaike
1752 Groot, Daphne Lentjes, and Jørn Zeiler, 59–80. Leiden: Sidestone Press.
- 1753 Canty, Angelo, and B. D. Ripley. 2021. *Boot: Bootstrap r (s-Plus) Functions*.
- 1754 Carruthers, Denise. 2005. “Hunting and Herding in Central Anatolian Prehistory: The Sites at Pinarbasi.”
1755 Book Section. In *Archaeozoology of the Near East VI*.
- 1756 ———. 2006. “Pınarbaşı 1994: Animal Bones.” Dataset. <https://opencontext.org/projects/1677643a-15d0-1b2c-3a35-37f04c765387>.
- 1757 Corporation, Microsoft, and Steve Weston. 2022. *doParallel: Foreach Parallel Adaptor for the 'Parallel'
1758 Package*. <https://CRAN.R-project.org/package=doParallel>.

- 1763 Corti, Paulo, and David M. Shackleton. 2002. "Relationship Between Predation-Risk Factors and Sexual
1764 Segregation in Dall's Sheep (*Ovis Dalli Dalli*).” Journal Article. *Canadian Journal of Zoology* 80 (12):
1765 2108–17. <https://doi.org/10.1139/z02-207>.
- 1766 Crabtree, Pam J. 1990. *West Stow: Early Anglo-Saxon Animal Husbandry*. Book. East Anglian Archaeology.
1767 Suffolk: Suffolk County Planning Dept.
- 1768 Crema, Enrico R. 2012. "Modelling Temporal Uncertainty in Archaeological Analysis." Journal Article.
1769 *Journal of Archaeological Method and Theory* 19 (3): 440–61. <https://doi.org/10.1007/s10816-011-9122-3>.
- 1770 Dahl, Gudrun, and Anders Hjort. 1976. *Having Herds: Pastoral Herd Growth and Household Economy*.
1771 Book. Stockholm Studies in Social Anthropology. Stockholm: University of Stockholm.
- 1772 Daly, K. G., V. Mattiangeli, A. J. Hare, H. Davoudi, H. Fathi, S. B. Doost, S. Amiri, et al. 2021. "Herded
1773 and Hunted Goat Genomes from the Dawn of Domestication in the Zagros Mountains." Journal Article.
1774 *Proc Natl Acad Sci U S A* 118 (25). <https://doi.org/10.1073/pnas.2100901118>.
- 1775 Davis, Simon J. M. 1981. "The Effects of Temperature Change and Domestication on the Body Size of Late
1776 Pleistocene to Holocene Mammals of Israel." Journal Article. *Paleobiology* 7 (1): 101–14.
- 1777 ———. 1982. "Climatic Change and the Advent of Domestication: The Succession of Ruminant Artiodactyls
1778 in the Late Pleistocene-Holocene in the Israel Region." Journal Article. *Paléorient* 8 (2): 5–15.
- 1779 ———. 1996. "Measurements of a Group of Adult Female Shetland Sheep Skeletons from a Single Flock:
1780 A Baseline for Zooarchaeologists." Journal Article. *Journal of Archaeological Science* 23: 593–612.
- 1781 ———. 2000. "The Effect of Castration and Age on the Development of the Shetland Sheep Skeleton and
1782 a Metric Comparison Between Bones of Males, Females and Castrates." Journal Article. *Journal of
1783 Archaeological Science* 27 (5): 373–90. <https://doi.org/10.1006/jasc.1999.0452>.
- 1784 Davison, A. C., and D. V. Hinkley. 1997. *Bootstrap Methods and Their Applications*. Cambridge: Cambridge
1785 University Press. <http://statwww.epfl.ch/davison/BMA/>.
- 1786 De Cupere, Bea, Wim Van Neer, Hervé Monchot, Elina Rijmenants, Mircea Udrescu, and Marc Waelkens.
1787 2005. "Ancient Breeds of Domestic Fowl (*Gallus Gallus f. Domestica*) Distinguished on the Basis of
1788 Traditional Observations Combined with Mixture Analysis." Journal Article. *Journal of Archaeological
1789 Science* 32 (11): 1587–97. <https://doi.org/10.1016/j.jas.2005.04.015>.
- 1790 Degerbøl, Magnus. 1970. "The Urus (*Bos Primigenius Bojanus*) and Neolithic Domesticated Cattle (*Bos
1791 Taurus Domesticus Linné*) in Denmark." Journal Article. *Det Kongelige Danske Videnskabernes Selskab
1792 Biologiske Skrifter* 17 (1): 5–177.
- 1793 Dong, Zhuan. 1997. "Mixture Analysis and Its Preliminary Application in Archaeology." Journal Article.
1794 *Journal of Archaeological Science* 24: 141–61.
- 1795

- 1796 Dowle, Matt, and Arun Srinivasan. 2021. *Data.table: Extension of ‘Data.frame’*. <https://CRAN.R-project.org/package=data.table>.
- 1797
- 1798 Driesch, Angela von den. 1976. *A Guide to the Measurement of Animal Bones from Archaeological Sites*.
1799 Book. Peabody Museum Bulletins. Cambridge, MA: Harvard University.
- 1800 Evershed, Richard P., Sebastian Payne, Andrew G. Sherratt, Mark S. Copley, Jennifer Coolidge, Duska
1801 Urem-Kotsu, Kostas Kotsakis, et al. 2008. “Earliest Date for Milk Use in the Near East and Southeastern
1802 Europe Linked to Cattle Herding.” Journal Article. *Nature* 455 (7212): 528–31. <https://doi.org/10.1038/nature07180>.
- 1803
- 1804 Fernández, Hélène, and Hervé Monchot. 2007. “Sexual Dimorphism in Limb Bones of Ibex (*Capra Ibex*
1805 l.): Mixture Analysis Applied to Modern and Fossil Data.” Journal Article. *International Journal of
1806 Osteoarchaeology* 17: 479–91. <https://doi.org/10.1002/oa.876>.
- 1807 Fernée, Christianne L., and Konstantinos P. Trimmis. 2021. “Detecting Variability: A Study on the
1808 Application of Bayesian Multilevel Modelling to Archaeological Data. Evidence from the Neolithic Adri-
1809 atic and the Bronze Age Aegean.” Journal Article. *Journal of Archaeological Science* 128. <https://doi.org/10.1016/j.jas.2021.105346>.
- 1810
- 1811 Gabry, Jonah, and Rok Cešnovar. 2022. *Cmdstanr: R Interface to ‘CmdStan’*.
- 1812 Gabry, Jonah, Daniel Simpson, Aki Vehtari, Michael Betancourt, and Andrew Gelman. 2019. “Visualization
1813 in Bayesian Workflow.” Journal Article. *Journal of the Royal Statistical Society, Series A* 182 (2): 389–
1814 402.
- 1815 Galik, Alfred. 2013. “Barçın Höyük Zooarchaeology.” Dataset. <http://opencontext.org/projects/74749949-4FD4-4C3E-C830-5AA75703E08E>. <https://doi.org/10.6078/M78G8HM0>.
- 1816
- 1817 Gelman, Andrew. 2006a. “Multilevel (Hierarchical) Modeling: What It Can and Cannot Do.” Journal
1818 Article. *Technometrics* 48 (3): 432–35. <https://doi.org/10.1198/004017005000000661>.
- 1819 ———. 2006b. “Prior Distributions for Variance Parameters in Hierarchical Models (Comment on Article
1820 by Browne and Draper).” Journal Article. *Bayesian Analysis* 3: 515–34.
- 1821 Gelman, Andrew, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin. 2020.
1822 *Bayesian Data Analysis*. Book. Third Edition.
- 1823 Gelman, Andrew, Aleks Jakulin, Maria Grazia Pittau, and Yu-Sung Su. 2008. “A Weakly Informative
1824 Default Prior Distribution for Logistic and Other Regression Models.” Journal Article. *The Annals of
1825 Applied Statistics* 2 (4): 1360–83. <https://doi.org/10.1214/08-aos191>.
- 1826 Gelman, Andrew, Aki Vehtari, Daniel Simpson, Charles C. Margossian, Bob Carpenter, Yuling Yao, Lauren
1827 Kennedy, Jonah Gabry, Paul-Christian Bürkner, and Martin Modrák. 2020. “Bayesian Workflow.”
1828 Journal Article. *ArXiv*.

- 1829 Gerritsen, Fokke, and Rana Özbal. 2019. “Barcin Höyük, a Seventh Millennium Settlement in the Eastern
1830 Marmara Region of Turkey.” Journal Article. *Documenta Praehistorica* 46: 58–67. <https://doi.org/10.4312/dp.46.4>.
- 1832 Gillis, R., R. M. Arbogast, J. F. Piningre, K. Debue, and J. D. Vigne. 2015. “Prediction Models for Age-
1833 at-Death Estimates for Calves, Using Unfused Epiphyses and Diaphyses.” Journal Article. *International
1834 Journal of Osteoarchaeology* 25 (6): 912–22. <https://doi.org/10.1002/oa.2377>.
- 1835 Gillis, R., I. Carrère, M. Saña Seguí, G. Radi, and J. D. Vigne. 2014. “Neonatal Mortality, Young Calf
1836 Slaughter and Milk Production During the Early Neolithic of North Western Mediterranean.” Journal
1837 Article. *International Journal of Osteoarchaeology* 26 (2): 303–13. <https://doi.org/10.1002/oa.2422>.
- 1838 Gourichon, Lionel, and Daniel Helmer. 2008. “Etude de La Faune Neolithique de Mentesé.” Book Section. In
1839 *Life and Death in a Prehistoric Settlement in Northwest Anatolia. The Ilipinar Excavations, Volume III.*,
1840 edited by Jacob Roodenberg and Songül Alpaslan Roodenberg, 435–48. Leiden: Nederlands Instituut
1841 voor he Nabije Oosten.
- 1842 ———. 2013. “Faunal Data from Neolithic Mentesé.” Dataset. <https://opencontext.org/projects/42ead4dbbaed-4a58-9b9b-7ec85266d2a9>. <https://doi.org/https://doi.org/10.6078/M7S46PVN>.
- 1844 Grau-Sologestoia, Idoia, and Umberto Albarella. 2019. “The ‘Long’ Sixteenth Century: A Key Period of
1845 Animal Husbandry Change in England.” Journal Article. *Archaeological and Anthropological Sciences* 11
1846 (6): 2781–2803. <https://doi.org/10.1007/s12520-018-0723-6>.
- 1847 Grigson, Caroline. 1982. “Sex and Age Determination of Some Bones and Teeth of Domestic Cattle: A
1848 Review of the Literature.” Book Section. In *Ageing and Sexing Animal Bones from Archaeological Sites*,
1849 edited by Bob Wilson, Caroline Grigson, and Sebastian Payne, 7–23. BAR British Series. Oxford: British
1850 Archaeological Reports.
- 1851 ———. 1989. “Size and Sex: Evidence for the Domestication of Cattle in the Near East.” Book Section.
1852 In *People and Culture in Change: Preceedings of the 2nd Symposium on Upper Paleolithic, Mesolithic,
1853 and Mesolithic Populations of Europe and the Mediterranean Basin*, edited by Israel Hershkovitz, 77–109.
1854 BAR International Series. Oxford: Archaeopress.
- 1855 Hammer, Oyvind. 2013. *PAST PAleontological STatistics Reference Manual*. Book.
- 1856 Helmer, Daniel, Lionel Gourichon, Hervé Monchot, Joris Peters, and Maria Saña Segui. 2005. “Identifying
1857 Early Domestic Cattle from Pre-Pottery Neolithic Sites on the Middle Euphrates Using Sexual Dimor-
1858 phism.” Book Section. In *First Steps of Animal Domestication: New Archaeozoological Approaches*,
1859 edited by Jean-Denis Vigne, Joris Peters, and Daniel Helmer, 86–95. Proceedings of the 9th Conference
1860 of the International Council of Archaeozoology, Durham, August 2002. Oxford: Oxbow Books.
- 1861 Hill, Matthew E., Matthew G. Hill, and Christopher C. Widga. 2008. “Late Quaternary Bison Diminution

- 1862 on the Great Plains of North America: Evaluating the Role of Human Hunting Versus Climate Change.”
1863 Journal Article. *Quaternary Science Reviews* 27 (17-18): 1752–71. <https://doi.org/10.1016/j.quascirev>.
1864 2008.07.002.
- 1865 Hongo, Hitomi, Jessica A. Pearson, Banu Öksüz, and Gülcin Ilgezdi. 2009. “The Process of Ungulate
1866 Domestication at Çayönü, Southeastern Turkey: A Multidisciplinary Approach Focusing on Bos Sp.
1867 And *Cervus Elaphus*.” Journal Article. *Anthropozoologica* 44 (1): 63–78.
- 1868 Jasra, A., C. C. Holmes, and D. A. Stephens. 2005. “Markov Chain Monte Carlo Methods and the Label
1869 Switching Problem in Bayesian Mixture Modeling.” Journal Article. *Statistical Science* 20 (1): 50–67.
1870 <https://doi.org/10.1214/088342305000000016>.
- 1871 Kabukcu, Ceren. 2017. “Woodland Vegetation History and Human Impacts in South-Central Anatolia
1872 16,000–6500 Cal BP: Anthracological Results from Five Prehistoric Sites in the Konya Plain.” Journal
1873 Article. *Quaternary Science Reviews* 176: 85–100. <https://doi.org/10.1016/j.quascirev.2017.10.001>.
- 1874 Karul, Necmi. 2019. “Early Farmers in Northwestern Anatolia in the Seventh Millennium.” Book Section.
1875 In *Concluding the Neolithic: The Near East in the Second Half of the Seventh Millennium BC*, edited by
1876 Arkadiusz Marciniak, 269–86. Atlanta, Georgia: Lockwood Press.
- 1877 Kassambara, Alboukadel. 2020. *Ggpubr: 'Ggplot2' Based Publication Ready Plots*. <https://CRAN.R-project.org/package=ggpubr>.
- 1878 Kay, Matthew. 2022. *ggdist: Visualizations of Distributions and Uncertainty*. <https://doi.org/10.5281/zenodo.3879620>.
- 1879 Koch, Paul L. 1986. “Clinal Geographic Variation in Mammals: Implications for the Study of Chronoclines.”
1880 Journal Article. *Paleobiology* 12 (3): 269–81.
- 1881 Lebennon, Roxanne, and Natalie D. Munro. 2022. “Body Size Variation in a Modern Collection of Mountain
1882 Gazelle (*Gazella Gazella*) Skeletons.” Journal Article. *Journal of Archaeological Science: Reports* 41.
1883 <https://doi.org/10.1016/j.jasrep.2021.103285>.
- 1884 Lyman, R. Lee. 2008. *Quantitative Paleozoology*. Book. Cambridge Manuals in Archaeology. Cambridge:
1885 Cambridge University Press.
- 1886 Madgwick, Richard, and Jacqui Mulville. 2015. “Reconstructing Depositional Histories Through Bone
1887 Taphonomy: Extending the Potential of Faunal Data.” Journal Article. *Journal of Archaeological Science*
1888 53: 255–63. <https://doi.org/10.1016/j.jas.2014.10.015>.
- 1889 Manning, Katie, Adrian Timpson, Stephen Shennan, and Enrico Crema. 2015. “Size Reduction in Early
1890 European Domestic Cattle Relates to Intensification of Neolithic Herding Strategies.” Journal Article.
1891 *PLoS One* 10 (12): e0141873. <https://doi.org/10.1371/journal.pone.0141873>.
- 1892 Marin, Jean-Michel, Kerrie Mengerson, and Christian P. Robert. 2005. “Bayesian Modelling and Inference
1893 2008.07.002.

- 1895 on Mixtures of Distributions.” Journal Article. *Handbook of Statistics* 25 (Bayesian Thinking, Modeling
1896 and Computation): 459–509.
- 1897 Marom, Nimrod, and Guy Bar-Oz. 2013. “The Prey Pathway: A Regional History of Cattle (*Bos Taurus*)
1898 and Pig (*Sus Scrofa*) Domestication in the Northern Jordan Valley, Israel.” Journal Article. *PLoS One*
1899 8 (2): e55958. <https://doi.org/10.1371/journal.pone.0055958.g001>.
- 1900 Martín, Patricia, and Ricardo García-González. 2015. “Identifying Sheep (*Ovis Aries*) Fetal Remains in
1901 Archaeological Contexts.” Journal Article. *Journal of Archaeological Science* 64: 77–87. <https://doi.org/10.1016/j.jas.2015.10.003>.
- 1903 Martín, Patricia, Ricardo García-González, Jordi Nadal, and Josep Maria Vergès. 2015. “Perinatal Ovi-
1904 caprine Remains and Evidence of Shepherding Activities in Early Holocene Enclosure Caves: El Mirador
1905 (Sierra de Atapuerca, Spain).” Journal Article. *Quaternary International*. <https://doi.org/10.1016/j.quaint.2015.08.024>.
- 1907 Marwick, Ben. 2017. “Computational Reproducibility in Archaeological Research: Basic Principles and a
1908 Case Study of Their Implementation.” Journal Article. *Journal of Archaeological Method and Theory* 24
1909 (2): 424–50. <https://doi.org/10.1007/s10816-015-9272-9>.
- 1910 Marwick, Ben, and Suzanne E. Pilaar Birch. 2018. “A Standard for the Scholarly Citation of Archaeological
1911 Data as an Incentive to Data Sharing.” Journal Article. *Advances in Archaeological Practice* 6 (02):
1912 125–43. <https://doi.org/10.1017/aap.2018.3>.
- 1913 McCarthy, Michael A., and P. I. P. Masters. 2005. “Profiting from Prior Information in Bayesian Analyses
1914 of Ecological Data.” Journal Article. *Journal of Applied Ecology* 42 (6): 1012–19. <https://doi.org/10.1111/j.1365-2664.2005.01101.x>.
- 1916 McElreath, Richard. 2020. *Statistical Rethinking: A Bayesian Course with Examples in r and Stan*. Book.
1917 Texts in Statistical Science. Boca Raton, FL: CRC Press.
- 1918 McGrory, S., E. M. Svensson, A. Götherström, J. Mulville, A. J. Powell, M. J. Collins, and T. P. O’Connor.
1919 2012. “A Novel Method for Integrated Age and Sex Determination from Archaeological Cattle Mandibles.”
1920 Journal Article. *Journal of Archaeological Science* 39 (10): 3324–30. <https://doi.org/10.1016/j.jas.2012.05.021>.
- 1922 Meadow, Richard. 1999. “The Use of Size Index Scaling Techniques for Research on Archaeozoological
1923 Collections from the Middle East.” Book Section. In *Historia Animalium Ex Ossibus: Beiträge Zur
1924 Paläoanatomie, Archäologie, Ägyptologie, Ethnologie Und Geschichte Der Tiermedizin*, edited by Cornelia
1925 Becker, Henriette Manhart, Joris Peters, and Jörg Schibler, 285–300. Leidorf: Verlag Marie.
- 1926 Meier, Jacqueline. 2020. “The Contextual Taphonomy of Middens at Neolithic Kfar HaHoresh.” Journal
1927 Article. *Journal of Archaeological Science: Reports* 33. <https://doi.org/10.1016/j.jasrep.2020.102531>.

- 1928 Middleton, Caroline. 2018. "The Beginning of Herding and Animal Management: The Early Development
1929 of Caprine Herding on the Konya Plain, Central Anatolia." Journal Article. *Anatolian Studies* 68: 1–31.
1930 <https://doi.org/10.1017/s0066154618000017>.
- 1931 Millard, A. R. 2006. "A Bayesian Approach to Ageing Sheep/Goats from Toothwear." Book Section. In
1932 *Recent Advances in Ageing and Sexing Animal Bones*, edited by Deborah Rusillo, 145–54. Oxford:
1933 Oxbow Books.
- 1934 Milner, J. M., E. B. Nilsen, and H. P. Andreassen. 2007. "Demographic Side Effects of Selective Hunting in
1935 Ungulates and Carnivores." Journal Article. *Conserv Biol* 21 (1): 36–47. <https://doi.org/10.1111/j.1523-1739.2006.00591.x>.
- 1937 Monchot, Hervé, and Jacques Léchelle. 2002. "Statistical Nonparametric Methods for the Study of Fossil
1938 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).
- 1940 Monchot, Hervé, Marjan Mashkour, and Jean-Denis Vigne. 2005. "Kernel Smoothing and Mixture Analyses
1941 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
1942 Jean-Denis Vigne, Joris Peters, and Daniel Helmer, 55–60. Oxford: Oxbow Books.
- 1944 Munro, Natalie D., Guy Bar-Oz, and Austin C. Hill. 2011. "An Exploration of Character Traits and Linear
1945 Measurements for Sexing Mountain Gazelle (*Gazella Gazella*) Skeletons." Journal Article. *Journal of
1946 Archaeological Science* 38 (6): 1253–65. <https://doi.org/10.1016/j.jas.2011.01.001>.
- 1947 Munro, Natalie D., Roxanne Lebennon, and Lisa Sapir-Hen. 2022. "Revisiting Late Pleistocene-Early
1948 Holocene Mountain Gazelle (*Gazella Gazella*) Body Size Change in the Southern Levant: A Case for
1949 Anthropogenic Impact." Journal Article. *PLoS One* 17 (8): e0273024. <https://doi.org/10.1371/journal.pone.0273024>.
- 1951 Otárola-Castillo, Erik, and Melissa G. Torquato. 2018. "Bayesian Statistics in Archaeology." Journal Article.
1952 *Annual Review of Anthropology* 47 (1). <https://doi.org/10.1146/annurev-anthro-102317-045834>.
- 1953 Otárola-Castillo, Erik, Melissa G. Torquato, Jesse Wolfhagen, Matthew E. Hill, and Caitlin E. Buck. 2022.
1954 "Beyond Chronology, Using Bayesian Inference to Evaluate Hypotheses in Archaeology." Journal Article.
1955 *Advances in Archaeological Practice*.
- 1956 Özdoğan, Mehmet. 2011. "Archaeological Evidence on the Westward Expansion of Farming Communities
1957 from Eastern Anatolia to the Aegean and the Balkans." Journal Article. *Current Anthropology* 52 (S4):
1958 S415–30. <https://doi.org/10.1086/658895>.
- 1959 ———. 2019. "Early Farmers in Northwestern Turkey: What Is New?" Book Section. In *Concluding
1960 the Neolithic: The Near East in the Second Half of the Seventh Millennium BC*, edited by Arkadiusz

- 1961 Marciniak, 307–28. Atlanta, Georgia: Lockwood Press.
- 1962 Payne, Sebastian. 1969. “A Metrical Distinction Between Sheep and Goat Metacarpals.” Book Section. In
1963 *The Domestication and Exploitation of Plants and Animals*, edited by Peter J. Ucko and Geoffrey W.
1964 Dimbleby, 295–305. London: Duckworth.
- 1965 ———. 1973. “Kill-Off Patterns in Sheep and Goats: The Mandibles from Aşvan Kale.” Journal Article.
1966 *Anatolian Studies* 23: 281–303.
- 1967 Payne, Sebastian, and Gail Bull. 1988. “Components of Variation in Measurements of Pig Bones and Teeth,
1968 and the Use of Measurements to Distinguish Wild from Domestic Pigs.” Journal Article. *Archaeozoologia*
1969 II (1,2): 27–66.
- 1970 Pebesma, Edzer. 2018. “Simple Features for R: Standardized Support for Spatial Vector Data.” *The R
1971 Journal* 10 (1): 439–46. <https://doi.org/10.32614/RJ-2018-009>.
- 1972 Pérez-Barbería, F. Javier, Iain J. Gordon, and M. Pagel. 2002. “The Origins of Sexual Dimorphism in Body
1973 Size in Ungulates.” Journal Article. *Evolution* 56 (6): 1276–85.
- 1974 Peters, Joris, Angela von den Driesch, and Daniel Helmer. 2005. “The Upper Euphrates-Tigris Basin:
1975 Cradle of Agro-Pastoralism?” Book Section. In *The First Steps of Animal Domestication*, edited by
1976 Jean-Denis Vigne, Joris Peters, and Daniel Helmer, 96–124. Proceedings of the 9th Conference of the
1977 International Council of Archaeozoology, Durham, August 2002. Oxford: Oxbow Books.
- 1978 Popkin, Peter R. W., Polydora Baker, Fay Worley, Sebastian Payne, and Andy Hammon. 2012. “The Sheep
1979 Project (1): Determining Skeletal Growth, Timing of Epiphyseal Fusion and Morphometric Variation
1980 in Unimproved Shetland Sheep of Known Age, Sex, Castration Status and Nutrition.” Journal Article.
1981 *Journal of Archaeological Science* 39 (6): 1775–92. <https://doi.org/10.1016/j.jas.2012.01.018>.
- 1982 Post, Diane M., Trent S. Armbrust, Eva A. Horne, and Jacob R. Goheen. 2001. “Sexual Segregation
1983 Results in Differences in Content and Quality of Bison (*Bos Bison*) Diets.” Journal Article. *Journal of
1984 Mammalogy* 82 (2): 407–13.
- 1985 Pozo, Jose M, Valenzuela-Lamas, Silvia, Trentacoste, Angela, Nieto-Espinet, Ariadna, Guimarães Chiarelli,
1986 and Silvia. 2021. “Zoolog: Zooarchaeological Analysis with Log-Ratios.” *Journal of Statistical Software*
1987 In preparation. <https://josempozo.github.io/zooolog/>.
- 1988 Proaktor, G., T. Coulson, and E. J. Milner-Gulland. 2007. “Evolutionary Responses to Harvesting in
1989 Ungulates.” Journal Article. *J Anim Ecol* 76 (4): 669–78. <https://doi.org/10.1111/j.1365-2656.2007.01244.x>.
- 1991 R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R
1992 Foundation for Statistical Computing. <https://www.R-project.org/>.
- 1993 Ram, Karthik, and Ben Marwick. 2018. “Building Toward a Future Where Reproducible, Open Science Is

- 1994 the Norm.” Book Section. In *The Practice of Reproducible Research: Case Studies and Lessons from the*
1995 *Data-Intensive Sciences*, edited by Justin Kitzes, Daniel Turek, and Fatma Deniz, 69–78. Oakland, CA:
1996 University of California Press.
- 1997 Redding, Richard W. 1984. “Theoretical Determinants of a Herder’s Decisions: Modeling Variation in the
1998 Sheep/Goat Ratio.” Book Section. In *Animals and Archaeology*, edited by Juliet Clutton-Brock and
1999 Caroline Grigson, 3. Early Herders and their Flocks:223–41. BAR International Series. Oxford: British
2000 Archaeological Reports.
- 2001 Roodenberg, Jacob. 2012a. “Change in Food Production and Its Impact on an Early 6th Millennium
2002 Community in Northwest Anatolia. The Example of Ilipinar.” Journal Article. *Praehistorische Zeitschrift*
2003 87 (2). <https://doi.org/10.1515/pz-2012-0015>.
- 2004 ———. 2012b. “Ilipinar: A Neolithic Settlement in the Eastern Marmara Region.” Book Section. In
2005 *The Oxford Handbook of Ancient Anatolia (10,000-323 BCE)*, edited by Gregory McMahon and Sharon
2006 Steadman. <https://doi.org/10.1093/oxfordhb/9780195376142.013.0044>.
- 2007 Roodenberg, Jacob, A. van As, L. Jacobs, and M. H. Wijnen. 2003. “Early Settlement in the Plain of
2008 Yenisehir (NW Anatolia): The Basal Occupation Layers at Mentese.” Journal Article. *Anatolica* XXIX
2009 (1): 18–58.
- 2010 Rosenthal, Robert. 1979. “The ”File Drawer Problem” and Tolerance for Null Results.” Journal Article.
2011 *Psychological Bulletin* 86 (3): 638–41.
- 2012 RStudio Team. 2022. *RStudio: Integrated Development Environment for r*. Boston, MA: RStudio, PBC.
2013 <http://www.rstudio.com/>.
- 2014 Ruckstuhl, K. E. 2007. “Sexual Segregation in Vertebrates: Proximate and Ultimate Causes.” Journal
2015 Article. *Integr Comp Biol* 47 (2): 245–57. <https://doi.org/10.1093/icb/icm030>.
- 2016 Ruckstuhl, K. E., and P. Neuhaus. 2002. “Sexual Segregation in Ungulates: A Comparative Test of
2017 Three Hypotheses.” Journal Article. *Biological Reviews* 77 (1): 77–96. <https://doi.org/10.1017/s1464793101005814>.
- 2019 Russell, Nerissa, and Janet L. Griffitts. 2013. “Catalhoyuk Worked Bones: South and 4040 Areas.” Book
2020 Section. In *Substantive Technologies at Catalhoyuk: Reports from the 2000-2008 Seasons*, edited by Ian
2021 Hodder, 277–306. Los Angeles.
- 2022 Said, Sonia, Vincent Tolon, Serge Brandt, and Eric Baubet. 2011. “Sex Effect on Habitat Selection in
2023 Response to Hunting Disturbance: The Study of Wild Boar.” Journal Article. *European Journal of*
2024 *Wildlife Research* 58 (1): 107–15. <https://doi.org/10.1007/s10344-011-0548-4>.
- 2025 Sasson, Aharon, and Susan Arter. 2020. “Earliest Utilization of Chicken in Upper California: The Zooar-
2026 chaeology of Avian Remains from the San Diego Royal Presidio.” Journal Article. *American Antiquity*

- 2027 85 (3): 516–34. <https://doi.org/10.1017/aaq.2020.27>.
- 2028 Scrucia, Luca, Michael Fop, T. Brendan Murphy, and Adrian E. Raftery. 2016. “mclust 5: Clustering,
2029 Classification and Density Estimation Using Gaussian Finite Mixture Models.” *The R Journal* 8 (1):
2030 289–317. <https://doi.org/10.32614/RJ-2016-021>.
- 2031 Slowikowski, Kamil. 2021. *Ggrepel: Automatically Position Non-Overlapping Text Labels with 'Ggplot2'*.
2032 <https://CRAN.R-project.org/package=ggrepel>.
- 2033 Smith, A. F. M., and Alan E. Gelfand. 1992. “Bayesian Statistics Without Tears: A Sampling-Resampling
2034 Perspective.” Journal Article. *The American Statistician* 46 (2): 84–88.
- 2035 South, Andy. 2017a. *Rnaturalearth: World Map Data from Natural Earth*. <https://CRAN.R-project.org/>
2036 package=rnaturalearth.
- 2037 ———. 2017b. *Rnaturalearthdata: World Vector Map Data from Natural Earth Used in 'Rnaturalearth'*.
2038 <https://CRAN.R-project.org/package=rnaturalearthdata>.
- 2039 Speth, John D. 1983. *Bison Kills and Bone Counts: Decision Making by Ancient Hunters*. Book. Chicago:
2040 University of Chicago Press.
- 2041 ———. 2013. “Thoughts about Hunting: Some Things We Know and Some Things We Don’t Know.”
2042 Journal Article. *Quaternary International* 297: 176–85. <https://doi.org/10.1016/j.quaint.2012.12.005>.
- 2043 Stan Development Team. 2021. “RStan: The R Interface to Stan.” <https://mc-stan.org/>.
- 2044 Stiner, Mary C. 1990. “The Use of Mortality Patterns in Archaeological Studies of Hominid Predatory
2045 Adaptations.” Journal Article. *Journal of Anthropological Archaeology* 9: 305–51.
- 2046 Stiner, Mary C., Natalie D. Munro, Hiljke Buitenhuis, G. Duru, and M. Ozbasaran. 2022. “An Endemic
2047 Pathway to Sheep and Goat Domestication at Asikli Hoyuk (Central Anatolia, Turkey).” Journal Article.
2048 *Proc Natl Acad Sci U S A* 119 (4). <https://doi.org/10.1073/pnas.2110930119>.
- 2049 Symmons, Robert. 2005. “New Density Data for Unfused and Fused Sheep Bones, and a Preliminary
2050 Discussion on the Modelling of Taphonomic Bias in Archaeofaunal Age Profiles.” Journal Article. *Journal*
2051 *of Archaeological Science* 32 (11): 1691–98. <https://doi.org/10.1016/j.jas.2005.05.011>.
- 2052 Tchernov, Eitan, and Liora Kolska Horwitz. 1991. “Body Size Diminution Under Domestication: Uncon-
2053 scious Selection in Primeval Domesticates.” Journal Article. *Journal of Anthropological Archaeology* 10:
2054 54–75.
- 2055 Team, Stan Development. 2022. “Stan Modeling Language Users Guide and Reference Manual, Version
2056 2.29.” Generic. <https://mc-stan.org>.
- 2057 Thissen, Laurens, Hadi Özbal, Ayla Türkekul Biyik, Fokke Gerritsen, and Rana Özbal. 2010. “The Land
2058 of Milk? Approaching Dietary Preferences of Late Neolithic Communities in NW Anatolia.” Journal
2059 Article. *Leiden Journal of Pottery Studies* 26: 157–72.

- 2060 Todd, Lawrence C. 1983. "The Horner Site: Taphonomy of an Early Holocene Bison Bonebed." Thesis.
- 2061 Trentacoste, A., A. Nieto-Espinet, S. Guimaraes, B. Wilkens, G. Petrucci, and S. Valenzuela-Lamas. 2021.
- 2062 "New Trajectories or Accelerating Change? Zooarchaeological Evidence for Roman Transformation of
- 2063 Animal Husbandry in Northern Italy." Journal Article. *Archaeol Anthropol Sci* 13 (1): 25. <https://doi.org/10.1007/s12520-020-01251-7>.
- 2065 Twiss, Katheryn C. 2019. *The Archaeology of Food*. Book. <https://doi.org/10.1017/9781108670159>.
- 2066 Twiss, Katheryn C., and Nerissa Russell. 2009. "Taking the Bull by the Horns: Ideology, Masculinity, and
- 2067 Cattle Horns at atalhöyük (Turkey)." Journal Article. *Paléorient* 35 (2): 19–32.
- 2068 Uerpmann, Margarethe, and Hans-Peter Uerpmann. 1994. "Animal Bones." Book Section. In *Qala'at Al-*
- 2069 *Bahrain*, edited by Flemming Hojlund and H. Hellmuth Andersen, 1: The Northern City Wall and the
- 2070 Islamic Fortress:417–54. Jutland Archaeological Society Publications. Aarhus: Aarhus University Press.
- 2071 Urbanek, Simon, and Jeffrey Horner. 2022. *Cairo: R Graphics Device Using Cairo Graphics Library for*
- 2072 *Creating High-Quality Bitmap (PNG, JPEG, TIFF), Vector (PDF, SVG, PostScript) and Display (X11*
- 2073 *and Win32) Output*. <https://CRAN.R-project.org/package=Cairo>.
- 2074 Vigne, Jean-Denis. 2011. "Le Mouton (Ovis Aries)." Book Section. In *Shillourokambos: Un Établissement*
- 2075 *Néolithique Pré-Céramique à Chypre: Les Fouilles Du Secteur i*, edited by Jean Guilaine, François Briois,
- 2076 and Jean-Denis Vigne, 1021–38. Paris: Editions Errance.
- 2077 Watkins, Trevor. 1996. "Excavations at Pinarbaşı: The Early Stages." Book Section. In *On the Surface:*
- 2078 *Çatalhöyük 1993-1995*, edited by Ian Hodder, 47–57. Cambridge: McDonald Institute for Archaeological
- 2079 Research.
- 2080 Weinstock, Jacobo. 2006. "Environment, Body Size and Sexual Dimorphism in Late Glacial Reindeer."
- 2081 Book Section. In *Recent Advances in Ageing and Sexing Animal Bones*, edited by Deborah Ruscillo,
- 2082 247–53. Oxford: Oxbow Books.
- 2083 Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>.
- 2085 Wickham, Hadley, and Jennifer Bryan. 2022. *Readxl: Read Excel Files*. <https://CRAN.R-project.org/package=readxl>.
- 2087 Widga, Chris. 2006. "Niche Variability in Late Holocene Bison: A Perspective from Big Bone Lick, KY."
- 2088 Journal Article. *Journal of Archaeological Science* 33 (9): 1237–55. <https://doi.org/10.1016/j.jas.2005.12.011>.
- 2090 Wolfhagen, Jesse. 2020. "Re-Examining the Use of the LSI Technique in Zooarchaeology." Journal Article.
- 2091 *Journal of Archaeological Science* 123: 105254. <https://doi.org/10.1016/j.jas.2020.105254>.
- 2092 Wolfhagen, Jesse, and Max D. Price. 2017. "A Probabilistic Model for Distinguishing Between Sheep and

- 2093 Goat Postcranial Remains.” Journal Article. *Journal of Archaeological Science: Reports* 12: 625–31.
2094 <https://doi.org/10.1016/j.jasrep.2017.02.022>.
- 2095 Wolfhagen, Jesse, Katheryn C. Twiss, Jacqui A. Mulville, and G. Arzu Demirergi. 2021. “Examining
2096 Caprine Management and Cattle Domestication Through Biometric Analyses at Çatalhöyük East (North
2097 and South Areas).” Book Section. In *Peopling the Landscape of Çatalhöyük: Reports from the 2009–2017
2098 Seasons*, edited by Ian Hodder, 181–98. Los Angeles: Cotsen Institute of Archaeology Press.
- 2099 Wolverton, Steve. 2008. “Harvest Pressure and Environmental Carrying Capacity: An Ordinal-Scale Model
2100 of Effects on Ungulate Prey.” Journal Article. *American Antiquity* 73 (2): 179–99.
- 2101 Wright, Elizabeth, and Sarah Viner-Daniels. 2015. “Geographical Variation in the Size and Shape of the
2102 European Aurochs (*Bos Primigenius*).” Journal Article. *Journal of Archaeological Science* 54: 8–22.
2103 <https://doi.org/10.1016/j.jas.2014.11.021>.
- 2104 Xie, Yihui. 2015. *Dynamic Documents with R and Knitr*. 2nd ed. Boca Raton, Florida: Chapman;
2105 Hall/CRC. <https://yihui.org/knitr/>.
- 2106 ———. 2022. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. <https://yihui.org/knitr/>.
- 2107 Zeder, Melinda A. 2012. “Pathways to Animal Domestication.” Book Section. In *Biodiversity in Agriculture:
2108 Domestication, Evolution, and Sustainability*, edited by P. Gepts, T. R. Famula, and R. L. Bettinger,
2109 227–59. Cambridge: Cambridge University Press.
- 2110 Zeder, Melinda A., and Brian Hesse. 2000. “The Initial Domestication of Goats (*Capra Hircus*) in the Zagros
2111 Mountains 10,000 Years Ago.” Journal Article. *Science* 287 (5461): 2254–57.
- 2112 Zeder, Melinda A., and Ximena Lemoine. 2020. “A Method for Constructing Demographic Profiles in Sus
2113 Scrofa Using Logarithm Size Index Scaling.” Journal Article. *Journal of Archaeological Science* 116.
2114 <https://doi.org/10.1016/j.jas.2020.105115>.
- 2115 Zhu, Hao. 2021. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. <https://CRAN.R-project.org/package=kableExtra>.
- 2116 Zohary, D., Eitan Tchernov, and Liora Kolska Horwitz. 1998. “The Role of Unconscious Selection in the
2117 Domestication of Sheep and Goats.” Journal Article. *Journal of Zoology, London* 245: 129–35.