

# Investigating the Effects of Co-housing with Animals and Age on Gut Microbiome Diversity

Nathan Kolbow  
nkolbow@wisc.edu

## 1 Summary

The SELEVER project is a five-year controlled study in Burkina Faso that seeks to evaluate the impacts and costs of nutritional and agricultural interventions on the health of women and children in the area. To gather baseline measurements prior to the interventions, researchers went to households selected for the study to survey a range of information about the participants as well as gut microbiome alpha diversity measures for each individual calculated from sequences stool samples.

Aside from the stool samples, the surveyed information includes general participant information, general information about the participant's parents and siblings, information concerning how clean the household is overall, and the quantities of various livestock owned by the household. Our primary interests for this report are how gut microbiome diversity in these individuals changes with age, and whether co-housing with animals has an effect on gut microbiome diversity. Our main findings are as follows:

1. Gut microbiome diversity increases with age and begins leveling off at around 53 months old.
2. There is no evidence in this data for an effect of co-housing with animals. This may be because only 3% of individuals in this dataset do not own animals, and because, despite not owning them, they are likely constantly in close proximity with the livestock of others in their village.

## 2 Introduction

### 2.1 Background

Burkina Faso is a West African country of around 22 million people and is one of the least developed countries in the world according to L’Institut national de la statistique et de la Démographie, Burkina Faso’s official statistical institute. Additionally, the residents of Burkina Faso have very low dietary diversification which plays a major role in their high rates of anemia in women, child stunting, child wasting, and child anemia, according to the International Food Policy Research Institute.

The SELEVER intervention, funded by the Bill and Melinda Gates foundation, is a five year controlled intervention study that will examine the effect various nutritional and agricultural interventions on the diets, health, and nutritional status of women and children in Burkina Faso. To this end, researchers will provide training and education on value chain services such as flock management, financing, and vaccinations, promote improved diets at key stages of life, and empower and strengthen women’s groups at a community level. This last point is important because women in these communities are often in charge of the children, livestock, finances, the household, etc., so empowering them should lead to positive downstream effects for entire communities.

### 2.2 Experimental Design

Three regions in Burkina Faso were considered for this study, within there are 79 total communes. From these communes, 60 were selected for the program with 120 rural villages total across all 60 communes. Detailed inclusion and exclusion criteria are described in the study’s protocol paper<sup>1</sup>. Treatments and controls were assigned to villages, but, in this report, we only have access to data collected at baseline, prior to any intervention. So, we are not interested in treatment assignments, but we are interested in potential random effects across regions, communes, and villages.

Prior to the intervention, researchers visited the selected households and gathered baseline information on one of the household’s children. This information included metadata discussed in Subsection 2.3, as well as stool samples that were used to calculate three gut microbiome alpha diversity measures for each individuals, namely Shannon’s entropy, Faith’s phylogenetic diversity (PD), and the total number of taxa in each sample (observed features).

### 2.3 The Dataset

There are 3 alpha diversity measures and 40 pieces of metadata connected to each of 778 individuals in our dataset. This metadata includes the region, commune and village of the individual, as well as general

---

<sup>1</sup>Gelli, A., Becquey, E., Ganaba, R. *et al.* Improving diets and nutrition through an integrated poultry value chain and nutrition intervention (SELEVER) in Burkina Faso: study protocol for a randomized trial. *Trials* 18, 412 (2017). <https://doi.org/10.1186/s13063-017-2156-4>

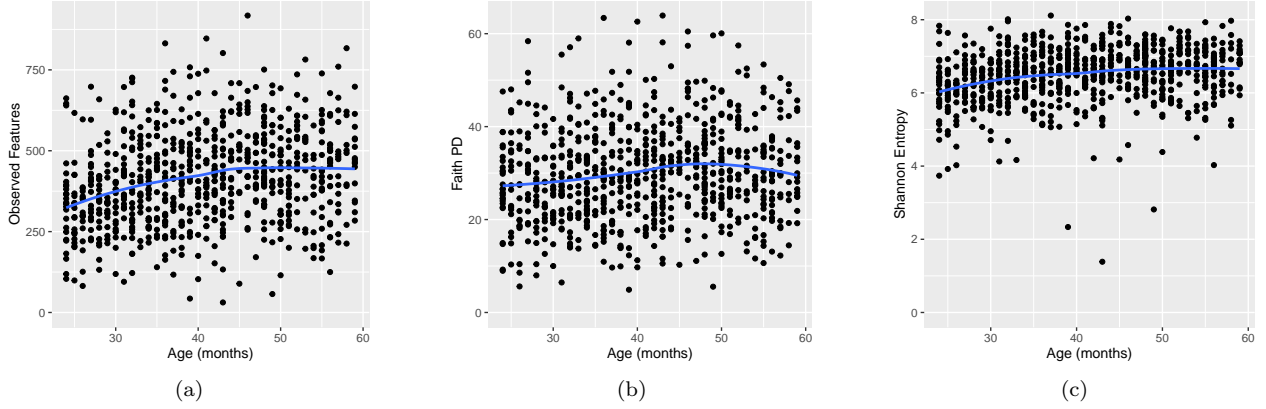


Figure 1: Each alpha diversity measure plotted against age with a trend line overlaid on each.

information such as age, height and weight, family information such as parent’s age, parent’s education and number of siblings, cleanliness information such as whether the house has a clean toilet and whether there is animal feces in the household, and information on the number of various livestock owned by the household. Nine variables have missing data, including each of the four  $z$ -score variables<sup>2</sup>, the individual’s temperature, hemoglobin concentration, and malaria status, as well as the mother and father’s ages. Early analyses indicated that these variables had no impact on the alpha diversity measures, and none of these are primary variables of interest, so the analyses that we discuss did not use these variables.

Each of the alpha diversity measures are plotted against age in Figure 1. We see that each measure seems to have a positive relationship with age, but that each also has a lot of variation around this trend. The correlations between Shannon’s entropy and observed features, Shannon’s entropy and Faith’s PD, and observed features and Faith’s PD are 0.856, 0.627, and 0.764 each. These are very high, so we do not expect our findings to vary significantly between each of the measures.

### 3 Methods

#### 3.1 Outlier and Highly Influential Point Detection

Highly influential points are data points that impact model coefficients significantly more than other data points. Thus, these individuals, along with outliers, can drastically alter the conclusions that we draw from analyses if they are included in our models, so they should be identified and removed.

Plotting our metadata, as in Figure 2, immediately reveals several outliers. These individuals do not differ from the norm in any of their other metadata, so including them in our analyses would skew our results away from any true underlying trends.

<sup>2</sup>These are complicated variables generated by SELEVER that do not have any direct interpretation, but generally speaking each a function of cleanliness and two of age, height, and weight.

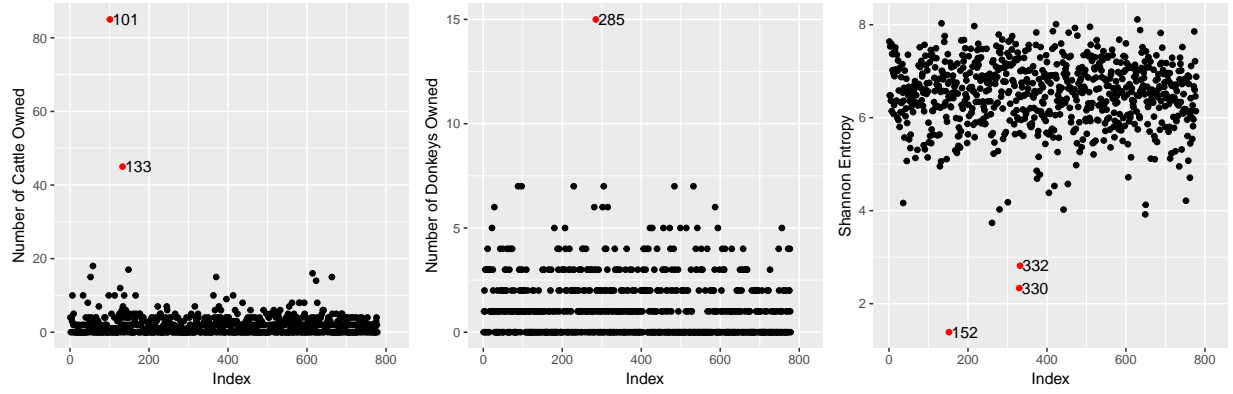


Figure 2: Number of cattle owned, number of donkeys owned, and Shannon’s entropy plotted against each individual’s index in the dataset with outliers highlighted and labeled.

In order to investigate which points are highly influential, we need to fit a model. The experimental design of this study lends itself very nicely to a mixed effects model, so we fit a stepwise mixed effects model<sup>3</sup> for each of our alpha diversity measured against all of our metadata. Figure 3 shows the diagnostics for such a model with Shannon’s entropy as the response variable. The same points were identified when observed features and Faith’s PD were used as well. All of the points highlighted in Figures 2 and 3 have been removed from subsequent analyses.

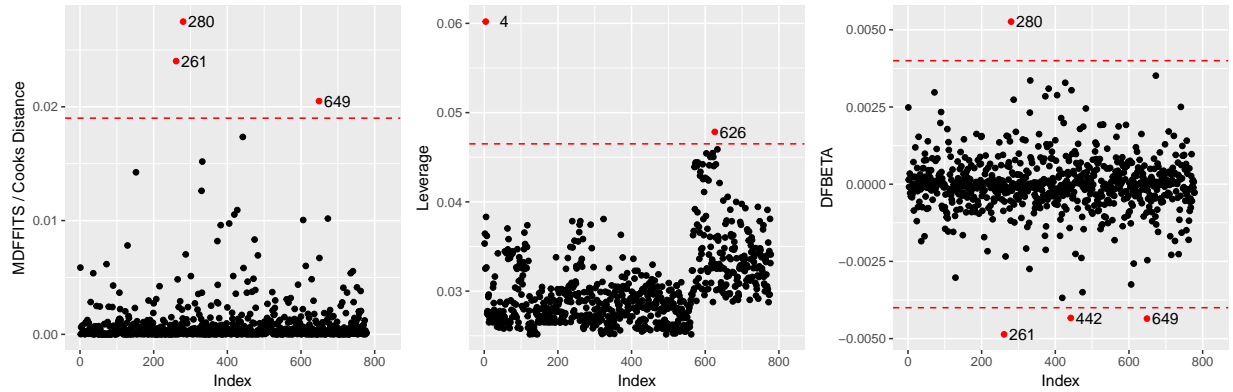


Figure 3: Influential points detected with Cook’s distance, leverage, and DFBETA as analyzed on a stepwise reduced mixed effects model with shannon entropy as the response variable.

### 3.2 Stepwise Mixed Effects Models

The experimental design of this study lends itself very nicely to a mixed effects model, and we are interested in which of our predictors are important, so stepwise mixed effects models are a very natural choice. We

<sup>3</sup>A stepwise mixed effects model is a model that first fits a mixed effects model with all of the predictors in the metadata, then iteratively adds or removes the predictor that leads to the most model AIC improvement until the model’s AIC cannot be improved any further.

fit one stepwise mixed effects model for each alpha diversity measure, resulting in three final models. Each model began with random effects for region, commune, and village, fixed effects for every other predictor in our dataset, and a quadratic term for age. Then, the model was iteratively adjusted, one variable at a time, until it could not be improved any further. Additionally, appropriate variance-stabilizing transformations (VSTs) were applied to each alpha diversity measure. VSTs are transformations to the response variable (in this case each of the alpha diversity measures) that lead to residuals that are more in line with model assumptions, thus improving a model’s fit and accuracy.

The VSTs and resulting model coefficients for each of the 3 models are given in Table 1; “-” indicates that the term was not found to be significant in modelling the given alpha diversity measure. To illustrate how the table should be read, the final model for Shannon’s entropy would be:  $S^2 = 43.855 + 0.9118 \text{ age} - 0.008616 \text{ age}^2$ , with a commune random effect standard error estimate of 2.186 and a residual standard error estimate of 8.046. These standard error estimates mean that given two individuals of the same age, the standard error of the difference in Shannon’s entropy squared between the individuals is 8.046 if they are from the same commune, and  $8.046 + 2.186 = 10.232$  if they are from different communes.

	Shannon’s Entropy	Faith’s PD	Observed Features
VST	$Y^2$	$\sqrt{Y}$	$Y^{2/3}$
Intercept	21.0149	4.901	21.2325
Age (months)	0.9118	0.01236	1.4081
Age <sup>2</sup> (months)	-0.008616	-	-0.01353
$\sigma_{\text{Commune}}$	2.186	-	-
$\sigma_{\text{Village}}$	-	-	3.241
$\sigma_{\text{Residual}}$	8.046	25.518	11.735

Table 1: Model coefficients and variance-stabilizing transformations for the stepwise mixed effects model corresponding to each alpha diversity measure.

Model diagnostics are provided in Figure 5 in the appendix; they all indicate that our models fit the data very well.

### 3.3 Mixed Effect Lasso Regression

Lasso regression is a popular model for variable selection, and there are various extensions to this model that provide support for random effects. In these models, a penalty is applied across the model’s coefficients that causes the coefficients of less important variables to move towards 0. This penalty is applied through a hyperparameter  $\lambda$  that is typically selected through cross-validation, where many  $\lambda$  values are tested and an ideal one is selected based off of model AIC or BIC. We perform this cross-validation for each of the 3 alpha diversity measures, with the same fixed effects, random effects, and VSTs as in Subsection 3.2.

The resulting models for Shannon’s entropy and observed features have the same terms as those in Subsection 3.2, with coefficients agreeing in magnitude and direction. The resulting model for Faith’s PD,

though, did not have any non-zero coefficients, whereas its model in Subsection 3.2 had a non-zero coefficient for age. In summary, the results of this method mostly agreed with those in Subsection 3.2, but were more conservative when it came to Faith’s PD.

### 3.4 Decision Trees

Decision trees naturally model complex interactions between predictors, so they are another choice for modelling our data. Using the `partykit` package in R, no matter how the decision tree’s hyperparameters are tuned, even allowing for infinite depth and leaves with only 1 sample, Figure 4 is the most complex tree produced for any of the alpha diversity measures<sup>4</sup>. This indicates that there is no signal in our data for alpha diversity from anything other than age, which agrees with our findings in Subsections 3.2 and 3.3.

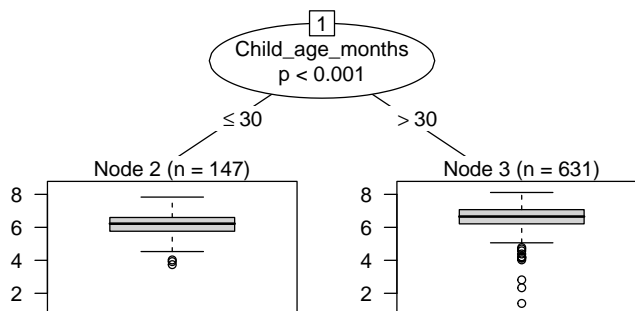


Figure 4: Decision tree fit for Shannon’s entropy on all metadata variables.

### 3.5 Fixed Age Range Analyses

There appears to be no signal for alpha diversity in our entire dataset from anything other than age, but there may be effects from other variables that aren’t showing up because they either vary across ages or are masked by age. To investigate this, we perform a sliding window analysis using the methods described in Subsections 3.2, 3.4, and 3.3.

In a sliding window analysis, we specify a window size, step size, and “step” this window across the range of ages, performing independent analyses at each step. For example, we have individuals ranging from 24 to 59 months old, so with a window size of 6 months and a step size of 4 months, we would perform the analyses described in Subsections 3.2, 3.4, and 3.3 first on all individuals 24-30 months old, then on all individuals 28-34 months old, then 32-38 months old, and so on until we perform the last analysis on all individuals 56-62 months old. If a variable consistently appears across these analyses, then it may be important and would warrant further investigation.

<sup>4</sup>`partykit` cannot handle categorical variables with more than 31 levels, so the commune and village IDs could not be included in the fit.

Performing these analyses across many combinations of window sizes and step sizes ranging from 2 to 12 and 1 to 12, respectively, on each of the three alpha diversity measures reveals that no variable in the dataset is consistently important. The decision trees generated in each window are *empty* around 96% of the time. For the stepwise mixed effects models and mixed effect Lasso regression models, no variable appeared in a given analysis more than 30% of the time, variables always appeared sporadically, e.g. no variable was consistently significant for all ages under 36 months old, and the mixed effects Lasso regression models and stepwise mixed effects models never agreed with one another.

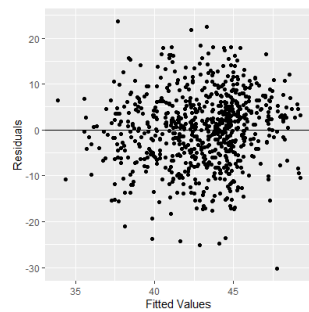
## 4 Conclusions

None of our methods found a significant effect for any of the animal-related metadata variables on any of the gut microbiome alpha diversity measures. This could be because co-housing with animals has no effect on gut microbiome, but there are many other possible explanations for this. One such explanation is that only 24 individuals in our dataset do not own any animals while the remaining 754 do. If there is a real effect on gut microbiome alpha diversity from co-housing with animals, it would be very difficult to extract that information when there is such a drastic imbalance. Additionally, livestock are a significant part of these individuals economics, so even if these 24 individuals don't live with animals, they still spend a significant amount of time around them, so it is reasonable to assume that they would experience similar effects.

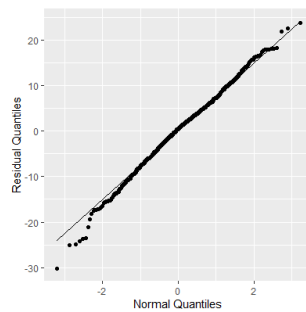
Our stepwise mixed effects models and mixed effects Lasso models found positive associations between age and each of the alpha diversity measures. A quadratic term was estimated for age for Shannon's entropy and observed features, but this is more likely because these measures increase more slowly as individuals age, and then level off at around 53 months old. Age squared was not significant for Faith's PD, but it is unclear whether this is because Faith's PD does not level off at around 53 months old, or because Faith's PD is much more variable than the other measures, as summarized by the random effect and residual error estimates in Table 1.

We conclude that there is a clear positive relationship between age and each of the given alpha diversity measures, that Shannon's entropy and observed features level off at around 53 months old, and that Faith's PD might level off at around this age or later, but we need more data to verify this. Additionally, we conclude that we cannot determine from these data whether co-housing with animals affects gut microbiome alpha diversity.

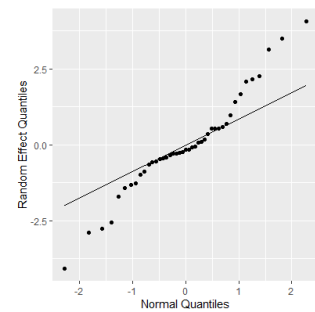
# Appendix



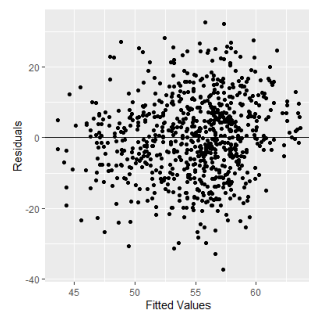
(a) Shannon's entropy SMEM fitted values vs. residuals.



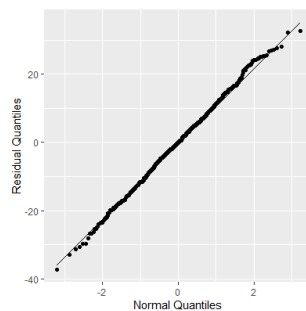
(b) Shannon's entropy SMEM QQ-plot of residuals.



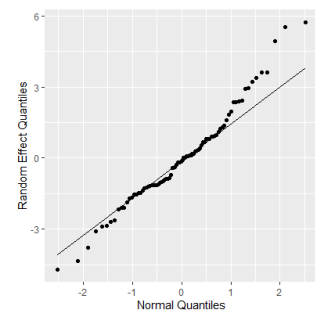
(c) Shannon's entropy SMEM QQ-plot of random effects.



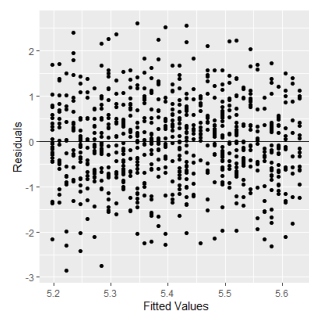
(d) Observed features SMEM fitted values vs. residuals.



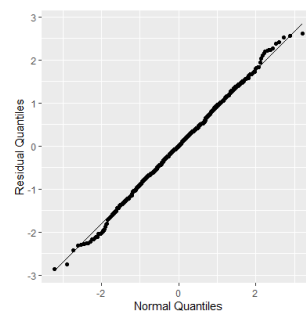
(e) Observed features SMEM QQ-plot of residuals.



(f) Observed features SMEM QQ-plot of random effects.



(g) Faith's PD SMEM fitted values vs. residuals.



(h) Faith's PD SMEM QQ-plot of residuals.

Figure 5: Diagnostic plots for all of the stepwise mixed effects models (SMEMs).