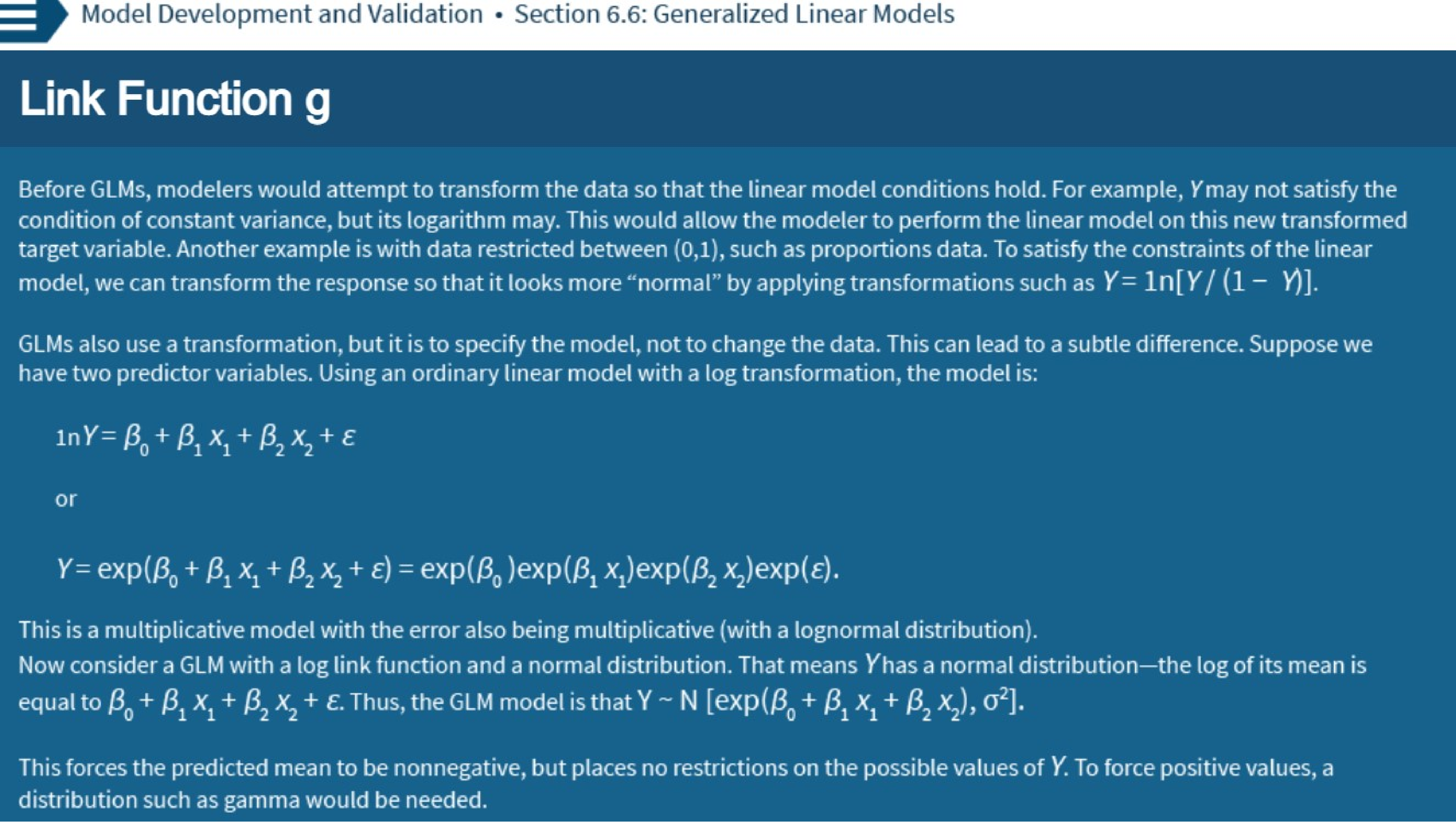

Link Function g

Before GLMs, modelers would attempt to transform the data so that the linear model conditions hold. For example, Y may not satisfy the condition of constant variance, but its logarithm may. This would allow the modeler to perform the linear model on this new transformed target variable. Another example is with data restricted between (0,1), such as proportions data. To satisfy the constraints of the linear model, we can transform the response so that it looks more "normal" by applying transformations such as $Y = 1\text{n}[Y / (1 - Y)]$.

GLMs also use a transformation, but it is to specify the model, not to change the data. This can lead to a subtle difference. Suppose we have two predictor variables. Using an ordinary linear model with a log transformation, the model is:

$$1\text{n}Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

or

$$Y = \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon) = \exp(\beta_0)\exp(\beta_1 x_1)\exp(\beta_2 x_2)\exp(\varepsilon).$$

This is a multiplicative model with the error also being multiplicative (with a lognormal distribution).
Now consider a GLM with a log link function and a normal distribution. That means Y has a normal distribution—the log of its mean is equal to $\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$. Thus, the GLM model is that $Y \sim N\,[\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2), \sigma^2]$.

This forces the predicted mean to be nonnegative, but places no restrictions on the possible values of Y. To force positive values, a distribution such as gamma would be needed.