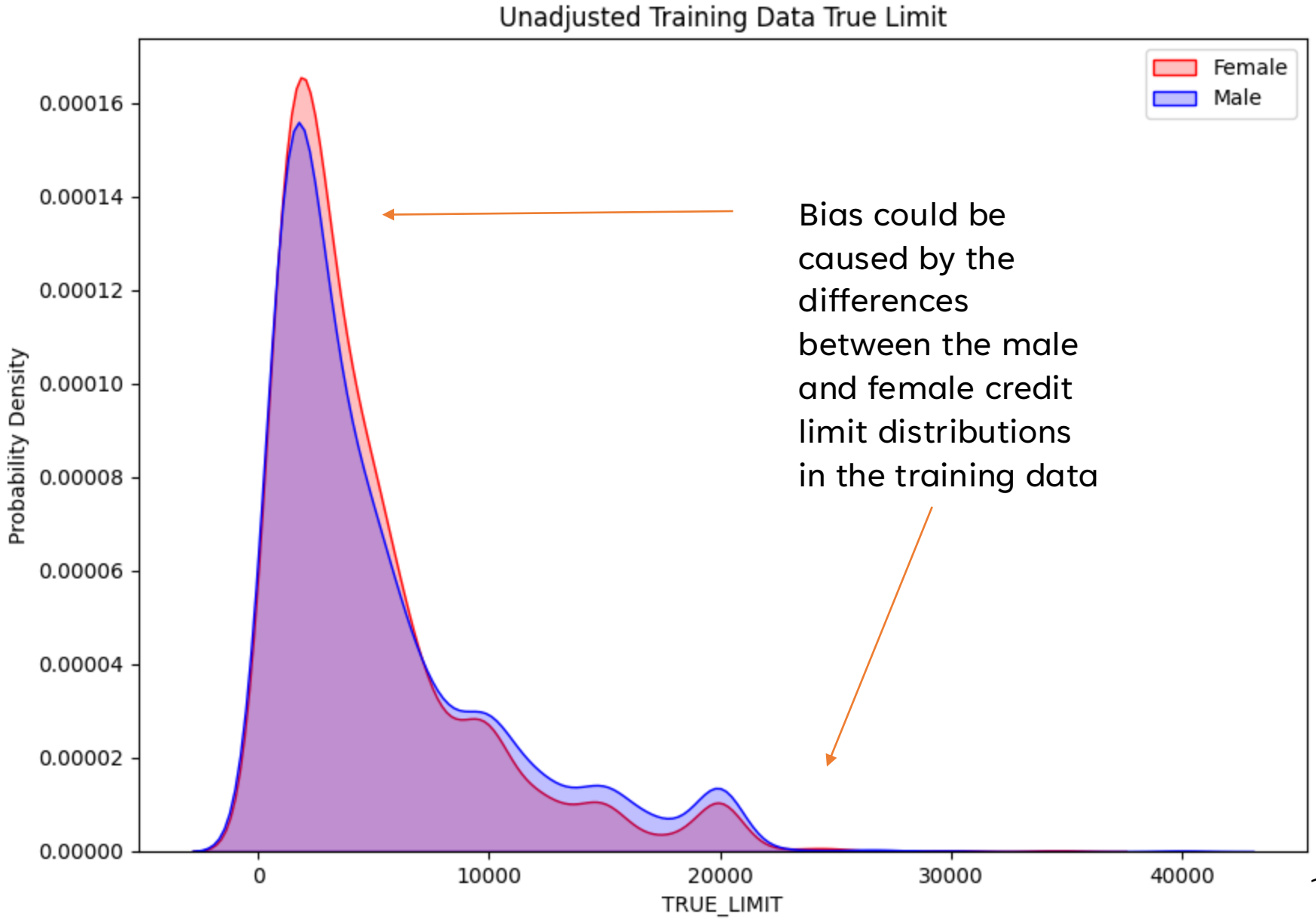
PREVENTING GENDER BIAS IN CREDIT LIMIT ALGORITHMS

TRAIN & EVALUATE RIDGE AND LASSO MODELS

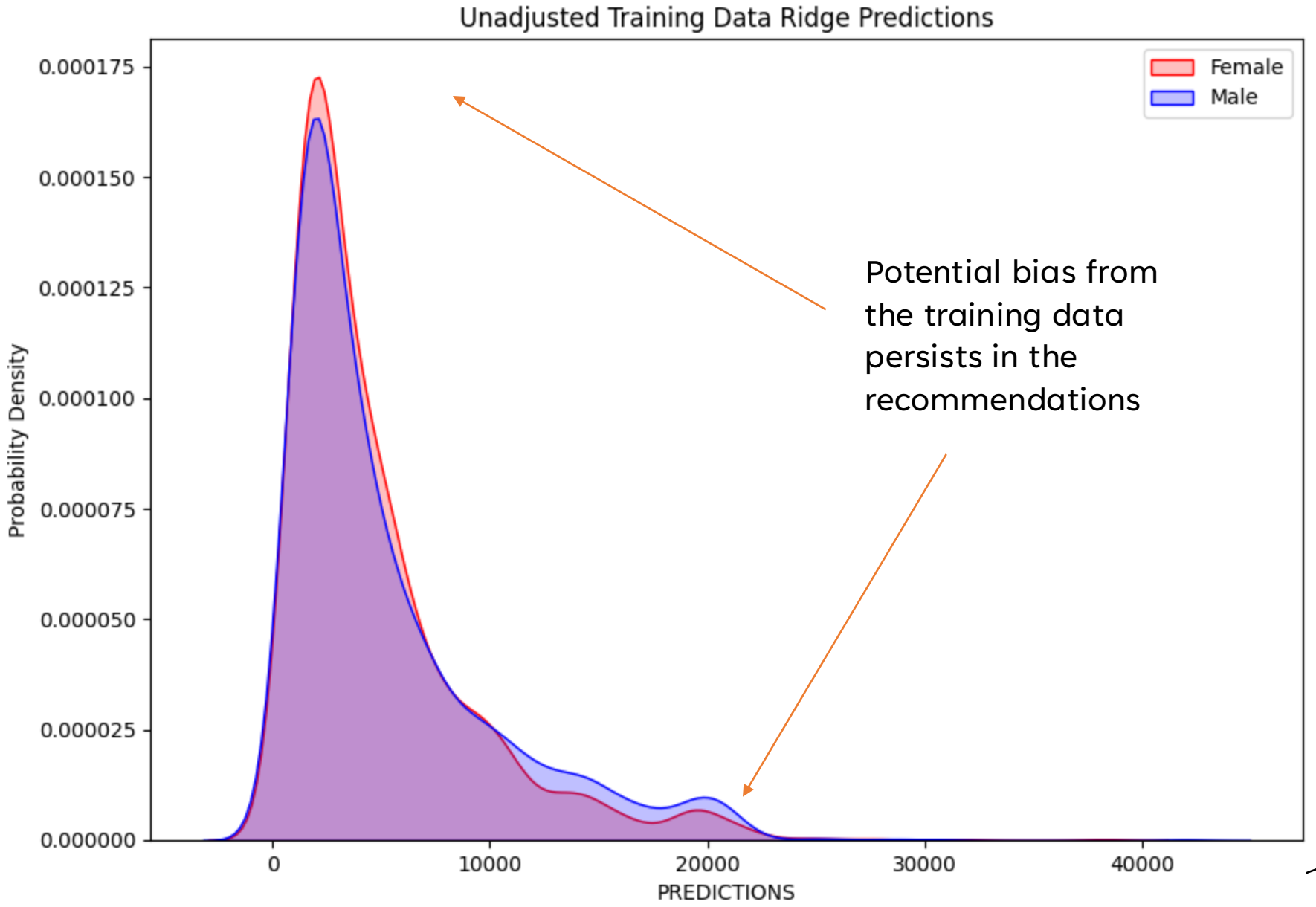
1. Train models to evaluate performance
2. Create Kernel Density Estimate (KDE) plot to understand the probability density of credit limits for each gender in the training data
3. Use the KDE plot to understand how the probability distribution of the training data impacts predictions

KDE – TRAINING DATA CREDIT LIMITS



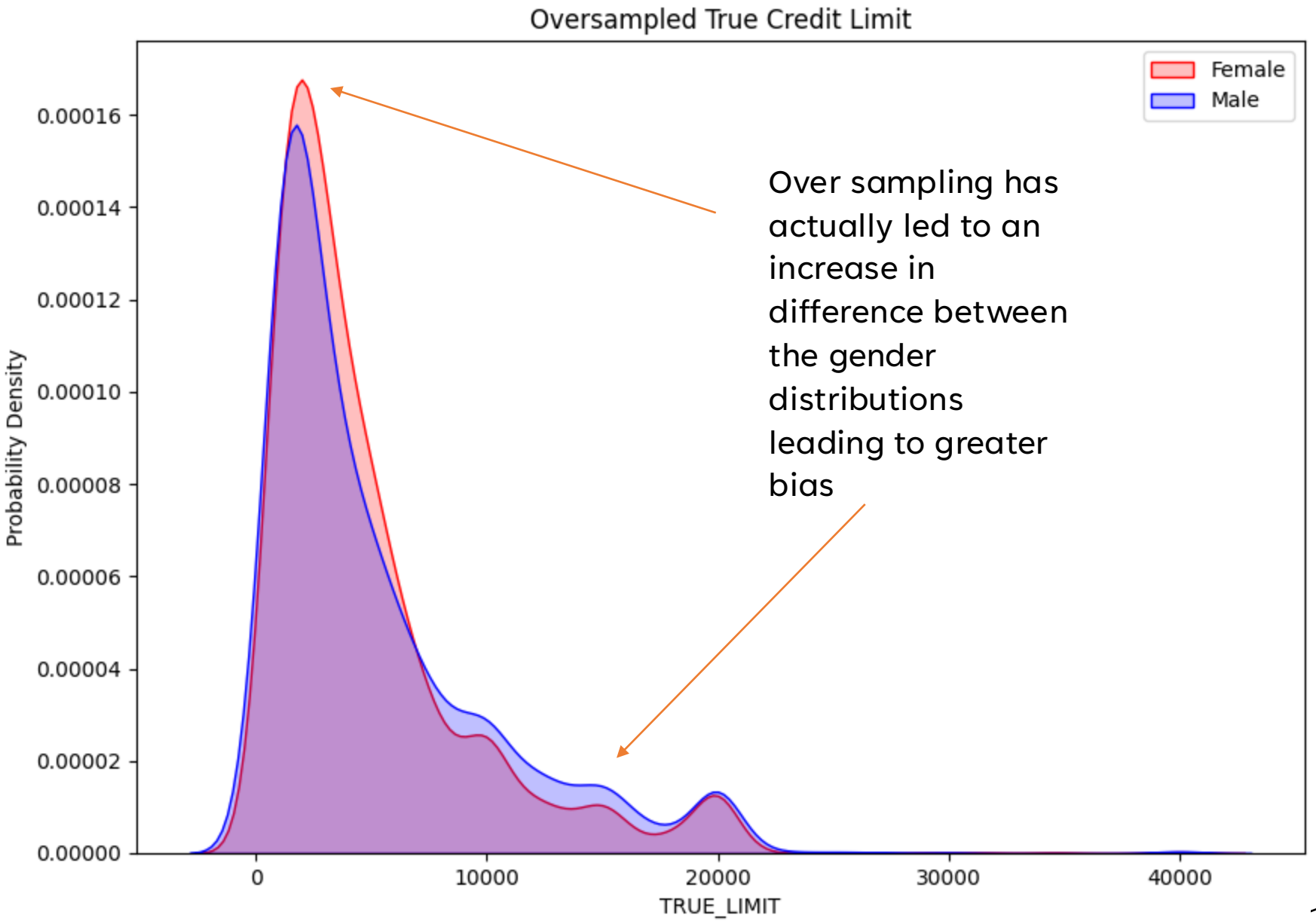
Bias could be caused by the differences between the male and female credit limit distributions in the training data

KDE – RIDGE MODEL RECOMMENDATIONS (BIASED)



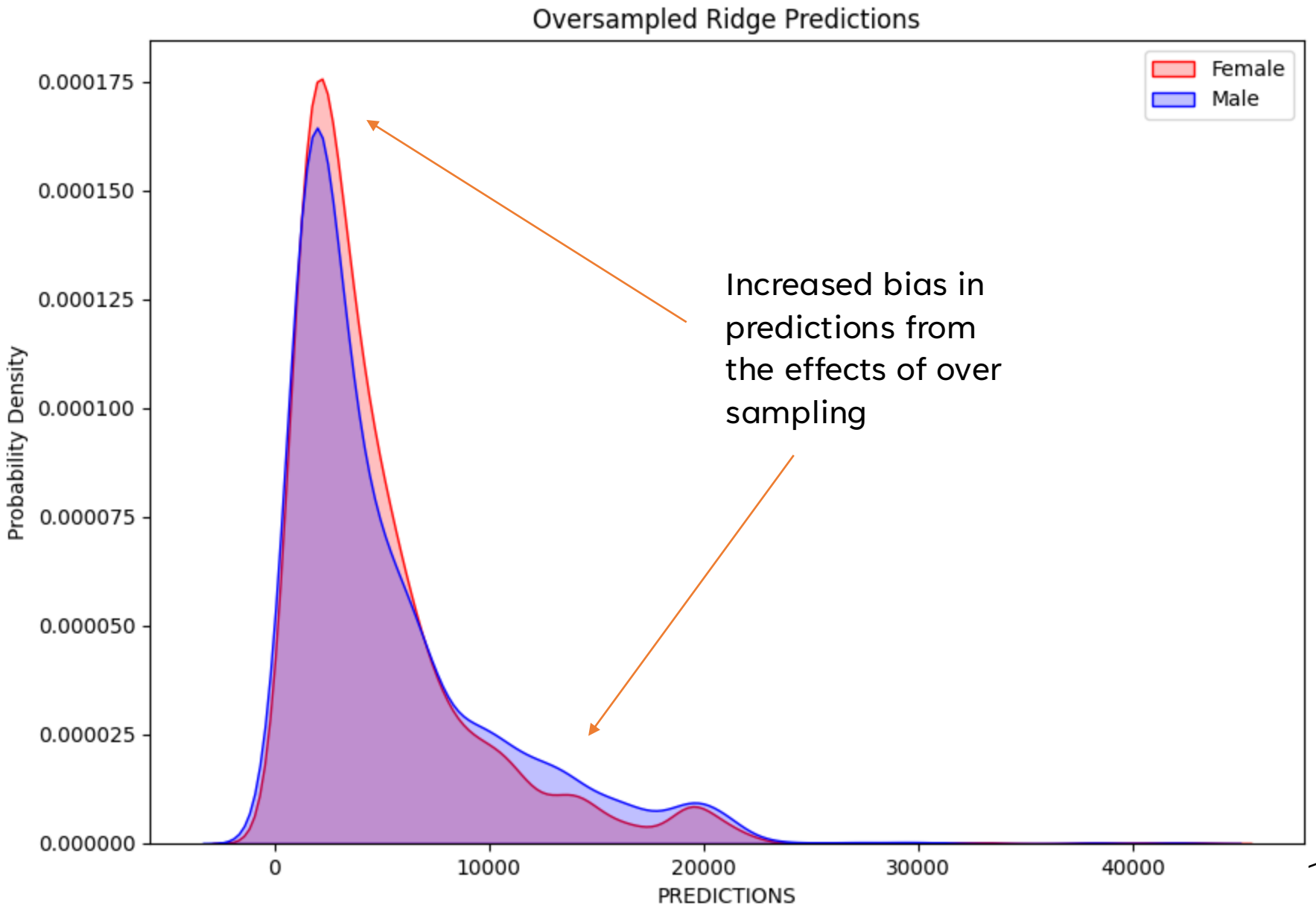
Potential bias from the training data persists in the recommendations

KDE – TRAINING DATA OVER SAMPLED

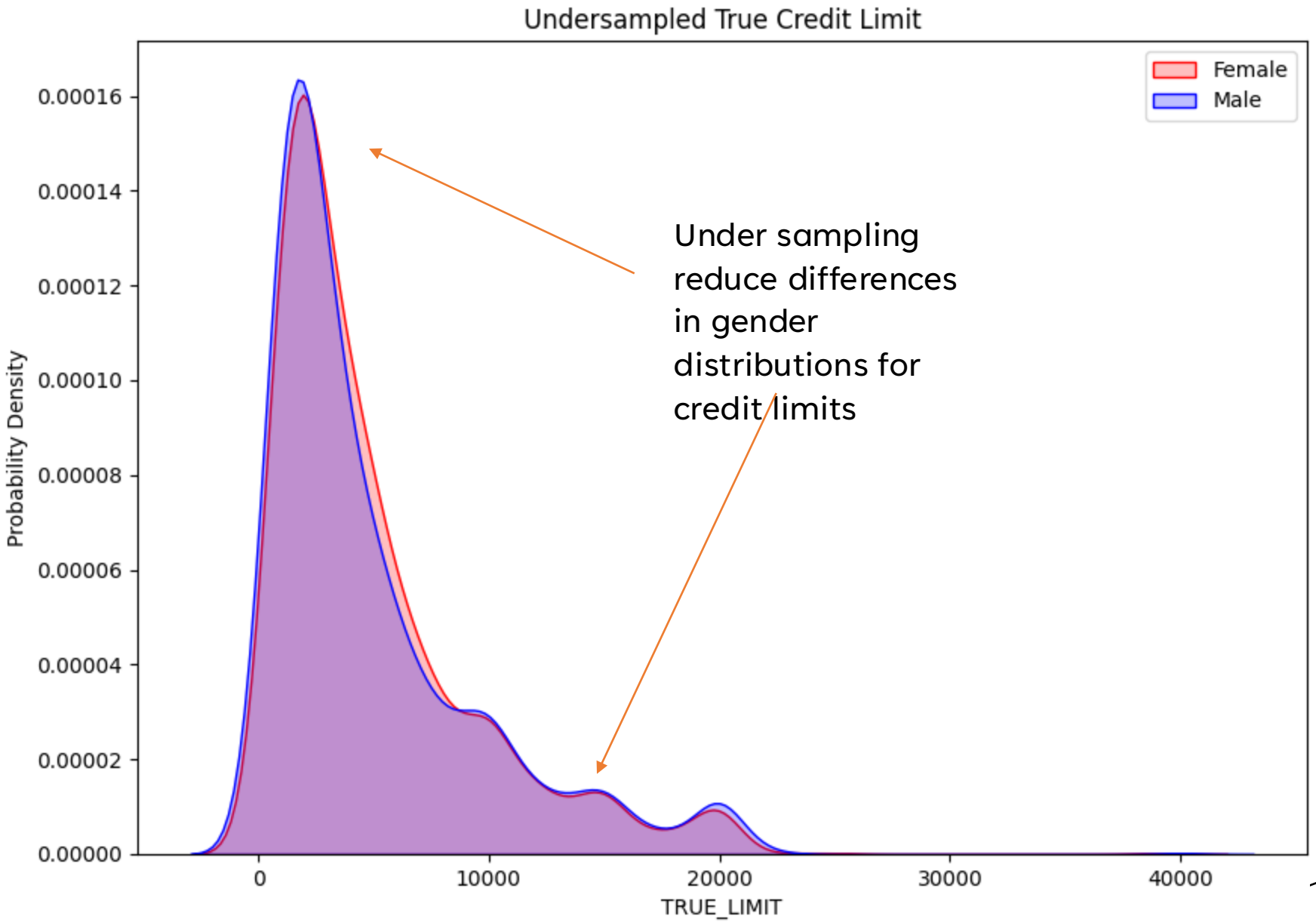


Over sampling has actually led to an increase in difference between the gender distributions leading to greater bias

KDE – RIDGE MODEL RECOMMENDATIONS (OVER SAMPLED)



KDE – TRAINING DATA TRUE LIMIT (UNDER SAMPLED)



KDE – RIDGE PREDICTION (UNDER SAMPLED)