Working Title: Comparison of Reject Inference Techniques using a simulation study

Credit Scorecards are commonly built using data available within an organizations transactional database. Such data, however, will only contain information for those applicants who were 'accepted' or previously awarded credit by the organization; data will not be available for those applicants who were 'rejected'. The use of *reject inference* to adjust credit scorecard models for the missing data represented by rejected loan applications is common practice and several approaches are used in today's financial industry. In this paper, we investigate the use of mixture models as an alternative approach for reject inference. In this approach, the underlying probability density function (PDF) for the total population of all applicants is modeled as a weighted average of the PDFs of the 'accepted' population and the 'rejected' population. The unknown parameters in this model will then be the mixing parameters of this weighted average. To estimate, these parameters, we use the EM algorithm (Laird, Rubin, and Dempster, 1977) wherein the data associated with the 'rejected' applicants is treated as missing completely at random (Rubin, 1976). Simulated data derived from actual case studies are used to assess the effectiveness of the mixture model approach and other reject inference approaches.