

# Specification and Informational Issues in Credit Scoring \*

NICHOLAS M. KIEFER <sup>†</sup> and C. ERIK LARSON <sup>‡</sup>

## Abstract

Lenders use rating and scoring models to rank credit applicants on their expected performance. The models and approaches are numerous. We explore the possibility that estimates generated by models developed with data drawn solely from extended loans are less valuable than they should be because of selectivity bias. We investigate the value of “reject inference” – methods that use a rejected applicant’s characteristics, rather than loan performance data, in scoring model development. In the course of making this investigation, we also discuss the advantages of using parametric as well as nonparametric modeling. These issues are discussed and illustrated in the context of a simple stylized model.

---

\*Disclaimer: *Portions of this paper were completed while Larson was employed by the Office of the Comptroller of the Currency. Views expressed herein are those of the authors, and do not necessarily represent the views of Fannie Mae, the Office of the Comptroller of the Currency, the U.S. Department of the Treasury, or their staffs.*

Acknowledgements: *We thank colleagues and workshop participants at Cornell University, the OCC and Syracuse University, as well as the referees, for helpful comments.*

*Received:* January 9, 2006; *Accepted:* October 3, 2006.

*Key words and phrases:* Logistic regression, specification testing, risk management, nonparametrics, reject inference.

*JEL Classification:* C13, C14, C52, G11, G32

<sup>†</sup>*Mailing Address:* Departments of Economics and Statistical Sciences, Cornell University; and Risk Analysis Division, Office of the Comptroller of the Currency, US Department of the Treasury. *Email:* nmk1@cornell.edu

<sup>‡</sup>Fannie Mae. *E-mail:* c\_erik\_larson@fanniemae.com