CRA Insights:
Financial Economics

CRA Charles River Associates

February 2014

Evaluating the fair lending risk of credit scoring models

Credit scoring models and other automated decision tools can limit the potential for credit applicants to be treated differently on a legally prohibited basis, whether deliberately or inadvertently, by reducing or eliminating the amount of judgment and discretion in credit decisions.[1] As a result, credit scoring models can be valuable tools to manage fair lending risk. Financial regulatory guidance has existed for some time on the use and implementation of credit scoring models.[2] The guidance notes how fair lending compliance risk can arise if scoring models are not properly developed, tested, implemented, managed, and monitored.

Conceptually, the potential fair lending risks associated with credit scoring models are clear. However, assessing, quantifying, and weighing those risks is a complex technical endeavor. We provide an overview of key considerations in evaluating credit scoring models for fair lending risk. First, we provide a brief review of the main fair lending considerations under the Equal Credit Opportunity Act (ECOA) that apply to credit scoring models. Next, we discuss some common fair lending risk issues and testing considerations related to scoring models.[3,4]

Sources of potential liability under the ECOA

All three theories of liability for discrimination under the ECOA may apply in the context of credit scoring models:[5]

- *Overt discrimination* refers to intentional discrimination on the basis of a legally prohibited factor.
- *Disparate treatment* refers to cases in which similarly situated individuals receive different treatment, and the only explanation for that difference, after legitimate factors have been accounted for, is a prohibited factor.

---

1 The prohibited factors under the Equal Credit Opportunity Act are: race, color, religion, national origin, sex, marital status, age (provided the applicant has the capacity to contract), an applicant's receipt of income from any public assistance program, and an applicant's exercise in good faith of rights under the Consumer Credit Protection Act. 15 USC 1601, *et seq*.

2 See OCC Bulletin, 97-24. Accessed February 7, 2014, http://ithandbook.ffiec.gov/media/resources/3672/occ-bl-97-24_credit_scor_models.pdf

3 Regulatory guidance regarding the examination of credit scoring systems for ECOA compliance can be found in "Considering Automated Underwriting and Credit Scoring," Part II of the Appendix to "Interagency Fair Lending Examination Procedures," Consumer Financial Protection Bureau Examination Manual v.2 (October 2012). Accessed February 13, 2014, http://www.consumerfinance.gov/guidance/supervision/manual/.

4 This brief is not intended to provide comprehensive coverage of all aspects of the ECOA relevant to credit scoring models, nor is it intended to provide legal opinions or advice.

5 See Consumer Financial Protection Bureau Bulletin 2012-04 (Fair Lending), April 18, 2012.

- *Disparate impact* refers to situations in which decision rules or model variables are facially neutral, but nevertheless have a disproportionate adverse impact on the basis of a prohibited factor, *in effect*, which cannot be justified by a business necessity and despite the availability of an equally effective, but less discriminatory, alternative.[6]

"Empirically derived, demonstrably and statistically sound"

Regulation B, which implements the ECOA, also prescribes the general standards that a credit scoring system must meet to qualify as an "empirically derived, demonstrably and statistically sound, credit scoring system." Forms of credit analysis that do not meet these standards are considered to be "judgmental" systems. Even a statistically-based credit scoring model might be considered "judgmental" if it does not meet the "empirically derived, demonstrably and statistically sound" (EDDSS) standard.

The distinction between a judgmental system and an EDDSS scoring system is important. Creditors that use an EDDSS scoring system may take applicant age directly into account as a predictive variable in a scoring model (provided that elderly consumers are scored at least as favorably as younger consumers) or may segment the population into multiple scorecards based on age, as long as elderly applicants are included in a narrow age range.[7] Judgmental systems may take age into account only to determine minimum legal requirements for a credit obligation, or to treat elderly applicants more favorably than younger applicants.

A credit scoring system must satisfy all of the following criteria to be classified as EDDSS:

- *Empirical:* Based on a rigorous statistical analysis and derives empirical ways to distinguish between more and less creditworthy consumers, using data for applicants who applied for credit within a reasonable preceding period of time.
- *Business justified:* Developed to evaluate the creditworthiness of applicants with respect to a specific, *legitimate business purpose* of the creditor; and directly related to a legitimate business objective or necessity, such as (but not limited to) maximizing profit, limiting the risk of default, or limiting loss exposure in the event of a default.
- *Statistically valid:* Developed and validated based on generally accepted statistical practices and methodologies.
- *Periodically revalidated:* Re-evaluated for statistical soundness from time to time, and adjusted as necessary, using appropriate statistical methods and the creditor's own data, to maintain predictive ability. If a model's predictive power deteriorates significantly over time, the foundation for its statistical validity and business justification may erode.

Generally, EDDSS models are less likely to cause fair lending compliance issues.

---

[6] The availability of disparate impact (or "effects test") liability under the ECOA does not appear to be an entirely settled question. A footnote in Regulation B (12 CFR 202.6, footnote 2) asserts that Congress intended for disparate impact liability to be available. The Department of Justice (DOJ), the CFPB, the Federal Deposit Insurance Corporation (FDIC), and other agencies have all pursued enforcement actions or reached settlements with creditors in cases that included allegations of disparate impact under the ECOA. In addition, the CFPB announced its intention to continue employing a disparate impact "effects test" under the ECOA in its fair lending examinations and enforcement actions (CFPB Bulletin 2012-04, April 18, 2012). However, some have argued that the text of the statute does not permit disparate impact liability claims. See, Peter N. Cubita and Michelle Hartmann, "The ECOA Discrimination Proscription and Disparate Impact - Interpreting the Meaning of the Words That Actually Are There," *The Business Lawyer*, v. 61:2, February 2006.

[7] See OCC Bulletin, 97-24, op. cit. "Elderly" consumers are defined under the ECOA as age 62 or older.

Potential proxies for prohibited factors

Ostensibly neutral variables that predict credit risk may nevertheless present disparate impact risk on a prohibited basis if they are so highly correlated with a legally protected demographic characteristic that they effectively act as a substitute for that characteristic. For example, geographic location and income of the consumer, each of which may have some power to predict credit risk, may also be highly correlated with race or ethnicity, which could create a disparate impact risk. In both examples, alternative predictors of credit risk may be available to capture the predictive power of these variables, but with less fair lending risk.

In the case of income, the debt-to-income ratio might be used as a predictive factor instead of the level of income, and it may even be found to have better predictive power than income itself. In the case of geographic predictive variables, location may simply be a proxy for some underlying economic factor(s), rather than inherently a predictor of credit risk. For example, state-level differences in economic performance or cost of living, as well as differences in state laws, may be associated with differences in the average risk of default. Where this is the case, the risk factor(s) underlying geography (if they can be identified) would likely have greater predictive power for credit risk, and may entail less fair lending risk.

Finally, some credit card issuers have moved beyond using past credit performance in scoring models to also including "non-traditional" predictive variables such as:

- the types of creditors used (e.g., the identity of a consumer's mortgage lender as a prime versus subprime lender, or the consumer's use of payday loans);
- average credit risk of people residing in a particular area;
- area economic factors (e.g., unemployment rates or property appreciation rates); and
- indicators of an applicant's customer relationship status (e.g., whether the applicant is an existing customer of the bank or an affiliate, the number of the bank's products or services the customer uses, or the amount of funds on deposit).

The use of non-traditional attributes in credit scoring models has the potential to create disparate impact risk if the attributes are correlated with a prohibited factor. Therefore, the attributes should be assessed for fair lending risk, and issuers using these sorts of factors in custom scoring models should develop rigorous evidence of their business justification.

Risks arising from mismanagement or misuse of models

The mismanagement or misuse of credit scoring models can be an issue for both fair lending compliance and credit risk management. For example, *ad hoc* and judgmental adjustments to models, such as changes to score thresholds for underwriting approval, changes to weights on predictive variables, or the addition or removal of explanatory variables, increase fair lending risk by undermining the demonstrable statistical validity of a model. This could convert a model that was originally EDDSS into a "judgmental scoring system." If a model uses age as a predictive variable, then the loss of EDDSS status could be a serious regulatory issue. Therefore, management of fair lending compliance risk requires that all changes to a model be based on appropriately documented, rigorous empirical analysis; and that revised models be revalidated.

Fair lending risk can also arise from model risk management issues such as applying a model to a consumer population for which it was not developed, or failing to correctly implement a model . For example, if a model was developed based on a sample of consumers with prime credit, but then applied to a non-prime product or consumer population, it may lose its statistical validity.
As another example, a Consumer Financial Protection Bureau (CFPB) enforcement action in 2012 included settlement of an alleged ECOA violation resulting from a credit card issuer's failure to

correctly implement a credit scoring model. In that case, the credit card issuer allegedly developed an age-segmented scoring model but only implemented it on a staged basis. For a period of eight months, the issuer had implemented the model for applicants age 35 and younger, but not for applicants over the age of 35.[8] The CFPB contended that this violated the ECOA because the law requires credit scoring systems that take age into account to be properly designed and implemented.[9]

Fair lending risk may also arise from a failure to appropriately monitor the use and performance of a model over time. First, model performance may degrade over time due to factors such as changes in the consumer population, consumer behavior or credit policy, among others. To the extent that a model loses its claim to being EDDSS due to such changes, fair lending risk increases because, in the event the model has a disproportionate adverse impact on a prohibited basis, it may not have a sufficiently strong business justification to counter a disparate impact claim. As noted above, a model using age as a predictive factor must continue to meet the EDDSS standard over time, not just when it was originally developed.

Second, insufficient oversight and management of overrides or exceptions can create fair lending risk in two ways: (1) an excessive number or frequency of overrides could undermine a model's claim to statistical validity as a predictor of credit risk, and (2) overrides could result in different treatment of similarly qualified applicants who differ in terms of prohibited characteristics. To the extent that judgmental score overrides are allowed, fair lending risk can be controlled by establishing clear guidelines regarding the allowable reasons for overrides, requiring that underwriters document the reasons for granting an override, and monitoring to ensure that the guidelines are followed and that the volume or frequency of exceptions remains within an acceptable range. This applies to both high side overrides (overrides of system approval decisions) and low side overrides (overrides of system decline decisions).

It is important to realize that the fair lending risk issues discussed above arise from basic problems with model risk management. Therefore, fair lending risk issues can sometimes be avoided or mitigated simply by following sound model risk management practices, including developing and maintaining documentation and evidence of each model's statistical validity, both at development and over time.

Testing a model for disparate impact risk

Federal financial institution regulators have not published guidance on how to test for the disparate impact risk of credit scoring systems, or how large an adverse impact on a protected class needs to be before it becomes a regulatory compliance concern. Absent official guidance, various reasonable approaches might be considered. The appropriate approach to testing for disparate impact may depend on the specific structure of the scoring model. In addition, some degree of judgment may be required to evaluate whether the size of any differential impact is large enough to warrant a compliance concern.

---

[8] An age-split scoring model uses different scorecards or models based on the age of an applicant, with each scorecard containing variables that are predictive for a given age group. See Regulation B, 12 CFR 202.6(b)(2).

[9] Federal Deposit Insurance Corporation and Consumer Financial Protection Bureau, In The Matter of American Express Centurion Bank, Joint Consent Order, Joint Order For Restitution, and Joint Order to Pay Civil Money Penalty, FDIC-12-315b, FDIC-12-316k, 2012-CFPB-0002, October 1, 2012. See also the related CFPB press release. Accessed on February 13, 2014, http://www.consumerfinance.gov/newsroom/cfpb-orders-american-express-to-pay-85-million-refund-to-consumers-harmed-by-illegal-credit-card-practices/.

Two key points should be noted about disparate impact testing of models. First, a disparate impact analysis should clearly distinguish between the impacts of predictive variables in a scorecard taken individually, and the impact of the scoring system as a whole. Scoring models attempt to capture a multivariate relationship between a measure of credit performance and a set of multiple predictive variables that may have varying degrees of correlation with demographic characteristics and with each other. Even if a particular variable has a disproportionate impact on a prohibited basis, the scorecard as a whole may be free of such an impact, because negative effects related to one factor in a model may be offset by positive effects of other factors.

Second, disparate impact risk cannot be evaluated simply by comparing the average credit scores or score distributions of different demographic groups, such as minority and white consumers, or by comparing rejection rates at a given cut-off score. Such comparisons often show that a minority group's score distribution is skewed toward lower scores compared to white consumers, or that the minority group has a higher rejection rate at a given cut-off score. However, such a pattern by itself is not inherently an indication of an illegal disparate impact. It may simply be that minority consumers objectively tend to have higher credit risk than white consumers, as measured by the credit criteria included in the model, due to systematic differences in average income, wealth, employment, credit experience, or other economic factors.

However, if a model tends to assign different scores to applicants of the same underlying risk profile (based on observed and unobserved risk factors), but belonging to different protected demographic groups, then there may be a risk of potentially illegal disparate impact. Various statistical methods can be used to formally test whether a model tends to score one class of applicants differently than another class to a significant degree.[10]

The finding of a differential adverse impact on a prohibited basis in a scoring model is not the end of the story with respect to the disparate impact question. As discussed above, a model or predictive variable with a disproportionate adverse impact may still be legally permissible if it has a demonstrable business justification and there are no alternative variables that are equally predictive and have less of an adverse impact. Therefore, the next steps in a disparate impact analysis should be to (1) identify the predictive variable or variables that are the source of the adverse impact, (2) evaluate whether the effect is large enough to be of statistical and practical significance, (3) scrutinize the statistical evidence from the development of the model regarding the predictive power of the variable(s) in question, and (4) evaluate whether alternative predictive variables are available that have less of an adverse impact but do not significantly degrade the predictive power of the model.

If there is evidence that a variable has a disproportionate adverse impact, investigation of the business justification for the variable should include evaluation of the empirical basis for its inclusion in the model, and its importance to the model's predictive power and business objectives (such as achieving a target charge-off rate or level of profitability). If a variable only marginally contributes to predictive power or business objectives, but produces a disproportionate adverse impact on a prohibited basis, then it is worth considering whether use of the variable presents an acceptable regulatory compliance risk.

Finally, even if the analysis concludes that a variable with some level of adverse impact has a sufficient business justification, and there are no close substitutes available with less of an adverse

[10] See, for example, Robert B. Avery, Kenneth P. Brevoort and Glenn Canner, "Does Credit Scoring Produce a Disparate Impact?" *Real Estate Economics*, 2012, v40 S1: pp. S65–S114; and Elaine Fortowsky and Michael LaCour-Little. "Credit Scoring and Disparate Impact," Working Paper, Wells Fargo Home Mortgage, December 13, 2001.

impact, some professional and legal judgment normally comes into play in assessing whether to use the variable. The fact that a proxy effect exists may increase regulatory risk (including the cost of defending the use of the variable in the face of regulatory scrutiny), and the trade-offs between such risk and the benefits of using the variable should be weighed.

Conclusion

Credit scoring models are important tools for managing credit risk and limiting fair lending compliance risk. However, if not properly constructed, managed, and monitored, scoring models can also be a source of fair lending risk. Scoring models should be monitored and tested by professionals with the requisite technical skills and regulatory compliance experience to ensure they have demonstrable statistical validity, and do not inadvertently create a risk of fair lending regulatory violations.

About the Financial Economics Practice at CRA

CRA's Financial Economics Practice provides economic and financial analysis and advice to financial institutions, financial regulators, and counsel representing financial institutions. Our experts are skilled in quantitative modeling and econometrics, particularly as applied to issues in credit and compliance risk in primary and secondary consumer lending markets. To learn more about the practice, visit **www.crai.com/financialeconomics**.

Contact

David Skanderson, PhD
Principal
Washington, DC
+1-202-662-3955
dskanderson@crai.com

Alex Stricker
Principal
Washington, DC
+1-202-662-7890
astricker@crai.com

CRA Charles River Associates