TransUnion Supports Federal Reserve Board’s Findings on Credit Scoring

The Federal Reserve Board’s division of Research and Statistics’ study concerning the effect credit scoring has on the availability and affordability of credit on various population segments was mandated by Section 215 of the 2003 Fair and Accurate Credit Transactions Act of 2003 (FACT Act).

The study considered public comments submitted specifically for the report and past research on the topic. In addition, the FRB used a representative sample of more than 300,000 de-personalized credit records of the same individuals from two distinct points in time (June 30, 2003 and December 31, 2004). This allowed the FRB to create its own generic scoring model to compare its findings with TransUnion’s proprietary generic risk scoring model and VantageScore™, the industry’s newest scoring model.

The study further concluded that credit scoring “increases the efficiency of consumer credit markets by helping creditors establish prices that are more consistent with the risks and costs inherent in extending credit.” The study also recognized, on average, the inclusion of specific individual credit characteristics used in credit scoring models has no adverse affect on credit scores for the general population.
Report to the Congress on Credit Scoring and Its Effects on the Availability and Affordability of Credit

Submitted to the Congress pursuant to section 215 of the Fair and Accurate Credit Transactions Act of 2003

August 2007
EXECUTIVE SUMMARY

Credit scoring is a statistical technology that quantifies the credit risk posed by a prospective or current borrower. The technique is widely used to evaluate applications for credit, identify prospective borrowers, and manage existing credit accounts. The large savings in cost and time that have accompanied the use of credit scoring are generally believed to have increased access to credit, promoted competition, and improved market efficiency.

The expansion of the use of credit scoring, including by the adaptation of its methodology to insurance markets, has been accompanied by concerns that it may affect the availability and affordability of credit and insurance and that factors included in credit-scoring models may have adverse effects on certain populations, particularly minorities. Section 215 of the Fair and Accurate Credit Transactions Act of 2003 (Fact Act) directs the Federal Reserve Board and the Federal Trade Commission (FTC) to study how credit scoring has affected the availability and affordability of credit and insurance, to determine the relationship between credit scores and actual credit losses and insurance claims, and to determine how these relationships vary for the population groups protected under the Equal Credit Opportunity Act (ECOA). In addition, section 215 directs the Board and the FTC to study the extent to which the consideration of certain factors included in credit-scoring and insurance-scoring models could have a negative or differential effect on populations protected under ECOA and the extent to which alternative factors could be used in credit scoring to achieve comparable results with less negative effect on protected populations.

In preparing the study, the Federal Reserve took the lead in assessing the effects of credit scoring on credit markets, the subject of the present document; the FTC took the lead in the area of insurance and has issued a separate report on that topic.

In the broadest terms, the findings of the Federal Reserve study are as follows:

1. The credit history scores evaluated here are predictive of credit risk for the population as a whole and for all major demographic groups. That is, over any credit-score range, the higher (better) the credit score, the lower the observed incidence of default. These conclusions are limited to credit history scores, that is, scores calculated exclusively on the basis of individuals’ credit records as assembled by the three national credit-reporting agencies (Equifax, Experian, and TransUnion). Other kinds of credit scores were not studied here.

2. Results obtained with the model estimated especially for this study suggest that the credit characteristics included in credit history scoring models do not serve as

---

1 The Fact Act, Public Law 108-159, enacted December 4, 2003; section 215 is reproduced in appendix A of this report.
When developing models, TransUnion takes a proactive approach to ensure that the characteristics used make sense strategically, logically and are defendable against adverse action.

Evidence also shows that recent immigrants have somewhat lower credit scores than would be implied by their performance. This finding appears to derive from the fact that the credit history profiles of recent immigrants resemble those of younger individuals, whose credit performance tends to be poor relative to the rest of the population. Expanding the information supplied to credit-reporting agencies to include rent, other recurring bill payments, nontraditional uses of credit, and the credit histories of the foreign-born in their countries of origin may provide a broader picture of the credit experiences of recent immigrants and other individuals.

(3) Different demographic groups have substantially different credit scores, on average. For example, on average, blacks and Hispanics have lower credit scores than non-Hispanic whites and Asians, and individuals younger than age 30 have lower credit scores than older individuals. Also, for given credit scores, credit outcomes—including measures of loan performance, availability, and affordability—differ for different demographic groups. Data limitations (for example, regarding individuals’ wealth, employment, and education) prevented a complete assessment of these differences in score averages and outcomes among groups. The study found that many of these differences were reduced, at least in part, by accounting for the limited factors available for this study; however, differences—sometimes substantial—often remained.

(4) Evidence provided by commenters, previous research, and the present analysis supports the conclusion that credit has become more available over the past quarter-century. Credit scoring, as a cost- and time-saving technology that became a central element of credit underwriting during that period, likely has contributed to improved credit availability and affordability. However, in part precisely because the use of credit scoring became widespread decades ago, only limited direct information could be obtained on the contribution of credit scoring regarding availability and affordability. The increase in credit availability appears to hold for the population overall as well as for major demographic groups, including different races and ethnicities. There is no compelling evidence, however, that any particular demographic group has experienced markedly greater changes in credit availability or affordability than other groups due to credit scoring.
Executive Summary

Data Used to Prepare the Report
Despite concerns about the potential effects of credit scoring on minorities or other groups, little research has been conducted on the issue, largely because of a lack of data linking credit scores to race, ethnicity, and other pertinent demographic information about individuals. With the exception of dates of birth, the credit records maintained by the credit-reporting agencies, which serve as the basis for most credit-scoring models, do not include any personal demographic information, and federal law generally prohibits the collection of such data on applications for nonmortgage credit. Even in the context of mortgage credit, for which some creditors are required to collect information on race, ethnicity, and sex, little information is publicly available.

This report was prepared using two types of information. The first type was gathered from public comments submitted for the report and from a review of previous research and surveys. The second type came from unique research conducted by the staff of the Federal Reserve Board specifically for this study. In that research, the Board’s staff created a database that, for the first time, combines information on personal demographics collected by the Social Security Administration (SSA) with a large, nationally representative sample of the credit records of individuals. The sample comprised the full credit records of 301,536 anonymous individuals drawn in June 2003 and updated in December 2004 by TransUnion LLC (TransUnion), one of the three national credit-reporting agencies.2

Because the data set consisted of the credit records of the same individuals on two dates (June 30, 2003, and December 31, 2004), the Federal Reserve’s staff was able to construct measures of loan performance, credit availability, and credit affordability and to create its own credit-scoring model (the FRB base model). Besides the FRB score created for this study, the data supplied by TransUnion for each individual in the database included two commercially generated credit scores—the TransRisk Account Management Score (from TransUnion) and the VantageScore (from VantageScore Solutions LLC).3 The design of the FRB base model followed general industry practice to the extent possible. The three credit scores, together with the unique combination of credit and demographic information in the data set created for this purpose, allowed the Federal Reserve to address the questions posed by the Congress.

Access to Credit
The limited available evidence, including from public comments and previous research, suggests that credit scoring has increased the availability and affordability of credit. The information, such as names and Social Security numbers, was not made available to the Federal Reserve.

2 TransRisk Account Management Score is a registered trademark of TransUnion LLC; and VantageScore is a service mark of VantageScore Solutions LLC. All other trademarks, service marks, and brands referred to in this report are likewise the property of their respective owners.
basic reason is that credit scoring allows creditors to quickly and inexpensively evaluate credit risk and to more readily solicit the business of their competitors’ customers regardless of location.

Credit scoring likely increases the consistency and objectivity of credit evaluation and thus may help diminish the possibility that credit decisions will be influenced by personal characteristics or other factors prohibited by law, including race or ethnicity. Credit scoring also increases the efficiency of consumer credit markets by helping creditors establish prices that are more consistent with the risks and costs inherent in extending credit. By providing a low-cost, accurate, and standardized metric of credit risk for a pool of loans, credit scoring has both broadened creditors’ access to capital markets and strengthened public and private scrutiny of lending activities.

TransUnion routinely makes anonymous credit files available to independent third-parties to ensure the validity of other models.

Credit Scores and Loan Performance, Availability, and Affordability across Populations

The data assembled for this study are used to investigate the variation in credit scores across populations and the relationship between credit scores and loan performance, availability, and affordability across populations.

Credit scores differ among subpopulations: Blacks, Hispanics, single individuals, those younger than age 30, and individuals residing in low-income or predominantly minority census tracts have lower credit scores than other subpopulations defined by race or ethnicity, marital status, age, or location. Group differences in credit scores are narrowed, but not always eliminated, when differences in personal demographic characteristics, in residential location, or in a census-tract-based estimate of an individual’s income are taken into account.

The analysis conducted for this study finds that credit scores consistently predict relative loan performance within all population groups; that is, for all populations, the percentage of individuals experiencing a serious delinquency on one or more of their credit accounts consistently declines as credit scores increase.

The analysis also finds that some groups perform worse (experience higher rates of serious delinquency) on their credit accounts, on average, than would be predicted by the performance of individuals in the broader population with similar credit scores. For example, on average, blacks perform worse than other racial and ethnic groups with similar credit scores. Similarly, single individuals and those residing in predominantly black or low-income census tracts perform worse on their loans than do their complementary demographic groups with similar credit scores. In contrast, the loan performance of Asians, married individuals, foreign-born individuals (particularly, recent immigrants), and those residing in higher-income census tracts was better than the performance predicted by their credit scores. The results hold after controlling for the other personal demographics of these individuals and for an estimate of the individuals’
Executive Summary

incomes and locations; other factors that could be important, such as differences in employment experience, were not available.

The study also finds that credit scores are consistently related to measures of loan pricing and loan denial rates inferred from credit inquiries. That is, for all populations, interest rates derived from the terms reported for closed-end loans and average inferred denial rates consistently decline as credit scores increase. As was the case for loan performance, some differences were observed across population groups after controlling for credit score. Most notably, younger individuals appear to experience somewhat higher inferred denial rates than older individuals; blacks appear to pay somewhat higher interest rates on auto and installment loans than do non-Hispanic whites; and Asians pay interest rates that, on average, are typically lower than, or about the same as, those paid by non-Hispanic whites across all loan categories for which rates could be estimated.

Data limitations prevent a full assessment of the reasons for the remaining differences in credit outcomes.

Individual Credit Characteristics and Their Effects across Populations

This study reviewed the extent to which the consideration or lack of consideration of certain factors by credit-scoring systems could result in a negative or positive differential effect for different populations. By law and regulation, an individual’s personal characteristics—such as race or ethnicity, national origin, sex, and, to a limited extent, age—must be excluded from credit-scoring models. A concern exists that, despite that prohibition, a credit characteristic may be included in a model not because it helps predict performance but because it is a substitute, or proxy, for a demographic characteristic that is correlated with performance.

The analysis of the data assembled for this report found that few credit characteristics, including those in the FRB base model, were correlated with personal demographics and that therefore they were unlikely to serve as proxies for demographic characteristics. Credit characteristics related to the age of an individual’s credit record are the primary exception. The data show that some of these characteristics are often highly correlated with age. In addition, certain pertinent aspects of the credit files of recent immigrants tend to resemble those of younger individuals because they have not had sufficient time to build an extensive credit history in the United States.

To examine more closely whether the credit characteristics appearing in the FRB base model are serving, at least in part, as proxies for race or age, the model was reestimated in race-neutral and age-neutral environments. In each case, the FRB base

---

4 Credit inquiries are requests by creditors for an individual’s credit report. The lending industry uses the presence of credit inquiries without the issuance of new credit as an indication of loan denial. The data on credit inquiries are likewise used in this study to infer whether an individual likely experienced a credit denial.
model was reestimated with samples limited to a single race or age population respectively; in those reestimations, any credit characteristics serving solely as a proxy for race or age should have little weight in the reestimated model. Credit characteristics that have both an independent effect on performance and a correlation with race or age would be expected to have significantly different weights (either larger or smaller) in the reestimated models.

Reestimating the FRB base model in a race-neutral environment had little effect on credit scores. The result suggests that none of the credit characteristics included in the model serve, to any substantive degree, as proxies for race or ethnicity. However, when the FRB base model was reestimated in an age-neutral environment, credit scores did change: Scores for recent immigrants and younger individuals rose, and scores for older individuals fell. These results were traced to the inclusion of a specific credit characteristic, namely, that which specifies the length of an individual’s credit history. Further analysis showed that this credit characteristic served in part as a proxy for age. However, because the characteristic also had significant predictive power in an age-neutral environment, the effect could not be mitigated simply by excluding the credit characteristic from the FRB base model. An alternative means of mitigating the differential effect of this characteristic would be to use the weights derived from the age-neutral model. Use of the credit characteristic in this manner removes the differential effects relating to age with less loss of model predictiveness than would occur if this credit characteristic were excluded from the model entirely.
OVERVIEW OF THE REPORT

In recent decades, consumer credit markets in the United States have become increasingly national in scope as lenders have been better able to expand their geographic reach. These trends have been facilitated by the development of statistically derived credit-scoring models to mechanically evaluate credit risk, help establish loan prices, and manage consumer credit accounts. As a cost-saving technology, credit scoring has greatly affected consumer credit markets by allowing creditors to more inexpensively and readily gauge credit risk and expand their reach to consumers beyond the limits of their local offices.

The data maintained by credit-reporting agencies on the credit-related experiences of the majority of adults in the United States are at the heart of most credit-scoring models. Although credit scoring has been a feature of consumer lending markets for some time, its role has expanded in recent years, in part because the data maintained by those agencies have become more comprehensive. Indeed, many credit-scoring models, particularly those used for screening users of unsecured revolving consumer credit, such as credit card customers, are now sometimes based entirely on information contained in the records of the credit-reporting agencies. The scores generated by those models, referred to here as credit history scoring models, have helped to substantially reduce the cost and time needed to make credit decisions and to identify prospects for new credit.

The evaluation of creditworthiness, whether done judgmentally or on the basis of a credit score, is an inherently inexact science in that it attempts to predict the future: whether a loan will be repaid according to the agreed-upon terms. In building a credit-scoring model, the goal is to identify and use only those factors that have a proven relationship to borrower payment performance. By law and regulation, an individual’s personal characteristics—such as race or ethnicity, national origin, sex, and, to a limited extent, age—must be excluded from credit-scoring models. In this way, credit scoring promotes consistency and objectivity in credit evaluation and may help diminish the possibility that such personal characteristics are considered in the lending process.

As the use of credit scoring has expanded, so have concerns about the extent to which it may affect access to credit and about whether scoring may have adverse effects on certain populations, particularly minorities or those that rely more heavily on nontraditional sources of credit. These concerns reflect, among other things, a belief that the effect of including certain credit-record items in the development of credit-scoring

\[1 \text{ Under the Fair Credit Reporting Act, these organizations are referred to as consumer-reporting agencies. Although these agencies are sometimes elsewhere referred to as credit bureaus, that term includes firms that do not collect information on credit accounts, and such firms are not considered in this report.} \]

\[2 \text{ Industry participants often refer to credit history scoring models as credit-bureau-based-scoring models.} \]
models may have a differential effect on certain groups, particularly on racial and ethnic minority groups relative to non-Hispanic whites.

Little research has been conducted on the potential effects of credit scoring on minorities or other groups. Reliable data for conducting such research are not readily available. Creditors are generally prohibited from collecting race, ethnicity, and other personal demographic information on applications for credit, except in the case of mortgage credit. Even in the context of mortgage credit, only limited information is collected. Consequently, with the exception of dates of birth, the credit records maintained by the credit-reporting agencies do not include any personal demographic information.

The Fair and Accurate Credit Transactions Act of 2003 (Fact Act) addressed the need for research in this area. Section 215 of the Fact Act (reproduced in appendix A of the present report) directed the Federal Reserve Board and the Federal Trade Commission (FTC), in consultation with the Office of Fair Housing and Equal Opportunity of the Department of Housing and Urban Development (HUD), to study

1. the effects of the use of credit scoring on the availability and affordability of credit
2. the statistical relationship between credit scores and the quantifiable risks and actual losses experienced by businesses after accounting for personal demographics and other known risk factors
3. the extent to which the use of credit scores and credit-scoring models may affect the availability and affordability of credit to protected populations under the Equal Credit Opportunity Act (ECOA)
4. the extent to which the consideration or lack of consideration of certain factors by credit-scoring systems could result in negative or differential treatment of protected classes under ECOA
5. the extent to which alternative factors could be used in credit scoring to achieve comparable results with less negative effect on protected populations
6. the extent to which credit-scoring systems are used by businesses, the factors considered by such systems, and the effects of variables that are not considered by such systems

Section 215 also directed the study to include an analysis of these same questions for the use of credit scoring in insurance markets. In preparing the study, the Federal Reserve took the lead in assessing the effects of credit scoring on credit markets; the FTC

---

3 Under the Home Mortgage Disclosure Act of 1975, as amended in 1989, covered lenders are required to collect and disclose information about the race or ethnicity and sex of individuals applying for mortgages covered by the law.

4 The Fact Act, Public Law 108-159, was passed by the Congress on December 4, 2003.
Overview of the Report

took the lead in the area of insurance and is preparing a separate report on this subject. The present document focuses on credit scoring and credit markets.

Scope of the Analysis
Section 215 of the Fact Act essentially asks for four related analyses regarding the use of credit scoring in credit markets. The first is an analysis of the effect of credit scoring on the availability and affordability of financial products to consumers in general. The second is an analysis of the empirical relationship between credit scores and actual losses experienced by lenders. The third is an evaluation of the effect of scores on the availability and affordability of credit to specific population groups. The fourth is an evaluation of whether credit scoring in general, and the factors included in credit-scoring models in particular, may result in negative or differential effects on specific subpopulations and, if so, whether such effects could be mitigated by changes in the model development process.

Different approaches were taken to conduct each of these four analyses. The approach used to assess the general effect of credit scoring on the availability and affordability of credit was to rely on evidence from public comments, including those from government agencies, industry representatives, community organizations, and fair lending and fair housing organizations. The analysis also drew on evidence from previous studies on the topic and from indirect evidence obtained from an analysis of the Federal Reserve Board’s Survey of Consumer Finances.

The approach taken to examine the empirical relationship between credit scores and actual losses experienced by lenders and to examine the effect of scores on the availability and affordability of credit to specific population groups relied on a nationally representative sample of individuals drawn from credit-reporting agency files at two points in time. Importantly, we were able to obtain information on race, ethnicity, sex, and other demographics from the Social Security Administration (SSA) that could be matched to the credit-record data. Such demographic data has not been available for previous research on credit scoring. The data set also included two commercially available generic credit history scores. In part because of the important role they play in credit markets and in part because of data issues, the analysis here focuses on generic credit history scores.

The data assembled here were also used to estimate a credit history scoring model emulating the process used by industry model developers. This model was used to investigate whether the factors included in credit-scoring models result in negative or differential effects on specific subpopulations and, if so, whether such effects could be mitigated by changes in the model development process.
Background
Before the introduction of credit scoring, the evaluation of creditworthiness was conducted manually and judgmentally by loan officers relying primarily on experience and subjective assessments of credit risk. Because loan officers differ in their experience and subjective assessments of different credit-risk factors, judgmental underwriting can be inconsistent and difficult to manage. Moreover, manual credit evaluation is time consuming and thus costly.

Both credit scoring and judgmental underwriting tend to be opaque processes. In the case of credit-scoring models, they are proprietary, and firms that develop them typically provide the public with only general information about how they were created and how well they perform. In the case of judgmental underwriting, methods are not likely consistent, even within a firm, because evaluators differ in their experience and judgment about credit risk and because it is difficult to establish clear guidelines that would address the numerous factual differences in the credit profiles of consumers.

After a period of rather slow acceptance, credit scoring had, by the 1970s, become widely used by most national lenders. Subsequently, the use of credit scoring expanded greatly with the development of generic credit history scores by Fair Isaac Corporation (FICO scores) and by Management Decisions Systems (the MDS Bankruptcy Score) in the 1980s. Some time after the introduction of these scores, the three national credit-reporting agencies (Equifax, Experian, and TransUnion) developed their own proprietary generic credit history scores, and recently the three agencies jointly developed a new generic credit history score named the VantageScore. Credit scores derived from each of these models are marketed to lenders, and together they have become an important tool not only for credit evaluation but also for the prescreening and solicitation of new customers.

The Effects of Credit Scores on the Availability and Affordability of Financial Products
Although many of the broad effects of credit scoring are well understood, quantifying the effects of credit scoring on the availability and affordability of credit is difficult. The available evidence comes from three sources: comments received from the public on this study and previous research, original analysis of credit records obtained for this study, and an assessment of consumer survey data. Little specific evidence on these topics was provided in public comments or is available from earlier studies.

The available evidence indicates that the introduction of credit-scoring systems has increased the share of applications that are approved for credit, reduced the costs of underwriting and soliciting new credit, and increased the speed of decisionmaking. It has

---

5 Trademarks, service marks, and brands referred to in this report are the property of their respective owners.
also made it possible for creditors to readily solicit the business of their competitors. Although credit-scoring systems can be expensive to develop, they can be operated at low marginal cost. To the extent that the lower costs and time savings are passed through to consumers, they will lead to lower interest rates and greater access to credit.

Credit scoring also increases the consistency and objectivity of credit evaluation and thus may diminish the possibility that credit decisions will be influenced by personal characteristics or other factors prohibited by law, including race or ethnicity. In addition, quicker decisionmaking also promotes increased competition because, by receiving information on a timelier basis, consumers can more easily shop for credit. Finally, credit scoring is accurate; that is, individuals with lower (worse) credit scores are more likely to default on their loans than individuals with higher (better) scores.

Credit scoring increases the efficiency of consumer credit markets by helping creditors establish prices that are more consistent with the risks and costs inherent in extending credit. Risk-based pricing reduces cross-subsidization among borrowers posing different credit risks and sends a more accurate price signal to each consumer. Reducing cross-subsidization can discourage excessive borrowing by risky customers while helping to ensure that less-risky customers are not discouraged from borrowing as much as their circumstances warrant. Finally, risk-based pricing expands access to credit for previously credit-constrained populations, as creditors are better able to evaluate credit risk and, by pricing it appropriately, offer credit to higher-risk individuals.

By providing a low-cost, accurate, and standardized metric of credit risk for a pool of loans, credit scoring has broadened creditors’ access to capital markets, reduced funding costs, and strengthened public and private scrutiny of lending activities.

To better understand the potential effects of credit scoring on the availability and affordability of credit, data from the Survey of Consumer Finances were used to examine how the use of credit has changed from 1983 (the first year for which the survey results are comparable with those of later years) to 2004 (the most recent survey year). During this time, the first generic credit history models were introduced, so it is an appropriate period in which to assess at least some of the effects of credit scoring. However, such an analysis of credit use can provide only indirect evidence of the possible effects of credit scoring on access to credit. Moreover, other factors, including changes in the economic and demographic circumstances of households, technological innovations, and financial deregulation also have affected access to credit, making it difficult to distinguish the effects of credit scoring.

The survey data show that the share of families with any debt rose for nearly all populations; the steepest growth was in the ownership of bank-type or travel and entertainment cards. These trends are in broad alignment with the conjecture that credit scoring has helped increase the availability of credit since the early 1980s. It is difficult to draw a strong inference regarding changes in differences in credit use by race or
ethnicity, age, and income. On the whole, the data do not provide clear and compelling evidence that the broader adoption of credit scoring disproportionately benefited populations that historically had lower rates of debt ownership; for the most part, differences in credit use across groups appear to have changed only slightly or even to have widened.

Assessment of Credit Scoring, Performance, Availability, and Affordability and Differential Effect

The remainder of the study focuses on the analysis of a data set assembled and analyzed by the Federal Reserve specifically for this study. The data, which do not have personally identifying information, are unique in that they combine information on credit accounts and credit scores with information on loan performance and a wide variety of demographic characteristics of a nationally representative sample of individuals. As noted above, legal restrictions have made it difficult to assemble a nationally representative database containing these three elements. The data are used to address several of the requirements of the section 215 study request.

The analysis and results are summarized as follows. Background information on the definition of differential effect and its specific use in this study is followed by a description of the data and the credit-scoring model developed for this study. The results are then presented in four parts: (1) a description of differences found in credit scores for different populations, (2) the relationship between credit scores and loan performance for different groups, (3) additional findings on the effect of credit scores on the availability and affordability of credit for different populations, and (4) findings on differential effect using a credit-scoring model developed by the Federal Reserve staff specifically for that purpose. The concluding section of this summary (and of the report) discusses limitations and qualifications of the research.

Discrimination and the Law

Under ECOA, it is unlawful for a lender to discriminate against a credit applicant on a prohibited basis in any aspect of a credit transaction. Under both ECOA and the Fair Housing Act (FHA), it is unlawful for a lender to discriminate on a prohibited basis in a transaction related to residential real estate.

Despite the existence of federal anti-discrimination laws, longstanding concerns about discrimination in credit markets persist regarding essentially all aspects of the lending process—marketing, credit evaluation, establishment of loan terms, and loan servicing.

---

6 Among the prohibited bases under ECOA are race, color, religion, sex, national origin, age, and marital status.

7 Race, color, religion, sex, and national origin are prohibited bases under the FHA, as under ECOA. Additional prohibited bases under the FHA are handicap and family status but, unlike under ECOA, not age and marital status.
Overview of the Report

Analyses by the courts and federal regulators of credit discrimination often distinguish between discrimination that involves “disparate treatment” and “disparate impact.” Disparate treatment involves treating similarly situated applicants differently on the basis of one of the prohibited factors (for example, offering less-favorable terms to minority applicants). Disparate impact refers to the outcome of a practice that the lender applies uniformly to all applicants but which has a discriminatory effect on a prohibited basis and does not have a sufficient business justification.

Some observers maintain that reliance on automated credit-evaluation systems such as credit scoring serves to reduce the potential for discrimination in lending because the automated nature of the process reduces the potential for bias to influence lending outcomes. Others contend that the credit-scoring process may have a disparate impact on protected populations because some of the factors used in credit-scoring models may disadvantage minorities or other segments of the population protected by fair lending laws.

The Federal Reserve’s Regulation B, which implements ECOA, considers two broad types of credit evaluation: (1) traditional judgmental credit-evaluation systems, which may rely on the subjective evaluation of loan officers; and (2) credit-scoring systems that are empirically derived and demonstrably and statistically sound. Apart from the limited exception of age, which may be used as a predictive factor provided that those aged 62 or older are not assigned a negative factor or value, no prohibited factor may be used in a credit-scoring model.

Except, again, for age, credit-record data do not include personal or demographic characteristics, so such personal characteristics are unlikely to be an explicit part of a model. Of course, disparate treatment could arise if lenders fail to apply credit scores evenhandedly, ignore them, or exercise “overrides” for some populations or in some circumstances.

Under court and regulatory agency interpretations, the test for disparate impact requires that a practice both have a disproportionate effect on a protected population and lack a sufficient business justification. An empirically derived, demonstrably and statistically sound credit-scoring model is likely to have a sufficient business rationale for the characteristics that constitute the model. Even a model that is empirically derived and demonstrably and statistically sound may, however, embody some avoidable disparate impact on a protected population in one or both of the following ways: (1) An alternative approach or specification might achieve the business goal with less discriminatory effect,

---

8 Some courts and agencies have referred to certain forms of particularly blatant discriminatory treatment on a prohibited basis as “overt discrimination.”
10 Some credit records include the date of birth or age.
and (2) the predictiveness of a variable in the model might stem primarily from the fact that it is serving as a proxy for a protected population.

**Differential Effect Analyzed in this Study**

In the previous section, the phrase *disparate impact* was used to refer to the possible differential adverse effects that credit-scoring models may have on various groups in a legal context. In this section, we define more precisely the meaning of the term *differential effect* as used in the statistical analysis of this study. Although related, the legal definition and the term “differential effect” used here are not the same. The concept of disparate impact embodies specific legal criteria and must be applied on a case-by-case basis after considering all relevant facts and circumstances, including any business justification. The concept of differential effect used here is a statistical concept and does not necessarily correspond to the legal concept.

In the present study, a credit-scoring model, or a credit characteristic used in the model, is said to have a statistical differential effect based on a demographic characteristic—say, age—if the model’s predictiveness or the credit characteristic’s contribution to the model’s predictiveness stems, at least in part, from the fact that the score or the credit characteristic serves as a proxy for age. That is, if the model were estimated in an age-neutral environment, the resulting model would be less predictive of performance, or the credit characteristic’s contribution to the model’s predictiveness would decrease.

At a minimum, two conditions must hold for a demographic group to experience a differential effect from the presence of a credit characteristic in a credit-scoring model. First, the demographic characteristic must be correlated with performance; second, it must also be correlated with the credit characteristic used in the model. This relationship is a purely statistical one and does not imply causality in the relationship between the demographic characteristic and credit performance.

Defined this way, differential effect will generally be a zero-sum outcome. For example, if credit performance improves with age, then the less the credit characteristics in a credit-scoring model serve as a proxy for (“absorb”) age, the higher the scores of younger individuals will be and the lower the scores of older individuals will be. Alternatively, the more a model absorbs the positive effect of age on performance, the higher the scores of older individuals will be. When younger individuals are the focus of attention, however, the use of a credit-scoring model that absorbs a substantial portion of the positive effect of age on performance is described here as having a “differential effect” on younger individuals as compared with a model in which less of the age effect is absorbed.

The congressional requirement for the present study focuses on the differential effects that the estimation and application of credit scores and credit-scoring models may
While these 312 credit characteristics are used in many models, TransUnion has thousands of characteristics available to use in model development.

Overview of the Report

have on individuals with different demographic characteristics, including, but not limited to, the demographic characteristics attributable to protected populations under ECOA. Some of those effects could raise questions about illegal discrimination under ECOA and the Fair Housing Act, but some clearly do not.

Data Used in the Study

Before the beginning of this study, the Federal Reserve had already obtained, for other purposes, a nationally representative sample of the credit records of 301,536 anonymous individuals as of June 30, 2003. The data were obtained from TransUnion LLC (TransUnion). This data set included two commercially available generic credit history scores for each individual in the sample—the TransRisk Account Management Score (TransRisk Score) and the VantageScore. The TransRisk Score was generated by TransUnion’s proprietary model for assessing the credit risk of existing accounts. The VantageScore was developed by VantageScore Solutions LLC as a joint venture by Equifax, Experian, and TransUnion to create a measure of credit risk that scores individuals consistently across all three companies. The credit-record data include 312 credit characteristics that are representative of the credit characteristics used by the industry to develop generic credit history scoring models (appendix B lists the 312 credit characteristics).

The only personal demographic information included in an individual’s credit record is the individual’s date of birth (and many records do not even show that item). Thus, other demographic characteristics of the individuals in the credit records had to be obtained elsewhere. It was determined that the most accurate and comprehensive information on race or ethnicity, age, sex, and national origin could be obtained from records maintained by the SSA. Except for race and ethnicity, which are provided on a voluntary basis, all of that information must be provided by individuals who apply for Social Security cards. The SSA supplied this information for the individuals in the credit-record sample because the Federal Reserve Board is a federal agency and because conditions necessary to ensure the anonymity of the individuals were maintained.

Additional data were obtained for the individuals in the sample from a match between the census-block or census-tract place of residence derived from the credit records and Census 2000 data at the census-block and census-tract level of geography. Finally, demographic information, most importantly marital status, was obtained from one of the leading demographic information companies for the individuals in the sample, again through a process that ensured individuals’ anonymity.

To address the congressional directive, it was also necessary to construct measures of credit performance, availability, and affordability. A standard method used

---

11 No specific addresses of individuals included in the sample of credit records were included in the data made available to the Federal Reserve.
by the industry to measure credit performance for model-building purposes is to draw credit records for individuals on two separate dates. The time between those dates is called the “performance period.” Information from the credit records drawn for the later date shows which accounts became seriously delinquent or otherwise exhibited bad performance during the performance period. Information from the earlier date is used to predict subsequent loan performance.

This methodology was adopted in measuring performance for this study. The Federal Reserve’s existing sample of credit records, drawn as of June 30, 2003, was updated as of December 31, 2004, to provide an 18-month performance period, a length of time within the range used by the industry in measuring performance.

This summary highlights two of the five performance measures used in the study: “any-account performance,” which reflects whether any of an individual’s new or existing accounts suffered some form of major shortfall in performance (major derogatory), such as becoming 90 days or more past due, over the performance period; and “modified new-account performance,” which is limited to accounts opened sometime during the first six months of the performance period (that is, July through December 2003). Because the latter measure excludes loans in existence at the beginning of the performance period, it ensures that the borrower performance being evaluated is not already incorporated in the borrower’s initial score.

Measures of the availability and affordability of credit were also developed by following typical industry practice. Information from the second draw of credit records was used to determine which individuals opened new credit accounts during the beginning of the performance period; for closed-end loans, information on loan terms was used to estimate the interest rate on these loans. Credit records do not include a direct measure of loan denials (a measure of credit availability); however, a proxy often used by the industry is to infer that individuals who have credit inquiries but who did not take out new credit during that period were denied. The presence of such inquiries during the beginning of the performance period was used to infer loan denials.

Thus assembled, the data set was still not sufficient to address the extent to which credit-scoring systems incorporate factors that result in differential effect for certain population groups. To address this aspect of the study, it was necessary to develop our own credit history scoring model, which we term the “FRB base model”; fortunately, the data that had been assembled were sufficient for us to undertake the development of the model. The FRB base model reflects closely the methodologies used by the credit-scoring industry in constructing generic credit history scoring models; however, it does not represent fully any particular model in use today. The estimated model was used to test for the potential for differential effects of credit scoring across groups in the context of model development.
Some information that may be relevant to understanding credit performance, availability, and affordability is not included in the data assembled for this study. Most notably, the data do not include the financial and nonfinancial circumstances of individuals, such as their wealth, income, employment experiences, or financial literacy.

**Estimating the FRB Base Model**

The FRB base model was developed using the large, nationally representative sample of the credit records of individuals described above. The data are of the same type used by the industry to build credit history scoring models. To be as transparent as possible, the FRB base model departed somewhat from industry models in that the process of developing it was based entirely on rules. The rules selected, however, mimic general industry practice to the extent possible.

The model was developed with standard statistical techniques and was constructed using the 312 credit characteristics included in the data provided by TransUnion. The model was designed to predict whether an individual would have at least one new or existing account that would become seriously delinquent during the 18-month performance period used in this study (the “any-account performance” measure). The credit-scoring industry typically segments the population into distinct subgroups and estimates separate credit-score models, or scorecards, for each group. In keeping with that industry practice, the FRB base model segments the population into three scorecards according to the number of credit accounts and past credit experience in each individual’s record, or file: The “thin file” scorecard is for individuals with relatively few credit accounts; the “clean file” scorecard is, broadly speaking, for individuals whose credit records show no major derogatories; and the “major-derogatory file” scorecard is for individuals with at least one major derogatory account, collection account, or public record. These three scorecards consist of the 19 credit characteristics (of the 312 available for this study) found to best predict loan performance. The ability of the FRB base model to predict loan performance appears to be on a par with that of other generic credit-scoring models.

**The Relationship of Credit Scores to Credit Performance, Availability, and Affordability for Different Populations**

This section presents an assessment of the relationship of credit scores to credit performance and to credit availability and affordability for different populations. The

---

12 The credit-record data excluded any personal identifying information.
13 A “credit characteristic” is a summary measure of an aspect of an individual’s credit record, such as the number of credit accounts or the months since the most recent delinquency.
14 A major-derogatory account, as used in this study, is any account delinquent 90 days or more or that was involved in a repossession or charge-off.
15 The 19 credit characteristics used in the FRB base model are listed in appendix C.
assessment focuses on (1) the distribution of credit scores across different populations; (2) the extent to which other demographic, credit, and economic characteristics explain differences in credit scores across populations; (3) the stability of the credit scores of individuals over time; (4) the relationship between credit scores and loan performance measured in a variety of ways; (5) the extent to which, given score, performance varies across populations; (6) the extent to which differences in credit availability and affordability across populations can be explained by credit score; and (7) whether differences in performance, credit availability, and pricing may be explained by factors not included in credit records.

Differences in Credit Scores
The data assembled provide information on the distribution of credit scores for different populations. Results are presented in this summary only for the TransRisk Score, though results for the VantageScore and the Federal Reserve’s own estimated score (FRB base score) were also calculated and are virtually identical.

To compare the credit scores derived from different credit-scoring models, it was decided to normalize the scores to a rank-order scale ranging from 1 to 100. Each score was normalized so that each individual’s score was defined by its rank order in the population; a score of 50 places that individual at the median of the distribution.

For the analysis here, nine different groupings of the sample population are considered: The nine population groups are determined by individuals’ race or ethnicity (measured two ways); sex; marital status; national origin (foreign-born or not); age; and the relative income, degree of urbanization, and percent minority population of the census block or tract where the individual resides.¹⁶

Univariate differences in credit scores. Credit scores differ widely across populations, with blacks, Hispanics, individuals younger than age 30, unmarried individuals, and individuals residing in low-income or predominantly minority census tracts having lower credit scores than other subpopulations within their broader demographic group. Males and females have very similar credit-score distributions, and foreign-born individuals

¹⁶ Information on the race or ethnicity of individuals is generally not available in the data used to develop credit scores. However, if place of residence is known, the racial or ethnic composition of the census tract (or census block) can be used as an approximation of an individual’s race or ethnicity. This approach has been used in previous studies that examine the relationship between credit scores and race or ethnicity.

For this report, in addition to the SSA classification of the individual’s race or ethnicity, the adult racial or ethnic composition of the individual’s census block (available for about 85 percent of the population) or census tract is used as an approximation of the individual’s race or ethnicity. The proportion of the block belonging to each racial or ethnic group can be viewed as the probability that a random adult drawn from the block will have that race or ethnicity. The probability is used as a weight in forming the estimates presented in this study.
appear to have a score distribution that is virtually the same as that of the general population.

Differences in credit scores among racial or ethnic groups and age cohorts are particularly notable because they are larger than for other populations. For example, the mean normalized TransRisk Score for Asians is 54.8; for non-Hispanic whites, 54.0; for Hispanics, 38.2; and for blacks, 25.6 (figure O-1). Credit scores by age increase consistently from young to old: The mean TransRisk Score for individuals younger than age 30 was 34.3; for those aged 62 or older, it was 68.1.

Cumulative distributions show that the population differences suggested by the credit-score means generally hold for the entire score distribution for each population. For each level of credit score, the cumulative distribution indicates the proportion of a population with that score or lower. For example, the cumulative distributions of scores for blacks and Hispanics are consistently higher than those for non-Hispanic whites and Asians (figure O-2). Cumulative distributions by age are also consistently ordered, with younger individuals having a higher distribution than that of individuals aged 62 or older. Cumulative distributions for census-tract groupings by racial or ethnic population composition are also consistent with the patterns implied by the race or ethnicity of individuals.

Multivariate analysis of score differences. The univariate relationships described in the preceding section may in part reflect differences across demographic groups in other characteristics. To better understand the source of the differences in credit scores across different populations, a series of multivariate analyses were conducted to identify the independent effects of race or ethnicity, age, and sex on credit-score differences across populations. For race or ethnicity, a regression model was fit using only the non-Hispanic white individuals in the sample, controlling for their age, sex, marital status, and a census-tract-based estimate of individual income and other census-tract characteristics. Predicted values from this equation were then used to predict the scores for blacks, Hispanics, and Asians. Differences between individuals’ actual credit scores and their predicted scores can be interpreted as “unexplained” racial or ethnic effects. Results of this statistical analysis show that the gross difference in the TransRisk Score between non-Hispanic whites and blacks falls by more than one-half; the gross difference between non-Hispanic whites and Hispanics falls by about three-fourths. For age, regressions from a similar analysis suggest that only a minor portion of the relatively wide differences across age cohorts can be explained by the other factors available in the data.

17 The term “unexplained” as used here is a statistical concept. The unexplained difference is defined as the difference in average scores in the scorable sample after other factors included in the multivariate regressions are accounted for. Thus, the size of the unexplained component depends on what other factors are included in the model. Adding factors to the model, or subtracting them, will affect the size of the unexplained differences.
Credit Scores and Loan Performance
Section 215 of the Fact Act requires an analysis of “the statistical relationship, utilizing a multivariate analysis that controls for prohibited factors under the Equal Credit Opportunity Act and other known risk factors, between credit scores . . . and the quantifiable risks and actual losses experienced by businesses.” Information on actual losses experienced by creditors was not available for the study, so the focus was on loans that became seriously delinquent or were in default as represented by the performance measures. Such loans nearly always result in some loss to the creditor.

In response to the congressional requirement, the analysis addresses the question of whether loan performance differs across population groups controlling for credit score. For the analysis here, individuals were grouped by their score as of June 30, 2003, and average performance on loans over the ensuing 18-month period was measured using the performance measures described above.

Univariate patterns of loan performance. Using each of the three credit scores available for this study and a number of different measures of loan performance, the analysis finds that, on average, credit scores are predictive of future loan performance for all groups and differentiate risk well within each population group (figure O-3).18 The general shapes of the performance curves (curves that show the relationship between credit scores and loan performance) are similar across groups. Specifically, loan performance improves with credit score so that the curve declines as scores increase. Within a demographic population, the performance curves are not identical. Of particular interest for this study are performance curves for populations that are uniformly above or below those for others. A performance curve that is uniformly above (below) means that the group consistently underperforms (overperforms), that is, on average performs worse (better) on its loans than would be predicted by the performance of individuals in the overall population with similar credit scores.

Blacks, single individuals, and individuals residing in lower-income or predominantly minority census tracts show higher incidences of bad performance than would be predicted by their credit scores. Similarly, Asians, married individuals, the foreign-born (particularly, recent immigrants), and those residing in higher-income census tracts perform better than predicted. Results for age were mixed: Younger individuals exhibited a higher incidence of bad performance than would be predicted for two of the three credit scores used in this study; for the third credit score, performance on some measures was better than predicted.

18 The figures in this overview present results only for the TransRisk Score. Figure O-4 is further restricted to the modified new-account performance measure.
Overview of the Report

Multivariate analysis of differences in loan performance. In interpreting the patterns of differential effect discussed above, it is important to recognize that the assessments of overperformance and underperformance are based on univariate statistics. It is possible that the performance assessments for one population at least partly reflect effects coming from other factors. To address this possibility, multivariate analyses were conducted. First, an analysis was conducted in a manner similar to that performed for score levels that sought to determine whether performance differences across groups were related to other personal demographics and census-tract-related characteristics. Results show that controlling for other personal demographic and census-tract characteristics has only a modest effect on the assessment of overperformance or underperformance for populations.

Another possible explanation for performance differences may be that different populations take out different types of credit, borrow from different types of lenders, and receive different loan terms even when they have similar credit scores. Consequently a second analysis was conducted that added to the multivariate performance regressions information on loan terms (including amounts borrowed and derived interest rates), date of the loan, type of lender, and type of loan. The analysis was restricted to performance on modified new accounts.

Results show that there are some differences in the types of loans taken out by different groups. Nevertheless, differences in loan terms and interest rates explain virtually none of the differences in overperformance and underperformance by race, sex, or age. This is true when loan terms and interest rates are considered without other controls or along with other demographic and location factors. Thus, despite differences in the kinds of loans used by different populations, this factor does not appear to be the source of differences in performance once credit score is taken into account.

Credit Scores and the Availability and Affordability of Credit across Populations

The study asks for an assessment of the extent to which, if any, the use of credit-scoring models and credit scores affect the availability and affordability of credit by geography, income, race, color, national origin, age, sex, or marital status. The credit-record data assembled for this study are used to provide evidence on the effects of credit scores on the availability and affordability of credit across populations. The analysis here considers several indicators of credit availability by credit score across populations, including differences in credit use patterns, in "inferred" denial rates for credit, and in estimated interest rates.

Credit-record data can provide only limited insights into the effects of credit scores on credit availability and affordability, particularly as it has changed over time. A limitation of the credit-record data is that, although they contain loans extended before June 2003, the data do not contain the credit scores used to underwrite those loans.
However, the credit scores as of June 2003 arguably are likely representative of the scores used to underwrite new loans acquired at the beginning of the performance period used for this study (July–December 2003). Thus, the analysis presented here focuses on the extent to which population differences in the incidence of new credit, inferred denial rates, and the interest rates derived from terms reported for closed-end credit (installment, auto, and mortgage), as described earlier, can be explained by the June 2003 credit score.

Of course, the incidence and pricing of new credit, as well as the decision to accept or deny a loan, are affected by many factors beyond credit score, including both demand and supply elements such as wealth, employment experience, the presence of collateral for the loan, and the loan-to-value ratio for mortgages or other loans.

Results indicate that individuals with credit scores in the lowest credit-score quintile are substantially less likely to have taken out a new loan over the first six months of the performance period than individuals with higher credit scores. The strong relationship between credit scores and the incidence of new credit holds across all populations.

Individuals with lower credit scores experience higher inferred denial rates. This relationship is found across all population groups; after controlling for credit score, however, blacks and Hispanics, younger individuals, and individuals that live in low-income areas show somewhat higher inferred denial rates than other groups (figure O-4). Credit scores and interest rates are inversely related, a relationship that holds for all populations. However, black borrowers experienced higher interest rates than non-Hispanic whites in each loan category for which interest rates can be determined (figures O-5 and O-6 show the rates on new mortgages and auto loans). Interest rates paid by Asians are, on average, typically lower than, or about the same as, those paid by non-Hispanic whites across all credit-score quintiles and each product category for which rates could be estimated.

Multivariate analyses were also conducted for inferred denials and estimated interest rates. Controlling for credit score, loan type, lender and amount borrowed, and location factors reduces differences in interest rates by race and ethnicity, although not completely. The multivariate analysis had less effect in accounting for differences in inferred denial rates.

**Accounting for Economic and Financial Factors Not Available for this Study**

The multivariate analyses in the previous sections were, perforce, restricted to information contained in the credit records, the SSA file match, and factors based upon an individual’s location. Thus, the data assembled for this study can provide only limited insights into the relationship of credit scores to credit performance, availability, and affordability (and essentially no insight into whether the relationship is one of cause and effect). The data do not contain key variables that would need to be taken into account.
Overview of the Report

Missing data include other underwriting factors, such as loan-to-value ratios in the case of mortgages, and the weight given to credit scores relative to these other factors. Missing data also include underlying differences in socioeconomic factors such as wealth and employment experience; only a rough estimate of individual income is available. Moreover, the credit-record data used here are for a brief period in time and therefore cannot reflect changes over time in the relationship between credit scores and the availability or affordability of credit.

The multivariate analysis found unexplained differences in performance residuals among racial and ethnic groups and among age groups. Unexplained differences in accessibility and affordability in the multivariate regressions were also found among racial and ethnic groups. In this section, we use information from the Federal Reserve Board’s 2004 Survey of Consumer Finances (SCF) to explore the possibility that differences in, for example, wealth, employment history, and financial experience might explain some, or perhaps all, of the remaining differences in performance, availability, and affordability across groups. Inferences from this analysis are only suggestive because the information cannot be linked either to the individuals in the study sample or to their credit-related performance or loan terms.

Assessment of the SCF data shows that younger families differ substantially from older families over a wide variety of financial dimensions. Variations across age groups in income, wealth, and their components and in debt-payment burdens and savings largely reflect the life-cycle pattern of income; that is, income rises as workers progress through their careers and falls sharply upon retirement. Also, younger individuals are more likely to experience recent bouts of unemployment. None of these factors were explicitly accounted for in the multivariate performance analysis conducted with the credit-record data.

The SCF data show that income, wealth, and holdings of financial assets are substantially lower for black and Hispanic families than for non-Hispanic white families. Debt-payment burdens and propensities for unemployment are also higher for blacks and Hispanics. These racial patterns generally hold even after accounting for age, income, and family type.

Differences in educational attainment and credit-market experience may relate to financial literacy. For example, high-school and college graduation rates among Hispanics are below those for blacks, which, in turn, are lower than those for non-Hispanic whites. Each of these factors, none of which were included in the credit-record analysis, may at least partially explain differences in performance across racial or ethnic groups.
The Relationship between Individual Credit Characteristics, Credit Scores, and Differential Effect across Populations

Another provision of section 215 of the Fact Act requires an assessment of “the extent to which the consideration or lack of consideration of certain factors by credit-scoring systems could result in negative or differential treatment of protected classes under the Equal Credit Opportunity Act.” This study uses a variety of approaches to address concerns about whether credit-scoring models, or the individual characteristics that constitute the models, embody differential effect.

The Fact Act requires an analysis of the potential for differential effects arising from the use of individual credit characteristics in a credit-scoring model. As noted earlier, it was determined that the best way to address this issue was to develop our own credit-scoring model, mimicking the process used by the credit-scoring industry. Only in this way would we be able to identify the specific credit characteristics included in a model that may have a differential effect by evaluating the consequences on different groups of adding, removing, or otherwise altering the way the characteristics are used. As discussed above, these steps are necessary to address the differential effect of a specific credit characteristic.

The estimated model is used to provide information of three types. The first type of information involves successively dropping each credit characteristic contained in the estimated model and evaluating the change in normalized credit scores for different populations and the overall model predictiveness when these changes are made. If large changes in credit scores for a population occur when a credit characteristic is dropped, there is an inference that the characteristic embodies differential effect. The second complementary type is to successively add credit characteristics that were not included in the estimated model and then evaluate how such additions would affect scores. Again, significant score changes would suggest differential effect.

These two types of information provide only inferential indicators about differential effect. As noted earlier, to fully assess differential effect, it is necessary to compare credit scores and weights assigned to credit characteristics derived from the FRB base model with those obtained from models estimated in demographically neutral environments. Thus, additional credit-scoring models were estimated in race- and age-neutral environments using several different methods to define neutrality. The third type of information focuses on the comparison of scores and weights from these models with those from the FRB base model and forms the basis of the assessment of differential effect.

The Effects of Dropping and Adding Characteristics

One way of drawing an inference about differential effect for a credit characteristic included in a model is to examine the effect on the credit scores of each demographic
Overview of the Report

TransUnion’s newly redeveloped models, TransRisk Account Management and TransRisk New Account, do not use this characteristic.

group of dropping the credit characteristic. This analysis, which was conducted separately for each scorecard of the FRB base model, proceeded by dropping each individual credit characteristic from the FRB base model, reestimating the model, renormalizing the scores, and comparing the scores with those produced by the FRB base model.

Results of this analysis indicate that for most populations dropping any single credit characteristic (even those found to be highly predictive of loan performance) has a very minimal effect on mean credit scores, typically 1 point or less. Thus, such changes have virtually no impact on mean score differences between population groups. The small change in mean scores when a single credit characteristic is dropped reflects the high degree of correlation among the credit characteristics in the scoring model.

The one exception to this pattern is the credit characteristic “average age of accounts on credit report” on the clean-file scorecard.19 Dropping this credit characteristic from the clean-file scorecard increases mean credit scores for individuals younger than age 30 (5.4 points) and recent immigrants (6.7 points). The net effect is to reduce the mean score differences on the clean-file scorecard between individuals younger than age 30 and those aged 62 or older by about one-fourth. The lower mean scores for the young and recent-immigrant populations when this credit characteristic is included on the clean-file scorecard suggest that including this credit characteristic in a model may have a differential effect on these two populations.

Dropping credit characteristics does not provide any information about credit characteristics not included in the model. An inference about differential effects from excluded credit characteristics can be derived by examining the effect on the credit scores of each demographic group of adding individual excluded credit characteristics. This analysis, which was conducted separately for each scorecard of the FRB base model, proceeded by adding an additional credit characteristic to the FRB base model, reestimating the model, renormalizing the scores, and comparing the new scores with those produced by the FRB base model.

Across population groups, credit scores change very little after the addition of a new credit characteristic. Changes in mean scores for all population groups are approximately 1 point or less regardless of the credit characteristic added. Among the credit characteristics examined were those related to finance company accounts. These characteristics deserve particular note because concerns have been raised about their inclusion in credit-scoring models. However, when added to the FRB base model, credit characteristics related to finance company accounts had essentially no effect on the mean credit scores of any racial, ethnic, or other demographic group. (Note that dropping the

19 Calculated as the average age of all credit accounts in an individual’s credit record.
one credit characteristic related to finance company accounts included in the FRB base model had little effect on mean score differences across populations.)

**Race- and Age-Neutral Models**
The analysis above points to only two broad demographic categories, race and age, that are potentially proxied for by credit characteristics in our model and, thus, may have the potential for a differential effect. (Tests were also run for sex, with no differential effects observed.) Consequently, we focus on these two population taxonomies in estimating a model in a “group neutral” environment. Two methods are used to define neutrality for each population taxonomy. The first method is to restrict the sample used in model estimation to a single race (the “white only” model that uses only non-Hispanic whites for estimation) or to an age range (the “older-age” model that uses only individuals aged 40 or older for model estimation). The second method uses the entire sample in estimation but includes racial-intercept or age-intercept shifts (referred to, respectively, as the “racial-indicator variable” and “age-indicator variable” models). We test for differential effect by freezing the credit characteristics and attributes of the FRB base model and reestimating the attribute weights in the four demographically neutral environments described above.

Reestimating the attribute weights in demographically neutral environments is not a complete test of the potential for differential effect. It is possible that the presence of a large differential effect could mute the importance of a credit characteristic, and consequently that credit characteristic might not be included in a model estimated in a demographically neutral environment. To test for this possibility, each of the credit characteristics not included in the FRB base model was added one at a time to the race- and age-neutral versions of the model, and their effects on scores for different populations were evaluated.

**Race-neutral models.** A comparison of the white-only and the racial-indicator-variable models with the FRB base model shows little difference in fit regardless of how model predictiveness is defined. There are also virtually no differences between the group mean and median credit scores for different populations. The overall assessment of differential effect can also be looked at by examining changes in the underperformance or overperformance (conditioned on credit score) for different population groups. For all

---

20 The choice of the population group (in this case, non-Hispanic whites) was driven by considerations of sample size alone. In principle, any group could serve as the base population for estimating a model. The non-Hispanic white population was the only population in the sample of sufficient size to provide a basis for model estimation. In general, the selection of the base group may affect conclusions reached regarding differential effect of various credit characteristics.
performance measures, the underperformance or overperformance for different demographic groups is virtually unchanged for the two racially neutral models.

The only evidence of differential effect for any racial or ethnic group is a slight negative differential effect for recent immigrants. That is, the credit scores of recent immigrants are somewhat lower for the FRB base model than would have been the case had the model been estimated in a racially neutral environment. However, the overall foreign-born population showed no evidence of such an effect. Further, as described below, recent immigrants show a differential effect going in the opposite direction when evaluated in an age-neutral environment.

Tests of adding credit characteristics to the white-only and the racial-indicator-variable models showed no evidence of important excluded credit characteristics. Results were similar to those described above regarding the addition of characteristics to the FRB base model.

Age-neutral models. As with estimations in a race-neutral environment, shifting from the FRB base model to an age-neutral model appears to lead to little decline in predictive power. However, unlike estimations in the racially neutral environment, mean credit scores and mean performance residuals change for certain age groups in the older-age model and the age-indicator-variable model. Overall, for individuals younger than age 30, the credit scores derived from these two models are somewhat lower than the scores derived from the FRB base model. Recent immigrants show a similar pattern. However, scores for individuals aged 62 and older are higher when estimated in an age-neutral environment. Changes in underperformance and overperformance are consistent with these score changes. Results from adding credit characteristics showed no evidence that important credit characteristics were left out of the FRB base model.

Overall, these results suggest that the FRB base model embeds a modest negative differential effect for individuals aged 62 and older and an even more modest (and opposite) differential effect for individuals younger than age 30 and recent immigrants. These effects derive primarily from the weights assigned to credit characteristics related to the length of an individual’s credit history. These characteristics have somewhat more muted effects in the FRB base model than would be the case had the model been estimated in an age-neutral environment.

Recent immigrants appear to have somewhat lower scores in the FRB base model than would be appropriate given their performance. However, this overperformance is not due to a negative differential effect (indeed, as just stated, recent immigrants experience a positive differential effect). Rather, it is attributable to the tendency of recent immigrants to have credit profiles similar to those of young people in terms of the lengths of their credit histories, as reflected in their U.S. credit records.
The scores of recent immigrants might be made more consistent with performance by changes in the credit-reporting process. For example, it might be possible to gather information on the credit histories of recent immigrants from their home countries to supplement the credit records maintained by the three national credit-reporting agencies in the United States. More generally, ongoing industry efforts to incorporate into credit records items traditionally not collected (such as utility and rental payments) and experiences with nontraditional sources of financing (such as payday lenders and pawn shops) would broaden the information included in credit records and might serve to lengthen the period over which individuals would be recorded as having a credit record.

**Limitations of the Analysis**

Section 215 of the Fact Act asks for four related analyses regarding the use of credit scoring in credit markets. The first is an analysis of the effect of credit scoring on the availability and affordability of financial products to consumers in general. The second is an analysis of the empirical relationship between credit scores and actual losses experienced by lenders. The third is an evaluation of the effect of scores on the availability and affordability of credit to specific population groups. The fourth is an evaluation of whether credit scoring in general, and the factors included in credit-scoring models in particular, may result in negative or differential effects on specific subpopulations and, if so, whether such effects could be mitigated by changes in the model development process.

Different approaches were taken to conduct each of these four analyses. The approach used to assess the general effect of credit scoring on the availability and affordability of credit was to rely on evidence from public comments and previous studies on the topic and to obtain indirect evidence from the Survey of Consumer Finances. The ideal way of addressing this question would have been to conduct a “before and after” study of the effects of the introduction of credit scoring on the availability and affordability of credit. Such an endeavor was not possible because credit scoring has been in use for many years, and the distinction between the effects of scoring and those of economic and other changes that took place over the same period is difficult to discern. Also, the available public research is quite limited, perhaps because most analytical studies were proprietary and are not part of the public record. The approach taken here cannot conclusively address these concerns. Thus, our conclusions in this area can only be suggestive.

The approach taken to examine the empirical relationship between credit scores and actual losses experienced by lenders and to examine the effect of scores on the availability and affordability of credit to specific population groups relied on a nationally representative sample of individuals drawn from credit-reporting agency files. There are several limitations to this approach. First, the analysis was limited to credit history...
scores. Second, the data included only two commercially available credit scores. Third, the definition of performance was dictated by the time periods for which the samples were drawn. The resulting 18-month performance period is on the short end of the time frames considered by many in the industry. Further, the time period used to evaluate performance represented a relatively favorable period of macroeconomic performance. Consequently, the absolute levels of performance observed here may overstate the performance one would expect in a less favorable economic climate.

The issues of loan performance and the availability and affordability of credit to different populations were addressed using multivariate analyses, which were restricted to information contained in the credit records supplemented by demographic information from the SSA and data based on location. However, population groups differ widely along many financial and nonfinancial dimensions not reflected in credit records that may affect credit performance and the conclusions one might draw about differences across populations. So, for example, the overperformance or underperformance of a demographic group may be attributable to financial or nonfinancial characteristics (such as employment experience or wealth) that bear on performance and that are correlated with the demographic characteristic but that are not included in the credit records.

Another issue in this section of the analysis is the fact that performance and loan terms could be ascertained only for individuals receiving credit. It is reasonable to expect that individuals denied credit would have experienced both worse performance and higher interest rates; however, these outcomes are not included in the data. To the extent that individuals experiencing denials disproportionately have low credit scores, inclusion of these outcomes would likely have made the performance or interest rate curves steeper.

The assessment of denial rates using the inquiry proxy is subject to the same limitation. Individuals who know that they have a low credit score, or believe that they do, may act under the assumption that they will be denied credit if they apply for it. If so, they are being “discouraged” from applying for credit, and the observed relationship between credit score and denial rate would then be less steep than it would be if everyone wanting credit applied for it. A final issue in this section is the fact that information on demographic characteristics had to be imputed for a portion of the sample. Tests suggest that the results here are generally robust. However, for some population segments, such as marital status, concerns may still remain.

The fourth analysis was conducted using a credit history scoring model developed by Federal Reserve staff. We attempted to emulate the process used by the credit industry’s model developers in estimating credit-scoring models. However, the industry adheres to no single methodology, so our approach was inevitably approximate. For example, data restrictions forced a number of limitations to our approach. Moreover, the fact that industry modelers may have made different decisions or relied upon different
samples clearly limits the generalizations that can be made from our results. These limitations would arise under any circumstances involving the construction of a new model.

Additional concerns are raised about our model development because of the relatively small sample used for estimation. The small sample size prevented evaluation of the FRB base model on an out-of-sample basis (that is, on a sample of individuals different from that used to develop it). Also because of the small sample, the FRB base model was developed with fewer scorecards than are typically used in the industry’s credit history scoring models; consequently, the model has fewer credit characteristics than is typical in the industry. Having relatively few scorecards makes it difficult to identify credit characteristics that might have a differential effect on populations that could constitute other possible scorecards.

A limitation that runs through all four of the analyses is the decision to focus on credit history scoring models, as opposed to the broader class of scoring models. Much of the underwriting and pricing of credit relies upon credit-scoring models that incorporate factors not included in the records of credit-reporting agencies. Further, the underwriting process may use other information that is judgmentally combined with credit scores in making final decisions on underwriting and pricing. The role of some of these other factors could mitigate or alter some of the conclusions reached in this study.
Overview of the Report

Figure O–1. Mean TransRisk Score, by Demographic Group

Note. For definition of characteristics, refer to notes to table 9.
Figure O–2. TransRisk Score: Cumulative Percentage by Demographic Group

Note: For definition of characteristics, refer to notes to table 9.
Overview of the Report

Figure O–3. TransRisk Score: Modified New–Account Performance (Percent Bad), by Demographic Group

Note. For definition of characteristics, refer to notes to table 9.
Figure O-4. TransRisk Score: Inquiry-Based Proxy for Denials, by Demographic Group

Note. For definition of characteristics, refer to notes to table 9.
Figure O–5. TransRisk Score: Mortgage Interest Rate, by Demographic Group

Note. For definition of characteristics, refer to notes to table 9.
Figure O–6. TransRisk Score: Auto Loan Interest Rate, by Demographic Group

Note. For definition of characteristics, refer to notes to Table 9.