2000 V28 3: pp. 523–547
REAL ESTATE
ECONOMICS

Credit Scoring: Statistical Issues and Evidence from Credit-Bureau Files

Robert B. Avery,* Raphael W. Bostic,** Paul S. Calem**
and Glenn B. Canner**

Although credit scoring offers benefits to lenders and borrowers, its use raises important statistical issues that may affect the ability of scoring systems to accurately quantify an individual's credit risk. The evidence from a national sample of credit-bureau records suggests that concerns about omitted-variable bias may be justified, as local economic factors show significant correlations with credit scores.

Credit scores are statistically derived measures of the credit risk associated with an application for consumer, mortgage or small-business credit and are widely used as predictors of future loan performance. Credit scoring makes underwriting less subjective, lowers the costs of originating loans, and increases the speed and consistency of underwriting decisions. Evidence indicates that credit scores are predictive of future loan performance and suggests that scoring increases the accuracy of risk assessment. These features clearly benefit lenders, but they can also serve the interests of borrowers by expanding credit opportunities and improving the efficiency of the credit review process (for example, Calem and Wachter 1999).

Lending institutions use two kinds of credit scores: those based solely on the credit histories of individuals, as reflected in credit-bureau records, and those that weigh other factors along with credit history. The former are generally referred to as "bureau scores" or "credit history scores," while the latter are often referred to as "application scores" or "origination scores." This paper focuses on issues related to the use of *bureau scores*, which are increasingly being used as an initial and sometimes primary screen for applicants seeking credit and are also a prescreening tool for credit solicitations.

Despite the clear benefits of credit scoring, the growing reliance on bureau scores in underwriting decisions raises statistical issues. First, the methods

*Board of Governors of the Federal Reserve System, Washington, DC 20551 or ravery@frb.gov.
**Board of Governors of the Federal Reserve System, Washington, DC 20551.

used to develop scoring models raise a concern about the potential effects of omitted variables. This concern is particularly acute for bureau scores, which are constructed only from information on an individual's credit history. Other contemporaneous or historical factors specific to individuals, such as income, employment experience or health status, are not considered. Moreover, in both the development and the use of bureau scores, typically no adjustment is made for local economic conditions (such as a regional recession) that may have affected the history of loan repayment in a local area but may be unrelated to future patterns of repayment. Potentially systematic distortions in risk assessments arising from such omissions could result in lenders accepting different (and unanticipated) levels of credit risk in different parts of the country. A second set of statistical issues that raises potential concerns includes the representativeness of the baseline populations used to develop bureau scoring models and the completeness of bureau files used to develop and apply bureau scoring models. Failure to develop bureau scores using a population representative of the population the credit product is targeted for can compromise the effectiveness of scoring models.

This paper examines these statistical issues and to our knowledge is the first study to do so. For the analysis, we develop a dataset that contains economic and demographic information for a nationally representative sample of ZIP codes and bureau scores for all individuals and households that reside in those ZIP codes. Although we focus on bureau scores, our results have implications for application scores as well, because they also may omit potentially important variables and often use bureau scores as an essential component.

The next section introduces credit scoring and considers the various statistical issues surrounding its use. The following section describes the data used in our analysis. The ensuing section summarizes our empirical results. The final section provides a discussion of our findings.

Statistical Issues Surrounding the Use of Credit Bureau Scores

Bureau scoring models are built on the premise that past performance in repaying debts is the best predictor of future performance. They are designed to (1) rank individuals on the basis of their relative creditworthiness and (2) quantify the likelihood that a given individual will default (become delinquent). Nearly all evaluations of bureau scoring models have focused on their ability to predict relative creditworthiness. Uniformly, this research has found that they are powerful predictors of default and delinquency (see Avery *et al.* 1996; Freddie Mac 1996; Pierzchalski 1996). However, these studies have focused on the relative performance of broad groups. The stability of rank orderings across regions and the robustness of model predictions to changes in local economic conditions have not typically been explored, nor is it clear that the models perform equally well for all subgroups of the population. Indeed, the general success in prediction found in previous studies does not rule out the existence of statistical problems with the development and implementation of credit-scoring models that could lead to incorrect assessments of an individual's risk.

The potential statistical concerns can be highlighted using the following formalization of how bureau scoring models are developed. Assume that the true function for calculating the probability of default (delinquency) for individual i from population j for product k at time t is

$$P_{i,j,k,t} = F_{j,k}(y_{i,j,k_1,t-1}, \ldots, y_{i,j,k_K,t-1}, X_{i,j,t-1}, M_{i,j,t-1}) \qquad (1)$$

where P is the probability of default (delinquency), y_{ki} represents an individual's credit history for product k_i, X are non-credit-history individual characteristics, M are regional and local market conditions, and t-1 represents past realizations. Bureau-scoring models approximate the true function $F_{j,k}$ with the function $F'_{j',k'}$, which is estimated by relating the repayment performance RP, of individuals i' in population j' for product k' to *ex ante* predictive variables for a historical sample:

$$RP_{i',j',k',t-1} = F'_{j',k'}(y_{i',j',k_1,t-2}, \ldots, y_{i',j',k_K,t-2}) + e_{i',j',k',t-1} \qquad (2)$$

where e is an error term. Note that j' and k' may differ from j and k in (1), and $k'_1,\ldots,k'_K$ may be a subset of $k_1, \ldots, k_K$.

Potential problems with this process can take three forms. First, the estimated function may not include variables included in the true function (*i.e.*, X, M or a specific y_{ki}). Second, the credit-history data used to apply the model for a particular individual may contain errors or inaccuracies or be incomplete. That is, instead of using true values of y, a mismeasured $\hat{y}$ may be used. Third, the function F' may have been estimated over a different population j' or product k' than the population j or product k that it will be used to assess.

Problems with Omitted Individual and Geographic Variables

The omission of measures of X and M in the construction and use of bureau scores presents two potential problems. First, if they are valid predictors of future creditworthiness, their omission can create inappropriate rankings of individuals.

A second concern may arise if future realizations of X and M will affect credit performance but current values of X and M are not reliable predictors of future values of X and M. Here, the failure to control for historical values of X and M—which have influenced credit repayment history—in developing the credit scoring model may assign weight to credit-history variables which may more properly be attributed to those economic circumstances. For example, by not including individual economic circumstances, the models implicitly treat someone who performs poorly while unemployed the same as someone with similar performance while employed. Similarly, by not controlling for local economic factors, bureau scoring models assign the same risk level to a person who performs poorly during a recession as to a person with similar performance during better times. It is not clear that these persons should be expected to perform equally well in the future. In general, the uniform application of credit scores without regard to economic circumstances and geographic differences may result in a single score implying different underlying probabilities of future default in different regions.[1]

In some cases, in recognition of this point, lenders adjust cutoff levels according to the risks inherent in local regions. While such adjustments are relatively easy to implement, they do not resolve the fundamental misspecification problem.

Problems Due to Mismeasurement of Credit History

Credit history may be mismeasured because certain products which may be important indicators of creditworthiness may be omitted. For example, rent and utility payment histories have not typically been included in credit scoring models, in part because credit-bureau records have not historically included such items comprehensively.[2] Another issue arises if credit providers only report derogatory information. For example, until relatively recently many mortgage companies only reported delinquent accounts.

---

[1] See Carlino and Sill (1997) and Samolyk (1994) for evidence that economic circumstances and forecasts vary across regions.

[2] These alternative measures are occasionally present in credit-bureau files, but typically only when derogatory information is associated with the account. It is argued by some that the omission of the full records of these alternative measures may result in less effective screening of the creditworthiness of individuals residing in lower-income communities and a reduced flow of credit to such areas. Recognizing this, some lenders allow for the consideration of this information in addition to bureau scores in evaluating credit applications.

Finally, problems may arise if there are data entry or other types of errors or omissions in the credit records used to develop credit-scoring models. Problems of this nature can affect both model development and usage. A prime example of this relates to the completeness of individuals' credit bureau records, often referred to as the *depth* of credit files. Various types of recurring payment obligations, such as those to certain local retailers and neighborhood lenders, have not traditionally been fully recorded by all credit bureaus. Only one or two bureaus might have this information, which can be used as additional indicators of creditworthiness. In model development, reliance on less complete records could distort the decision weights assigned to particular factors. This could result in less reliable relative rankings as well as inaccurate quantification of individual risk.

Problems Due to Inappropriate Populations or Products

Issues can arise if credit-scoring models are developed from populations that are different from the population for which the model is to be used, such as when bureau scores are developed for a general population but are applied to applicants from a nonrandom subset of the population. A related concern is if certain groups are underrepresented in the populations used to develop scoring models, which can occur if group members are less likely to have credit-bureau records (which we call low *coverage*). If either is the case, and if behavior of the underrepresented individuals does not mirror the behavior of the general population, then bureau scores may not accurately reflect the credit risks of such individuals.[3] A particular concern in this respect is that minorities and lower-income individuals may be systematically underrepresented in the baseline populations used to develop the scoring models.[4]

[3] Underrepresentation is also an issue for individuals without credit-bureau files, for such individuals cannot be scored. However, lack of a credit record has always been an obstacle to obtaining credit. A heightened reliance on credit scores may have implications for the availability of credit services, although in some cases this is unlikely to be significant. For example, mortgage lenders often explore other means for evaluating creditworthiness when an applicant does not have a credit bureau file. Coverage issues are also addressed in Fair, Isaac and Company (1996).

[4] Some empirical research has tested the predictiveness of credit-history scores across racial and ethnic groups and among borrowers with differing incomes. Analysis of foreclosure rates on conforming home mortgages purchased by Freddie Mac indicates that credit history scores differentiate borrowers by risk grade regardless of racial or ethnic characteristic or income group. In each case, borrowers with low credit history scores were much more likely to enter foreclosure than borrowers with higher credit history scores (Freddie Mac 1996). Fair, Isaac and Company also has conducted research with a similar focus that produced similar results (Marteli *et al.* 1997).

In addition, bureau scores are developed from payment experience over a given time horizon on all consumer credit, which may create predictive inaccuracies when they are used to forecast performance for narrowly defined products, such as mortgages, or for different time horizons. One example of this would be using bureau scores developed to predict the likelihood of default on mortgages to predict the risk of auto-loan or credit-card applicants. This broad issue is not addressed in this paper.

Implications of Statistical Problems Associated with Credit Scoring Methodologies

The potential effects of these statistical problems raise both economic and regulatory issues. Since lenders generally select score cutoffs based on the highest probability of default (delinquency) they are willing to accept, if bureau scores do not accurately quantify these likelihoods, then marginal applicants may be provided more or fewer services than is appropriate. If risks are understated, loans will be riskier than anticipated; if they are overstated, institutions may be rationing their credit and services unnecessarily. Moreover, many institutions introduce new products or special programs via targeted mailings or other promotional campaigns that are limited in scope. These programs often select individuals by a screening process that includes bureau scores. In such cases, access to or the price of these services for the marginal individual may depend critically on factors such as the region's recent economic performance or other idiosyncratic factors that may have lasting effects on bureau scores.

From a regulatory perspective, the cumulative effects of these potential biases may raise questions about the adequacy of loss reserves and capital. Examiners may find it difficult to evaluate the safety and soundness of portfolios of scored loans. If regional variation in scores exists, should examiners assign different risk weights for identically scored loans within a portfolio? If so, what are the appropriate weights to assign? Only with a better understanding of the sources of variation in credit scores can such questions be answered.

Public-policy concerns extend beyond issues of safety and soundness. For example, compliance with fair-lending laws may be affected by these biases if protected groups are disproportionately affected by statistical error or omissions. Similarly, compliance with the Community Reinvestment Act of 1977 may be affected if problems with scoring models adversely affect the flow of credit to lower-income communities.

Approach and Data

Our empirical analysis explores some of the potential statistical problems highlighted in the previous section. For our analysis, we constructed a geographically based data set including individual bureau score information. A nationally representative sample of ZIP codes was selected by stratifying all U.S. ZIP codes by Census region, suburban–central-city–rural location and median household income in the ZIP code. ZIP codes were randomly selected from each stratum.[5] This process generated a sample of 994 ZIP codes nationwide.

Equifax Credit Information Services, Inc. (Equifax) provided proprietary data on bureau scores as of early 1996 for all individuals with credit files residing in each of the 994 ZIP codes. The bureau score provided by Equifax was The Mortgage Score (TMS), which was developed by Equifax Mortgage Services to predict performance on mortgage accounts. The score is computed from the same variables as typical consumer credit bureau scores and has a very similar aggregate distribution. The final sample includes the credit scores of 3.4 million individuals.

Equifax also provided indicators that allowed for the aggregation of individuals into households (2.5 million) and the identification of whether an individual appeared to hold a mortgage.[6] We use these indicators to conduct separate analyses of individuals, households, and households with

---

[5] We first divided ZIP codes among 27 location-based strata (9 Census regions divided into central city, suburban and rural groups). Within each stratum, ZIP codes were stacked by median household income. A systemic random sample was drawn with an oversampling of smaller ZIP codes, as measured by population. Sampling weights are used throughout our analysis to ensure that our sample statistics are representative of the entire U.S. population.

[6] The existence of a mortgage was inferred by Equifax using an internal algorithm. Comparing the proportion of households identified as holding mortgages in our sample with estimates from national surveys, we find the rate of mortgage holding for our sample is nearly 50% lower (Kennickell, Starr-McCluer and Sunden 1997). The rate of homeownership in the sample is much lower than that estimated from other surveys because mortgages are not explicitly identified in the credit records of many individuals. Because we lack information about the characteristics of those mortgage holders not identified as such in our sample, care must be exercised in interpreting results presented for this group.

[7] To assign a credit score to a household, we conformed to standard industry practices. In cases where a household comprised two individuals, the lower score was taken to be the household score; where there were three individuals, the middle score was used. In the few cases where there were more than three individuals, the average score was used.

mortgages.[7] These are the three main units whose credit score is evaluated using bureau scores—individuals for consumer credit, households for home mortgages, and households with mortgages for mortgage refinancing and home equity loans.

The analysis is conducted at the ZIP code level. We examine the omitted variable issue by exploring how scores vary across ZIP codes with particular characteristics, using two measures: the median score in the ZIP code, and the proportion of individuals (households) in the ZIP code with low credit scores. Additionally, coverage (underrepresentation) issues are explored directly by using a measure that divides the number of individual credit bureau records in each ZIP code by the 1990 Census of Population and Housing estimate of the number of individuals 21 years or older.

We regress these measures on proxies for regional and local market conditions and individual economic circumstances [*i.e.*, M and X from (1)]. Since the ZIP code of the individual's residence is known, we represent M using economic and demographic characteristics of these ZIP codes and the broader region where they are located. These include the vacancy rate in the ZIP code, population growth in the ZIP code's Metropolitan Statistical Area (MSA), and the annual unemployment rate in the ZIP code's county for each of the past five years. These are included explicitly in an attempt to determine whether geographic differences in scores or coverage are a function of differences in local economic experiences. We also include measures of the racial composition of each ZIP code and the location of the ZIP code. Two location measures are used—the *urbanness* of the ZIP code and the Census region in which the ZIP code was located.[8]

Ideally we would also like to consider individual characteristics that may influence repayment probabilities [*i.e.*, X in (1)], such as income, age and the educational attainment of the borrower. However apart from the credit score and household and mortgage status identifiers, no individual-level data were provided by Equifax. We are thus unable to conduct any analyses based on individual characteristics. Instead, we use rough proxies for general individual characteristics created from data on ZIP-code demographics as controls. For example, the median income for the MSA of the ZIP code and the median income of the ZIP code relative to the MSA median income together offer an indication of the relative income of individuals in these areas. Other control variables include the ZIP code's relative median house

[8] *Urbanness* here distinguishes between ZIP codes located in central city, suburban and rural locations.

value and the poverty rate, the rental rate and measures of the educational attainment of residents in the ZIP code.

Finally, the importance of errors-in-variables issues as represented by the varying depth of credit-bureau records is explored by using a bank lookup table, an institution-specific table which provides guidance regarding which credit bureau to approach when making credit inquiries. We use this table to identify those areas where Equifax is likely to provide deeper (more complete) files, and re-estimate the original three regression equations including this proxy. In each case, separate empirical tests are conducted for all individuals, for households and for households with mortgages. U.S. Census statistics on the age of ZIP-code residents are also used as controls in examining the depth of individual credit files; younger applicants are on average less likely to have extensive credit histories.

Table 1 presents some basic statistics about bureau scores and their distribution in our sample for individuals and for households with mortgages. Figures shown are the mean and standard deviation of the median credit score in ZIP codes in a given category.[9] As shown, there is considerable variation in median scores across ZIP codes (for example, a standard deviation of 50.1 for individual scores). Comparing credit scores of individuals with those of households with mortgages, credit scores are consistently higher for households with mortgages, as expected, and the range of scores within a given category is smaller for such households.

Omitted Variables: Variation in Median Credit Scores

The results of regressions for the median credit-bureau score in the ZIP code for individuals and for households with mortgages are shown in Table 2. We do not present results for all households because they turned out to be virtually identical to those for individuals.

Looking first at individuals, the evidence suggests that omitted-variable issues are a potential concern. Controlling for other factors, median scores vary with measures of local economic conditions. ZIP codes located in areas with high county unemployment rates in 1994 have median credit scores 31 points lower than similar ZIP codes in counties with low rates, and this difference is statistically significant. The results also suggest that

[9] Credit scores for individuals or for households with mortgages were aggregated to determine the mean and median score in each ZIP code. Since the analytical results using the mean and median scores were qualitatively the same, only the results using median credit scores are presented.

Table 1 ■ Sample distribution and median score statistics, by ZIP-code characteristics.

ZIP Code Category	ZIP Code % Distribution	Individuals[a] Mean	Individuals[a] S.D.	Households with Mortgages[a] Mean	Households with Mortgages[a] S.D.
Percent of households in poverty					
Greater than 25%	11.1	690.5	50.2	729.4	69.8
Between 10% and 25%	50.0	758.9	37.2	769.3	48.1
Between 5% and 10%	26.2	782.6	33.0	789.5	37.1
Less than 5%	12.6	815.0	36.3	816.1	34.2
Percent of household that are minorities, 1990					
Greater than 25%	16.5	706.4	58.8	742.6	61.1
Between 10% and 25%	21.7	763.3	43.9	783.3	45.0
Between 5% and 10%	13.0	783.3	43.7	789.7	43.1
Less than 5%	48.8	779.8	33.8	780.9	50.0
Urbanization					
Central city	49.9	757.2	43.6	766.5	54.3
Suburban	36.1	778.2	47.3	787.2	44.1
Rural	14.0	755.7	71.0	781.9	58.9
Percent of area median family income, 1990					
Median income less than 80%	20.4	722.8	57.2	749.3	63.5
Median income between 80% and 100%	38.5	736.0	41.5	772.6	49.4
Median income between 100% and 120%	28.2	779.3	35.5	785.3	42.4
Median income greater than 120%	12.8	803.4	38.5	809.3	35.2
Percent of area median house value, 1990					
Median house value less than 80%	34.5	740.4	53.6	754.4	60.3
Median house value between 80% and 100%	51.1	769.2	41.0	778.9	42.9
Median house value between 100% and 120%	14.5	805.9	35.6	817.1	31.1
County unemployment rate, 1994					
More than 9%	12.3	735.3	52.3	758.1	55.3
Between 7% and 9%	18.3	751.1	51.8	762.1	57.8
Between 5% and 7%	36.2	766.5	48.3	777.9	48.0
Less than 5%	33.2	780.8	42.6	789.3	48.6

Table 1 ■ continued

ZIP Code Category	ZIP Code % Distribution	Individuals[a]		Households with Mortgages[a]	
		Mean	S.D.	Mean	S.D.
Census region					
New England	6.1	782.6	41.2	787.5	48.0
Middle Atlantic	12.5	788.9	47.7	791.4	42.1
East North Central	15.5	779.6	32.2	783.2	42.4
West North Central	15.8	784.6	32.5	787.4	60.4
South Atlantic	15.6	744.9	57.9	770.2	50.2
East South Central	7.5	719.5	46.1	736.6	57.8
West South Central	11.1	731.7	45.3	757.8	43.5
Mountain	6.4	764.7	44.1	769.2	57.2
Pacific	9.3	769.5	50.4	788.4	49.5
Median age of adults, 1990					
Less than 30 years	8.2	720.0	57.2	759.1	56.1
Between 30 and 40 years	82.5	765.5	47.7	776.0	51.5
Greater than 40 years	9.2	796.2	32.6	794.4	50.9
High school graduate rate, 1990					
Less than 50%	4.5	693.8	42.6	718.2	60.7
Between 50% and 70%	33.9	734.3	47.5	752.5	52.9
Between 70% and 90%	56.1	784.1	34.4	790.2	41.5
Greater than 90%	5.5	810.4	41.0	828.1	22.8
Percent of individuals over age 60					
Greater than 30%	4.4	795.8	34.1	796.1	48.1
Between 20% and 30%	35.4	771.1	38.8	774.3	55.0
Between 10% and 20%	53.4	759.5	54.1	775.0	50.8
Less than 10%	6.7	749.7	65.2	783.7	50.4
Total	100.0	764.6	50.1	776.3	52.3

[a] Mean and standard deviation of median score across ZIP codes in a category.

Table 2 ■ Median-score regression results.

	Individuals		Households with Mortgages	
	Coeff.	t-Stat.	Coeff.	t-Stat.
Intercept	803.59***	31.1	820.16***	19.9
Population growth rate, 1980–1990				
Growth rate (%)	0.28	1.4	0.42	1.4
Growth-rate spline at 5%	−0.30	−1.0	−0.57	−1.2
Growth-rate spline at 25%	0.03	0.1	0.01	0.0
Housing-unit vacancy rate, 1990				
Vacancy rate (%)	−5.59***	−2.9	−2.36	−0.8
Vacancy-rate spline at 5%	5.73***	2.8	2.58	0.8
Vacancy-rate spline at 25%	−0.32	−0.9	−0.16	−0.3
Housing-unit rental rate, 1990				
Rental rate (%)	0.17	0.5	0.61	1.1
Rental-rate spline at 5%	−0.37	−0.9	−0.45	−0.7
Rental-rate spline at 25%	0.30	0.9	0.59	1.0
MSA median family income, 1990				
MSA median family income ($1000s)	1.81**	2.0	1.03	0.7
MSA-median-family-income spline at 24	−2.16	−1.3	−1.77	−0.7
MSA-median-family-income spline at 28	2.47	1.2	3.26	1.0
MSA-median-family-income spline at 32	−0.52	−0.2	−1.18	−0.3
MSA-median-family-income spline at 36	−1.06	−0.7	−1.58	−0.7
Percent of households in poverty (dummies, <5% base group)				
Greater than 25%	−30.83***	−4.7	−10.32	−1.0
Between 10% and 25%	−12.51**	−2.7	−4.94	−0.7
Between 5% and 10%	−9.25**	−2.4	−7.25	−1.2

Table 2 ■ continued

	Individuals		Households with Mortgages	
	Coeff.	*t*-Stat.	Coeff.	*t*-Stat.
Percent of households that are minorities, 1990 (dummies, <5% base group)				
Greater than 25%	−38.64***	−10.7	−23.69***	−4.1
Between 10% and 25%	−9.21***	−3.2	−0.45	−0.1
Between 5% and 10%	−4.79	−1.5	−1.56	−0.3
Urbanization (dummies, rural base group)				
Central city	−0.57	−0.1	−0.96	−0.1
Suburban	0.32	0.1	0.12	0.0
Percent of area median family income, 1990 (dummies, >120% base group)				
Median income less than 80%	−24.42***	−4.5	−15.08*	−1.8
Median income between 80% and 100%	−10.70***	−2.7	−7.54	−1.2
Median income between 100% and 120%	−6.20*	−1.7	−5.76	−1.0
Percent of area median house value, 1990 (dummies, >100% base group)				
Median house value less than 80%	−26.15***	−6.6	−32.20***	−5.1
Median house value between 80% and 100%	−13.52***	−4.0	−17.83***	−3.3
County unemployment rate, 1994 (dummies, <5% base group)				
More than 9%	−30.81***	−3.8	−27.13**	−2.1
Between 7% and 9%	−7.79**	−2.4	−12.89**	−2.3
Between 5% and 7%	−2.54	−0.9	−0.40	−0.9
Census region (dummies, Pacific base group)				
New England	−13.36**	−2.5	−13.84	−1.6
Middle Atlantic	−4.81	−1.1	−6.30	−0.9
East North Central	−11.71***	−2.6	−13.05	−1.8
West North Central	−3.85	−0.8	−4.48	−0.6
South Atlantic	−18.57***	−4.4	−8.03	−1.2
East South Central	−31.46***	−5.6	−30.44***	−3.4
West South Central	−21.54***	−4.5	−15.54**	−2.0
Mountain	2.87	0.5	−14.54*	−1.7

Table 2 ■ continued

	Individuals		Households with Mortgages	
	Coeff.	*t*-Stat.	Coeff.	*t*-Stat.
Median age of adults, 1990 dummies, >40 years base group)				
Less than 30 years	−15.27**	−2.5	−4.06	−0.4
Between 30 and 40 years	−8.24	−1.9	−4.40	−0.6
High school graduate rate, 1990 (dummies, >90% base group)				
Less than 50%	−25.63***	−3.5	−41.98***	−3.6
Between 50% and 70%	−24.64***	−4.5	−30.23***	−3.4
Between 70% and 90%	−12.96***	−2.7	−14.89**	−2.0
Percent of individuals over age 60 (dummies, <10% base group)				
Greater than 30%	55.56***	7.1	26.73**	2.1
Between 20% and 30%	38.84***	7.4	9.72	1.2
Between 10% and 20%	21.62***	4.5	4.62	0.6
Dynamic county unemployment (dummies, high to low unemployment rate base group)				
Persistent high unemployment rate	20.40**	2.5	17.07	1.3
Persistent low unemployment rate	−0.75	−0.2	−3.20	−0.6
Low to high unemployment rate	−14.67	−0.8	−21.17	−0.7
Memo items				
Mean dependent variable		764.58		776.26
R^2		0.73		0.37
Number of observations		892		879

The dependent variable is the median Equifax mortgage score for individuals (households with mortgages) in the ZIP code. Unless otherwise indicated, independent variables are measured at the ZIP code level. Coefficients are estimates from regressions where each observation is weighted to ensure the sample is nationally representative. The first regression presents estimates where the median score is among all individuals in the ZIP code. The median score for the second regression is among all households in the ZIP code that have a mortgage. All "spline at *X*" variables take the value 0 if the continuous variable has a value of less than *X*, and are equal to the continuous variable minus *X* if the continuous variable has a value greater than *X*. "Dynamic county unemployment" documents cases where extreme levels of the average county unemployment rate over the 1991–1992 period moved to or remained at extreme levels in 1994. A high unemployment rate is a rate above 9%; a low rate is below 5%.

***Significant at 1% level; **significant at 5% level; *significant at 10% level.

unemployment dynamics play a role, although not in expected ways. Compared with ZIP codes in counties with changing or persistently low unemployment rates, areas in counties with persistently high unemployment rates have significantly higher (by 20 points) median credit scores.[10] Thus, an area that has gone from low to high unemployment has lower median scores than an area that has had persistently high unemployment. This relationship may indicate that creditors apply tighter standards in areas with persistently high unemployment rates, dampening the effects of high unemployment rates on borrower delinquency rates. The estimated coefficients for another proxy for economic conditions, the vacancy rate, further suggest that economic conditions are important. However, two other economic variables included in our specification, population growth and the MSA median income, are not found to be significantly associated with differences in credit scores.

Credit scores are also correlated with regional location, as median scores vary significantly across Census regions, all else equal. The magnitude of the largest of these differences, between the Mountain and East South Central regions, is 34 points. By contrast, median scores are nearly identical across urban, suburban and rural ZIP codes after controlling for other factors. This is somewhat surprising, particularly given that urban, suburban and rural areas often have different economic dynamics.

A number of our control variables also show significant relationships with the median credit score in a ZIP code. Median credit scores are significantly lower for areas with a high percentage of people living in poverty, areas with high minority populations, areas with low median incomes, areas with low house values and areas where a large percentage of the population hold high school diplomas.

Among households with mortgages, although fewer ZIP-code characteristics are significantly associated with differences in median credit scores, variables that proxy for economic and regional characteristics remain significantly associated with median credit scores in ZIP codes. Thus, omitted-variable issues are relevant for this subpopulation also.

---

[10] In other specifications, we included other measures of local unemployment conditions, including the average unemployment over the past 5 years, the highest unemployment rate over the past 5 years and whether the most recent level of unemployment was greater or less than the average unemployment rate. Patterns were similar for these estimates.

Omitted Variables: Variation in Proportions with Low Scores

To place the analysis in the context of lending decisions, we grouped TMS credit scores into three categories—high (above 685), medium (603–685) and low (602 and below). These groupings are designed to roughly conform to the three risk ranges for mortgage applicants identified by Freddie Mac and Fannie Mae (Fannie Mae 1995; Freddie Mac 1995).[11] Individuals with scores in the high range are thought, on the basis of credit history alone, to be likely to repay loans, those with medium scores are thought to be less likely to repay loans, and those with low scores are believed to be much less likely to repay loans.

The percentage of individuals with medium or high credit scores (scores above 602) in a ZIP code is regressed on our set of demographic and economic ZIP-code characteristics. These estimates are shown in the first two columns of Table 3. As was the case with median scores, a number of economic, regional and demographic variables are significantly associated with the percentage of individuals with scores above 602. The vacancy rate in the ZIP code and the unemployment rate for the ZIP code's county are both negatively associated with the percentage of individuals with high or medium scores. Both locational measures, Census region and degree of urbanness, also show significant relationships, with urban areas having a higher percentage of individuals with high and medium scores after controlling for other factors.

These results imply that omitted-variable issues may be important in considering the effects of credit scoring on credit flows. However, roughly 68% of individuals and 75% of households with mortgages are creditworthy based on their credit histories. Thus, for most, the effects of omitted variables on access to credit are unlikely to be significant. For the 20% of individuals (15% of households with mortgages) with low TMS scores, however, obtaining credit may be more of a problem.[12]

---

[11] Score thresholds do not perfectly correspond to the Fannie Mae and Freddie Mac cutoffs, because the TMS has a slightly different scale. We aligned the two scoring scales using a dataset which contained both scores for the same population (see Avery *et al.* 1996). Though appropriate for mortgage credit, the thresholds defining the ranges may differ when considering other types of credit. Thus, observed effects may not be generalizable to other credit markets.

[12] During the early 1990s, large numbers of homeowners refinanced their outstanding mortgages. In other data obtained from Equifax, we observed a number of households with mortgages that did not refinance despite having mortgages with interest rates well above those prevailing in the marketplace (Avery *et al.* 1996). Our analysis here suggests that one reason such households may not have refinanced is poor credit histories, as reflected by low credit scores.

Table 3 ■ Percent high/medium scores and coverage regression results for individuals

	% High/Medium Scores		Coverage Ratio	
	Coeff.	t-Stat.	Coeff.	t-Stat.
Intercept	91.19***	19.5	92.83***	5.2
Population growth rate, 1980–1990				
Growth rate (%)	0.02	0.7	0.07	0.5
Growth-rate spline at 5%	−0.08	−1.4	0.08	0.4
Growth-rate spline at 25%	0.09**	2.1	−0.01	−0.1
Housing-unit vacancy rate, 1990				
Vacancy Rate (%)	−0.99***	−2.8	−1.22	−0.9
Vacancy-rate spline at 5%	1.06***	2.9	0.67	0.5
Vacancy-rate spline at 25%	−0.11*	−1.7	0.37	1.6
Housing-unit rental rate, 1990				
Rental rate (%)	−0.04	−0.6	−0.36	−1.6
Rental-rate spline at 5%	−0.05	−0.7	0.56*	1.9
Rental-rate spline at 25%	0.19***	3.0	−0.91***	−3.7
MSA median family income, 1990				
MSA-median family income ($1000s)	0.14	0.9	0.33	0.5
MSA-median-family-income spline at 24	−0.24	−0.8	−0.27	−0.2
MSA-median-family-income spline at 28	0.24	0.7	−0.15	−0.1
MSA-median-family-income spline at 32	0.18	0.5	0.33	0.2
MSA-median-family-income spline at 36	−0.21	−0.8	−0.45	−0.4
Percent of households in poverty (dummies, <5% base group)				
Greater than 25%	−6.77***	−5.7	−8.35*	−1.9
Between 10% and 25%	−2.35***	−2.8	−5.18	−1.6
Between 5% and 10%	−1.85***	−2.6	−4.04	−1.5

Table 3 ■ continued

	% High/Medium Scores		Coverage Ratio	
	Coeff.	t-Stat.	Coeff.	t-Stat.
Percent of households that are minorities, 1990 (dummies, <5% base group)				
Greater than 25%	−7.92***	−12.1	−2.73	−1.1
Between 10% and 25%	−2.40***	−4.5	0.27	0.1
Between 5% and 10%	−2.14***	−3.7	−0.71	−0.3
Urbanization (dummies, rural base group)				
Central city	1.69**	2.1	−8.67***	−2.8
Suburban	0.56	0.9	−3.30	−1.4
Percent of area median family income, 1990 (dummies, >120% base group)				
Median income less than 80%	−3.08***	−3.2	2.39	0.7
Median income between 80% and 100%	−1.02	−1.4	2.61	0.9
Median income between 100% and 120%	−0.82	−1.2	0.31	0.1
Percent of area median house value, 1990 (dummies, >100% base group)				
Median house value less than 80%	−3.56***	−4.9	−8.91***	−3.3
Median house value between 80% and 100%	−2.10***	−3.4	−5.39**	−2.3
County unemployment rate, 1994 (dummies, <5% base group)				
More than 9%	−4.46***	−3.0	−1.14	−0.2
Between 7% and 9%	−1.53***	−2.5	0.79	0.4
Between 5% and 7%	−0.65	−1.3	1.70	0.9
Census region (dummies, Pacific base group)				
New England	−2.44***	−2.5	7.35**	2.0
Middle Atlantic	−1.01	−1.3	−1.03	−0.3
East North Central	−1.40*	−1.7	−2.44	−0.8
West North Central	−1.71*	−1.9	−0.18	−0.1
South Atlantic	−4.91***	−6.4	0.34	0.5
East South Central	−6.72***	−6.5	0.26	0.1
West South Central	−4.48***	−5.0	4.63	1.4
Mountain	1.16	1.2	1.58	0.4

Table 3 ■ continued

	% High/Medium Scores		Coverage Ratio	
	Coeff.	t-Stat.	Coeff.	t-Stat.
Median age of adults, 1990 (dummies, > 40 years base group)				
Less than 30 years	−0.97	−0.9	−0.34	−0.1
Between 30 and 40 years	−1.34*	−1.7	−3.95	−1.3
High school graduate rate, 1990 (dummies, > 90% base group)				
Less than 50%	−3.81***	−2.9	−12.35***	−2.5
Between 50% and 70%	−4.47***	−4.4	−5.10	−1.3
Between 70% and 90%	−3.01***	−3.5	−2.31	−0.7
Percent of individuals over age 60 (dummies, <10% base group)				
Greater than 30%	9.88***	6.9	−2.94	−0.5
Between 20% and 30%	7.74***	8.1	1.32	0.4
Between 10% and 20%	5.09***	5.9	−0.56	−0.2
Dynamic county unemployment (dummies, high to low unemployment rate base group)				
Persistent high unemployment rate	2.22	1.5	2.91	0.5
Persistent low unemployment rate	0.70	1.2	−2.46	−1.1
Low to high unemployment rate	−3.46	−1.0	17.50	1.3
Memo items				
Mean dependent variable		80.34		66.13
R^2		0.71		0.28
Number of observations		892		892

The dependent variable for the first regression is the percentage of individuals in the ZIP code with Equifax mortgage scores above 602. Unless otherwise indicated, independent variables are measured at the ZIP code level. The dependent variable for the second regression is the number of individual credit-bureau records in the ZIP code divided by the 1990 Census of Population and Housing estimate of the number of individuals in the ZIP code aged 21 or older. Coefficients are estimates from regressions where each observation is weighted to ensure the sample is nationally representative. All "spline at X" variables take the value of 0 if the continuous variable has a value less than X and if the continuous variable has a value greater than X. "Dynamic county unemployment" documents cases where extreme levels of the average county unemployment rate over the 1991–1992 period moved to or remained at extreme levels in 1994. A high unemployment rate is a rate above 9%; a low rate is below 5%.

*** Significant at 1% level; **significant at 5% level; *significant at 10% level.

Underrepresentation: Analysis of Coverage

To investigate coverage, we divide the count of credit-bureau records in a ZIP code by the 1990 Census population estimates of the number of persons age 21 or older for that ZIP code and regress this coverage ratio on our locational and economic factors.[13] Such a test suffers from two limitations. First, population estimates are dated; actual 1995 ZIP-code populations may differ from 1990 estimates (which were the most current available). Second, the age 21 cutoff is arbitrary; some individuals establish credit at an earlier age, and others begin their use of credit at later ages. The net effect of these limitations on our estimates is difficult to predict.

For the 994 ZIP codes in our sample, the aggregate coverage ratio for individuals is 0.661.[14] This ratio masks considerable variability across ZIP-code characteristics. Regression results show substantial variation in coverage across ZIP codes with populations with different degrees of education, poverty rates and urbanness, and with different relative house values (last two columns of Table 3).[15] Very poor, less educated, and more urban populations appear to be less represented in the population of credit-bureau files. Significantly, neither age or minority composition nor relative family income appears to be related to credit bureau coverage.

The analysis suggests that concerns raised about the possible effects of incomplete coverage on credit access for minorities and individuals with lower relative income may be somewhat overstated. However, the significant variation in coverage rates by education, poverty rate and urbanness suggests that alternative methods of evaluating credit histories may have value in many lending contexts.[16]

Errors in Variables: Analysis of Depth

In the analysis in Table 2, age of population in the ZIP code is one measure which proxies for the depth of the bureau file. The results show that age is

[13] This analysis does not consider coverage among households with mortgages, because the 1990 Census data by ZIP code did not include information on the number of households with mortgages.

[14] The comparable ratio for all households was 0.920.

[15] To minimize the effect of outliers on the regressions, the sample was pruned of those ZIP codes with extremely high or low coverage ratios. This reduced the sample by approximately 100 ZIP codes.

[16] For example, to be effective, affordable-housing programs may need to rely on information not routinely included in credit-bureau files, such as rent and utility payment histories.

an important covariate with median credit score. ZIP codes whose older population (over 60 years of age) exceeded 30% of the total population have much higher median scores (by 56 points) than those in which the percentage was under 10%. Additionally, even after controlling for the percentage of older individuals, ZIP codes with young populations have much lower scores. This may reflect the fact that younger people might lack sufficient credit experience to establish a strong credit record or that younger people are more susceptible to employment disruptions and have fewer assets to draw upon in the face of financial difficulties (Kennickell, Starr-McCluer and Sunden 1997).

To explore this issue further, we repeated the individual regressions presented in Tables 2 and 3, adding an additional proxy for file depth to the specifications. This proxy was developed using a bank lookup table—an institution-specific table which provides guidance to underwriters regarding which credit bureau to approach when making credit inquiries for 1997. Based on the location of the applicant, the table rank-orders the three national credit bureaus when making credit inquiries. The lookup table we use ranks the credit bureaus on a state-by-state basis, although other breakdowns, such as at the ZIP-code level, are used by some institutions.

We take the rankings in the lookup table as a signal of lender beliefs about which credit repository has the best information about the applicant's credit history, which we interpret to be the most information about the applicant's history. Given this assumption, we construct a dummy variable which equals one if the ZIP code is located in a state in which Equifax is ranked as the best credit bureau and zero otherwise. Essentially, this variable is a flag suggesting whether the credit scores in the ZIP code are likely to be based on more complete credit-bureau records.

The depth proxy has no significant effect on the median score of individuals in a ZIP code, holding all else equal (regression results not shown). By contrast, depth does have a statistically significant effect on the proportion of individual scores in the medium and high range in a ZIP code. However, the size of the estimated coefficient is quite small, as depth is associated with only a 1.4 percentage-point drop in this proportion (the mean proportion in the overall sample is 80.3%). Similarly, depth is significantly associated with coverage, but the magnitude of the effect is again relatively small. Coverage is only about 3 percentage points higher in states where Equifax is thought to have better records. In all of these regressions, all other relationships are essentially the same as above.

Discussion

The results, which show significant variation in bureau scores across a number of economic, locational and demographic characteristics, suggest that omitted variable and, possibly, depth concerns warrant attention. Moreover, the results of the coverage test suggest that individuals with very low income or little education, but not those in minority groups, may be underrepresented in data used to develop credit-history scoring models. These results have economic and policy implications and raise questions deserving further study.

Omitted-Variable Issues

Although bureau scores are often developed without reference to local economic or broader regional information, our results suggest that they may be affected by such factors. The source of this variation in scores is not well understood. It could reflect inherent population differences in credit risk or be the legacy of differences in past experiences which are not necessarily predictive of future performance. Failure to make any adjustment for these factors is equivalent to assuming that all of these differences stem from the former. If this is not true, then failure to make adjustments may lead to inefficiencies.

One potential response to this problem is the development of credit-history scorecards based on subpopulations stratified by economic and regional considerations. To date, however, such information has generally not been used for the development of national scorecards, which are widely used in mortgage lending and other applications. One reason for this is that appropriate methods for incorporating macroeconomic and regional information into credit scoring models have not been established. This is a potential area for future research.

Another potential response is to adjust the use of scores according to the economic and regional context. For example, lenders may want to shift credit score thresholds for individual applicants who are unexpectedly affected by idiosyncratic events. Anecdotal information suggests that this is done in some circumstances. For instance, in mortgage lending, creditors routinely consider extenuating circumstances before denying credit to individuals with scores below established thresholds. However, in other cases, such as solicitations for consumer credit, such adjustments are less common.

Further, creditors may use bureau scores' predicted default (delinquency) likelihoods to help determine the interest rates charged to individual borrowers and to price portfolios. Our evidence suggests that such use of credit scores, in the absence of adjustments based on economic and regional information, could result in inappropriate pricing decisions.

Depth and Coverage Issues

Our results indicate that coverage and, by extension, representation may be an important issue for some segments of the population. For these underrepresented groups—the very poor, the less educated, and those in the central city—scoring models developed using credit-bureau records may not provide as accurate an indicator of risk as they do for other population groups, although to the extent that underrepresented groups perform similarly to the general population, scoring models will be equally powerful in assessing risks across such individuals. Future research might therefore focus on a comparative analysis of repayment behavior between underrepresented populations and the population at large. Importantly, our research suggests that some groups commonly thought to be underrepresented in credit-bureau records—individuals in high minority or lower income areas and areas with younger populations—do not appear to be so.

Our analysis further suggests that depth is not an important issue for most of the population. Depth is not associated with median score differences and only has a slight association with the percentage of individuals in a ZIP code with high or medium scores. Thus depth appears to be relevant exclusively at the margin; those with scores very close to credit risk cutoffs could be affected. Such individuals would benefit in cases when files with less depth are used and experience adverse effects when deeper files are referenced. Additionally, ZIP codes in states where credit-bureau records in our sample are likely to be deep had slightly higher coverage than other ZIP codes. This suggests that coverage issues and concerns are reduced to the extent that creditors and model developers use the deepest files.

Other Issues

Our analysis found a statistical relationship between bureau scores and the ZIP-code minority composition, after controlling for other locational characteristics. Consequently, concerns about potential *disparate impact* and its effects on the pricing and distribution of financial services merit further study.[17]

We caution, though, that simply finding such a statistical relationship (correlation) does not imply that disparate impact exists. The relationship between minority composition and credit scores may reflect either (1) minorities simply having, on average, weaker overall credit histories and being more likely to reside in neighborhoods with high minority compositions, or (2) their joint correlation with credit-related factors omitted from the scoring model, such that if these factors were accounted for the conditional correlation between the score and minority composition would disappear. Only in the second instance would disparate impact be a potential concern. Resolution of this question cannot be achieved with our data set and is an area for additional research.

Concluding Caveats

In considering the results, several caveats must be kept in mind. First, our analysis relies exclusively on geographic-based information, while many of the issues regarding scoring models concern individual characteristics. Variation in individual characteristics may account for much of the credit-score–locational relationships we observe. In order to fully separate the effects of individual demographics from locational effects one would need a data set that contained both types of information. For both legal and other reasons, such data is unavailable. This study therefore uses the most complete data set available under current circumstances.

Second, when lenders make credit decisions, in most cases they rely on more than simply a bureau score. Such considerations will lessen the relative importance of bureau scores and any attendant effects of reliance on such scores. Finally, in spite of the various concerns voiced regarding credit scoring, judgmental systems also have limitations which, in some cases, may make them inferior to credit scoring. Judgmental systems may be less consistent, may involve greater costs and may be less predictive than scoring models. Moreover, it may be more difficult to ensure that all loan officers comply fully with fair lending rules.

---

[17] Disparate impact arises when a facially neutral practice adversely affects members of a protected class. If scores (or credit outcomes) are correlated with membership in a protected group, such practices may be defended if they can be shown to be justified by business necessity. Even if justified, a practice may still be deemed discriminatory if it can be shown that an alternative practice would achieve the same business ends with a less discriminatory impact.

We thankfully acknowledge the able research assistance of John Matson and the comments of two anonymous referees. An earlier version of this paper was presented at the 33rd Annual Conference on Bank Structure and Competition at the Federal Reserve Bank of Chicago. The views expressed are those of the authors and do not necessarily represent the views of the Board of Governors of the Federal Reserve System or its staff.

References

Avery, R.B., R.W. Bostic, P.S. Calem and G.B. Canner. 1996. Credit Risk, Credit Scoring, and the Performance of Home Mortgages. *Federal Reserve Bulletin* 82: 621–648.

Calem, P.S. and S.M. Wachter. 1999. Performance of Mortgages in a Community Reinvestment Portfolio. *Real Estate Economics* 27: 105–134.

Carlino, G.A. and K. Sill. 1997. Regional Economies: Separating Trends from Cycles. *Business Review, Federal Reserve Bank of Philadelphia* (May/June): 19–31.

Fair, Isaac and Company, Inc. 1996. Low to Moderate Income and High Minority Area Case Studies. Discussion Paper. Fair, Isaac and Company: San Rafael, CA.

Fannie Mae. 1995. Measuring Credit Risk: Borrower Credit Scores and Lender Profiles. Fannie Mae Letter LL09-95. Fannie Mae: Washington, DC.

Freddie Mac. 1995. The Predictive Power of Selected Credit Scores. Freddie Mac Industry Letter. Freddie Mac: McLean, VA.

——. 1996. Automated Underwriting: Making Mortgage Lending Simpler and Fairer for America's Families. Freddie Mac Report (September). Freddie Mac: McLean, VA.

Kennickell, A.B., M. Starr-McCluer and A.E. Sunden. 1997. Family Finances in the U.S.: Recent Evidence from the Survey of Consumer Finances. *Federal Reserve Bulletin* 83: 1–24.

Marteli, J., P. Panichelli, R. Strauch and S. Taylor-Schoff. 1997. The Effectiveness of Scoring on Low-to-Moderate-Income and High-Minority Area Populations. Fair, Isaac and Company Inc. Research Paper. Fair, Isaac and Company: San Rafael, CA.

Pierzchalski, L. 1996. Guarding against Risk. *Mortgage Banking* (June): 38–45.

Samolyk, K.A. 1994. Banking Conditions and Regional Economic Performance: Evidence of a Regional Credit Channel. *Journal of Monetary Economics* 34: 259–278.