Joint Center for Housing Studies

WORKING PAPER SERIES

Credit Information Reporting and the Practical Implications of Inaccurate or Missing Information in Underwriting Decisions

Robert B. Avery, Paul S. Calem, and Glenn B. Canner

BABC 04-11

February 2004

Graduate School of Design

John F. Kennedy School of Government

HARVARD UNIVERSITY

Joint Center for Housing Studies

Harvard University

Credit Information Reporting and the Practical Implications of Inaccurate or Missing Information in Underwriting Decisions

Robert B. Avery, Paul S. Calem, and Glenn B. Canner

BABC 04-11

February 2004

This paper was produced for *Building Assets, Building Credit: A Symposium on Improving Financial Services in Low-Income Communities*, held at Harvard University on November 18-19, 2003.

Robert B. Avery is a Senior Economist for the Board of Governors of the Federal Reserve System. Paul S. Calem is a Senior Economist for the Board of Governors of the Federal Reserve System. Glenn B. Canner is a Senior Adviser for the Board of Governors of the Federal Reserve System

The authors would like to thank Shannon Mok and Jonathan Hershaff for their excellent research assistance. The views expressed are those of the authors and do not necessarily represent those of the Board of Governors of the Federal Reserve System or its staff.

Abstract

Concerns about the accuracy, timeliness, and completeness of information gathered by credit reporting agencies have grown as credit evaluation and decisions on access to other services have become ever more reliant on credit history scores and other automated systems for decision-making. Previous research indicates that key aspects of credit record data may be incomplete, duplicative, or ambiguous. This study assesses the degree to which such data limitations may potentially benefit or harm "at risk" consumers (those at the margin of creditworthiness).

The analysis suggests that many data quality issues have already been recognized and accounted for by credit score card developers and, thus "corrections" of these data items have little impact on the credit scores of "at risk" individuals. We show, though, that other data quality issues are likely to have a more significant impact on credit availability either because they potentially affect a larger portion of the "at risk" population or because of the magnitude of their effects on the credit scores of the individuals impacted.

Introduction

Credit reporting agencies gather information from creditors and other entities to develop a comprehensive and contemporaneous picture of the credit relationships of individuals primarily to facilitate credit evaluation. Available evidence indicates that these data and the credit history scoring models derived from them have substantially improved the overall quality of credit decisions while simultaneously reducing the costs of such decision-making. There is little doubt that consumers on the whole would receive less credit, and what they do receive would be at higher prices if not for the information provided by credit reporting agencies. Moreover, the credit reporting system has become more comprehensive over the past decade or so with notable improvements, such as the adoption of common formats for reporting information and enhanced reporting of mortgage credit.

Despite the benefits the credit reporting system offers, concerns have been raised about the completeness and accuracy of credit reporting agency records and the effects of these shortcomings on the ability of individuals to obtain credit, or on the prices they pay for such credit. These concerns have grown as the underwriting and pricing of consumer credit, and, to a lesser extent mortgage credit, have increasingly been based on credit history scores rather than on judgmental reviews of the underlying data in credit reports. Prescreening and initial solicitations for credit, as well as account monitoring, also have become particularly reliant on credit scores. Finally, concerns about data quality have been elevated as the use of credit records has grown in other areas, such as in insurance underwriting, employment, and housing.

Clearly, some minimum degree of accuracy and completeness of credit reports is required for the benefits of the credit reporting system to be realized. Moreover, the more accurate and complete the information assembled by credit reporting agencies, the greater the potential to enhance efficiency in the credit granting process, reducing costs to the advantage of both consumers and creditors. From a public policy perspective, it is necessary to determine whether the incidence of errors and their impact on the actual credit granting process is significant enough to warrant consideration of new regulatory initiatives.

Over the years, a number of studies have focused on the contents of credit records but have reached quite different conclusions about the degree of accuracy and completeness of such information and their importance on credit availability and pricing of certain types of data problems. This study is an attempt to build on and extend this earlier research. Specifically, we

explore the nature and extent of data limitations in credit reports and their potential effects on the availability and pricing of credit using a large, nationally representative sample of credit records from one of the three national credit reporting agencies.[1] Because changes in the assessment of credit risk are most likely to have a meaningful effect on the availability of credit for individuals at the margin, the analysis focuses on the group of individuals in the credit score range that marks the transition between the subprime and prime markets. Specifically, we estimate the proportion of individuals in this credit score range that are likely to be materially effected by a number of data limitations and quantify the likely effect of each limitation on the credit scores of individuals in this credit score range. These limitations include: (1) the non-reporting of credit limits; (2) ambiguity about the current status of accounts that show positive balances but are not currently reported; (3) the failure of some creditors to report nonderogatory accounts or minor delinquencies; and (4) potentially duplicative or inconsistent reporting of collection agency and public record accounts and incomplete data on inquiries.

The following section provides background information on issues related to the contents and quality of the information contained in credit records. This is followed by a discussion of the contents of the sample of credit records obtained by the Federal Reserve Board and aspects of the data that raise particular concerns with regards to completeness and accuracy. The next section describes the methodology used in this paper to assess the effects of data quality issues on the availability of credit. Empirical findings follow. The paper concludes with a discussion of the implications of our research.

Background: Content, Regulation and Quality of Credit Reporting Agency Records

Credit reporting agencies gather information on the experiences of individuals with credit, leases, non-credit-related bills, monetary-related public records, and credit inquiries, and compile it in a credit record.[2] The three national credit reporting agencies, Equifax, Experian, and Trans Union, each attempt to collect comprehensive information on all lending to individuals in the United States, and, as a consequence, the information maintained by each is vast. Each of the three national credit reporting agencies has records on perhaps as many as 1.5 billion credit

[1] For further details see, Robert B. Avery, Raphael W. Bostic, Paul S. Calem, and Glenn B. Canner, "An Overview of Consumer Data and Credit Reporting," *Federal Reserve Bulletin*, vol. 89, (February 2003), pp. 47-73.
[2] Non-credit-related bills include charges for such items as utility and medical services.

accounts held by approximately 210 million individuals.[3] Based on these data, the three national credit reporting agencies generate over 3 million credit reports each day. These reports are provided to creditors, employers, insurers and consumers themselves.[4] The reports, along with credit history scores derived from the credit reporting agency records, are a primary factor considered in credit evaluation and pricing decisions.

Credit reporting agencies collect information primarily from creditors, governmental entities, collection agencies, and third-party intermediaries generally every month, and they typically update their credit records within one to seven days of receiving new information. According to industry sources, each of the three national credit reporting agencies receives more than 2 billion items of information each month. Reporting entities submit information to credit reporting agencies on a purely voluntary basis; no state or federal law requires creditors or others to report data to the agencies.

Access to the information and maintenance of each credit record is governed by conditions spelled out in the Fair Credit Reporting Act (FCRA).[5] Under the FCRA, a consumer report can only be accessed by persons with a permissible purpose for obtaining the report, for example, in connection with a credit transaction, for employment purposes, or to underwrite insurance involving the individual. The FCRA prohibits a reporting institution from furnishing any information to a credit reporting agency if the institution knows or consciously avoids knowing that the information is inaccurate, and it requires institutions to participate in the process of correcting errors that are identified by consumers. A credit reporting agency must remove or correct inaccurate, incomplete, or unverified information from a consumer's credit record within 30 days after a dispute is filed. Also, the FCRA requires that those taking adverse action against a consumer based on a consumer report (for example, denying a request for credit) give notice to the consumer that a report was used in the decision. Consumers who are denied

[3] Testimony by John A. Ford, Chief Privacy Officer of Equifax Inc. before the House Committee on Financial Services Subcommittee on Financial Institutions and Consumer Credit, Hearings on the Fair Credit Reporting Act, June 4, 2003. Also see, "About CDIA," on the Consumer Data Industry Association web site www.cdiaoline.org.

[4] It is estimated that consumers request about 16 million credit reports each year. See, Loretta Nott and Angle A. Welborn, "A Consumer's Access to a Free Credit Report: A Legal and Economic Analysis," Report to the Congress by the Congressional Research Service, September 16, 2003, pp. 1-14.

[5] For a discussion of how the FCRA governs and encourages accurate credit reporting see Michael Staten and Fred Cate, "Does the Fair Credit Reporting Act Promote Accurate Credit Reporting," Paper presented at Building Assets, Building Credit: A Symposium on Improving Financial Services in Low-Income Communities, Joint Center for Housing Studies, Harvard University, November 18-19, 2003.

credit based on their credit reports are entitled to receive a free copy of their reports.[6] At the end of November 2003 the Congress amended the FCRA to expand consumer access to credit reports and credit scores, and to address issues of data accuracy and identity theft.[7]

Credit reporting agencies use a range of techniques and edit procedures to process the large volume of information they receive. When a credit reporting agency receives data from a creditor, government agency, or third-party provider, it first assesses its accuracy and completeness. If the data are found to or suspected to contain errors, they are returned to the reporting entity for resubmission with any necessary corrections.[8] Otherwise, the credit reporting agency compiles and reconfigures the newly received data to create or update the record of an individual's credit experiences. This reconfiguration can require technical sophistication. For example, credit reporting agencies have had to develop rules for deciding when to ignore variations in personal identifying information and techniques for recognizing that sometimes data items reported with the same identifying information, such as name, may actually be associated with different individuals.

Although credit reporting agencies' data are extensive, they are not complete. First, information on some credit accounts held by individuals is not reported. For example, some smaller retailers, mortgage and finance companies, and some government agencies do not report to the credit reporting agencies. Loans extended by individuals, employers, insurance companies, and foreign entities typically are not reported. Second, complete information is not always provided for each account reported. Sometimes creditors do not report, or update, information on the credit accounts of borrowers who consistently make their required payments as scheduled. Similarly, credit limits established on revolving accounts are not always reported or updated. Also, creditors may not notify the credit reporting agencies when an account is closed, transferred, or has had a change in payment status. For example, sometimes creditors do not report minor delinquencies and only report information on a change in payment status when a more serious payment problem arises.

---

[6] About 85 percent of the 16 million credit reports received by consumers each year are associated with an adverse action. See Loretta Nott and Angie A. Welborn, "A Consumer's Access to a Free Credit Report: A Legal and Economic Analysis," p. 10.

[7] Amendments to the FCRA are spelled out in the new law referred to as the "Fair Credit Transactions Act of 2003.

[8] For example, if a reporter submits a file that includes a much larger or smaller number of records than have historically been received the file would be flagged for review. Similarly, if the review of individual record items suggests unexpectedly large or small percentages of the data items have a given characteristic (for example, the number of accounts 60 or more days late exceeds a designated threshold) then the data are flagged for review.

The information reported on credit accounts reflects each account's payment status and outstanding balance shortly before it is forwarded to the credit reporting agencies. Thus, the information in a report is sensitive to the date on which the information is forwarded to the agencies. For example, a credit account reported to the credit reporting agencies on the day after a payment is posted to the account will show a smaller balance than one reported to the agencies on the day before the payment. Similarly, the payment status on an account reflected in a credit report is sensitive to timing; the record on an account may indicate no late payment problems on a given day but may reflect a delinquency if reported one or two days later.

In addition to issues about the timeliness and completeness of information in a given credit record, other concerns arise because information about an individual may differ across the three national credit reporting agencies. Such differences may arise because rules regarding the linkage of reports to a common individual and the treatment of items such as non-current data can vary across credit reporting agencies. Further, the credit record of an individual at any given point in time may differ some across the three agencies because the timing of receipt and posting of information differs among the three firms. Credit records may also differ across the agencies because not all reporters provide information to each agency. For example, a given creditor or collection agency may report to one or two of the credit reporting agencies, but not to all three. Finally, differences arise because changes made to disputed information may only be reflected in the credit records of the agency that received the disputed claim.

Although credit reporting agencies endeavor to maintain high-quality data, the degree to which consumer credit reports are accurate, complete, or consistent across agencies is in dispute. Similarly, there is disagreement over the importance of data errors and omissions when they exist. A recent study by Consumer Federation of America and the National Credit Reporting Association, for example, found evidence that the information included in the credit reports of any given individual can differ widely across the national credit reporting agencies.[9] This study found that credit scores based on data from the three national credit reporting agencies can vary substantially regardless of whether the individual had a generally good or bad credit history, and, as a consequence, "millions of consumers are at risk of being penalized by inaccurate credit

[9] "Credit Score Accuracy and Implications for Consumers," Report by Consumer Federation of America and the National Credit Reporting Association, December 17, 2002.

report information and inaccurate credit scores." [10] These inconsistencies arise from differences in the timing of receipt of information from reporters and the failure of some reporters to supply information to all three agencies. Inconsistencies also arise due to the different procedures used by the three agencies in assembling a unified credit record for an individual.

Errors and omissions also occur within the credit files of an individual credit reporting agency. An earlier investigation by a consumer organization suggests that as many as 70 percent of credit reports contain an error of some sort, and that one-third of all consumer credit reports may contain errors that could result in the denial of access to credit.[11] A study by Arthur Andersen & Company argues, however, that many errors may not have material significance on access to credit. The Andersen study, which reviewed the outcomes for individuals who were declined credit and then disputed information in their credit reports, concluded that only a small proportion of these individuals were denied credit on the basis of inaccurate information in their credit reports.[12]

Overall, research and creditor experience have consistently indicated that credit reporting agencies' information, despite any limitations that it may have, generally provides an effective measure of the relative credit risk posed by prospective borrowers.[13] Although credit reporting agency data are, on average, highly predictive of future payment behavior, data limitations may result in an incorrect risk ranking for any given individual and as a consequence effect that individual's access to or the price paid for credit. It should be noted that misclassification due to limitations of credit agency records can harm or benefit a given consumer. Some may be judged worse risks than is warranted while others are judged better than warranted. A general consensus exists that consumers should review their credit reports periodically, especially if they are in the

[10] The study found that the difference between the high and low credit score for an individual across the three national reporting agencies averaged 41 points, and that about 4 percent of individuals had score differences of 100 points or more.

[11] Jon Golinger, and Edmund Mierzwinski, "Mistakes Do Happen: Credit Report Errors Mean Consumers Lose," U.S. Public Information Research Group, March 1998, available at website, www.uspirg.org.

[12] See Consumer Data Industry Association, Press Release, March 12, 1998; also see Robert M. Hunt, "The Development and Regulation of Consumer Credit Reporting in America," Federal Reserve Bank of Philadelphia, Working Paper No. 02-21, November 2002. The study found that 8 percent of the consumers who were denied credit requested a copy of their credit reports. Of these consumers, 25 percent found and disputed errors. Of those consumers finding errors, about 12 percent (3 percent of those requesting a credit report) eventually received credit as a result of their dispute.

[13] See Robert B. Avery, Raphael W. Bostic, Paul S. Calem, and Glenn B. Canner, "Credit Risk, Credit Scoring, and the Performance of Home Mortgages," *Federal Reserve Bulletin* (July 1996), pp. 621-48.

market for new credit, if they have been denied credit, or if their creditor has changed the terms of an account on the basis of credit reporting agencies information.

Information Obtained by the Federal Reserve

To better understand the nature and content of credit files, the Federal Reserve Board obtained the full credit records (excluding any identifying personal or creditor information) for a nationally representative random sample of 248,000 individuals as of June 1999 from one of the national credit reporting agencies.[14] Information included in the credit records was provided by many thousands of entities, including more than 23,000 creditors reporting at the time the sample was drawn. A credit history score was provided for 203,000 individuals in the sample. The credit score was based on the credit reporting agency's proprietary credit risk scoring model as of the date the sample was drawn. This score is comparable to other commonly used consumer credit scores, with larger values indicating greater creditworthiness. Although the personal characteristics and address of the individual were not included in the sample data, the census tract, state, and county of residence were made available.[15]

The Federal Reserve sample (approximately 1 in 657 individuals from the credit reporting agency records) contains virtually all the information that would be available for the estimation and application of a credit history scoring model. This includes four general types of information: (1) detailed information reported by creditors (and some other entities such as utility companies) on current and past loans, leases, or non-credit-related bills, each of which is referred to here as a credit account; (2) information derived from monetary-related public records, such as records of bankruptcy, foreclosure, tax liens (local, state, or federal), garnishments, and other civil judgments, that are referred to here as public records; (3) information reported by collection agencies on actions associated with credit accounts and non-credit-related bills (such as unpaid medical or utility bills), referred to as collection agency accounts; and (4) the identities of individuals or agencies that request information from an individual's credit record, the date of the inquiry, and an indication of the purpose of the inquiry.

[14] For details about the sample of credit reporting agency data, see Avery, Bostic, Calem, and Canner (2003).

[15] Credit reporting agency files include personal identifying information that allows the agencies to distinguish among individuals and construct a full record of each individual's credit-related activities. Files include the consumer's name(s), current and previous addresses, and social security number. Other personal characteristics sometimes found in credit files include date of birth, telephone numbers, name of spouse, number of dependents, income, and employment information.

Not every individual in the sample had information of each type. Indeed, individuals can be in credit reporting agency files for a number of reasons: having a record of a credit account, being an authorized user on a credit account, having a monetary-related public record, having a record of a collection action, or having had an inquiry about their credit circumstances. Approximately 87 percent of the individuals in the Federal Reserve sample had a record of a credit account, and most of these (92 percent) had an open and active account as of the date the sample was drawn (table 1). A very small share of the individuals (well less than 1 percent) with a credit reporting agency file had only a public record item or only a record of an inquiry. However, about 11 percent of the sample had a credit reporting agency file only because of a collection action.

As will be shown below, examination of the data included in consumer reporting agency files reveals that the information is not complete in all regards and at times contains duplications and ambiguities. These omissions and limitations require users of the information to make assumptions about how to treat certain reported items in developing a credit profile for a consumer. These issues arise whether a user of the data is judgmentally reviewing the full credit file or relying on a credit history score to assess the individual's creditworthiness. In the latter case, however, the score card developer has addressed these issues in one fashion or another in developing the statistical model that underlies the credit history score.

The following discussion reviews the more important of these data limitations and quantifies their scope. Because the data are now somewhat dated, some of the patterns presented here may not reflect current circumstances.

Contents of Credit Records

To better understand issues related to the timeliness and completeness of credit records it is useful to begin with a review of the basic information that is included in such data. Because the majority of information related to an individual's credit history record is contained in the credit account, collection agency, and public record files, we discuss each of these in some detail below. Information about inquiries is discussed in a latter section.

Credit account information

Credit account records contain a wide range of details about each account. The data generally fall into five broad categories: account identification, account dates, account balances, account description, and payment performance. Each credit account record includes an account number, a unique identifier for each credit provider, and account ownership status (in particular, single or joint account or authorized user). Pertinent date information includes the date the account was established, the date it was closed or transferred (to collection or other major change in status), the date the account balance was paid down to zero, and the date when information was last reported to the credit reporting agency. The account records also provide current balance information, the largest amount ever owed on the account, the size of any credit limit applicable to the account, and any amount past due.

Credit account records include a variety of account descriptive information, including identification of the type of account; for example, a closed-end loan (mortgage or installment) or open-end loan (revolving, nonrevolving, or check credit), and the nature or purpose of the account; for example, credit card, charge account, automobile loan, or student loan. Finally, the credit account record provides information on the extent of current and historic payment delinquencies typically extending back 48 months as well as information on other account derogatories. Payment delinquency information is recorded in four classes of increasing severity—30 to 59 days, 60 to 89 days, 90 to 119 days, and 120 or more days past due. Other derogatories refer to accounts that have been charged-off or are in collection, or those associated with a judgment, bankruptcy, foreclosure, or repossession. Here, accounts that are 120 or more days past due and accounts with other derogatories are grouped together and termed "major derogatories" or "seriously delinquent." Accounts with less severe delinquencies are termed "minor delinquencies."

Public records and collection agency actions

In addition to personal characteristics and credit account information, credit reporting agency data include information derived from monetary-related public records and reports from collection agencies. Credit evaluators typically consider public records and collection agency actions to be adverse information on a par with credit account major derogatories when assessing the credit quality of individuals. The prevalence of these items is significant. Over one-half of

the individuals in the sample with at least one major derogatory (historic or current) did not have any credit account major derogatories; the only major derogatory items they had were collection agency actions or adverse public records.

Public record information includes records of bankruptcy filings, liens, judgments, and some foreclosures and lawsuits. The data distinguish between federal, state, and local tax liens and other liens. Otherwise, unlike credit account data, the public record data do not provide a classification code for the type of creditor or plaintiff. Although public records include some details about the action, such as the date the action was filed, the information available is much narrower in scope than that available on credit accounts. Overall, about 12 percent of the individuals in the sample had at least one public record item, and almost 37 percent of the individuals with a public record item had more than one item noted.

Credit reporting agency records also include information on non-credit-related bills in collection that are reported by collection agencies. In some cases, collections on credit-related accounts also are reported by collection agencies rather than by the original creditor. In this case, the information is grouped with the collection actions on non-credit-related bills rather than with the credit account information. Overall, about 30 percent of the individuals in the sample had at least one collection action reported by a collection agency. The most common types of collection actions reported involved unpaid bills for medical or utility services.

Collection agency records include only limited details about the action, including the date acquired by the collection agency, the original collection balance, and an indicator of whether the collection has been paid in full. There is no code indicating the type of original creditor.

Issues of timelines, completeness and duplication

Issues of timeliness, completeness and duplication arise in each of the four main components of the credit records data: the credit records, non-monetary public records, collections agency files and inquiry records. Issues can also arise in the personal identifying information, but because such data were not included in the sample the Federal Reserve acquired, these aspects of the data are not investigated here. It should be noted, however, that issues related to the completeness, accuracy, or consistency of the reporting of individual personal identifying information all can affect the credit record of an individual. If an individual uses different names or addresses on applications, or the social security number is either

erroneous or not provided, then fragments of information provided by different reporters may not all be properly linked to an individual. The result can be an incomplete credit record or, in some cases, the assignment of credit information to the wrong individual.

The discussion that follows focuses on each main component of the credit record in turn, beginning with the largest segment of the data, the records on credit accounts, and identifies issues that bear on the quality of the data and ultimately their use in assessing the credit history of an individual.

Account Status

Assessing the status of each credit account with respect to whether the credit relationship is ongoing (an "open account") or whether the account is closed and cannot be added to by the consumer is central to credit evaluation. Determining whether an account is open or closed is not always straightforward, partly because some creditors do not report all account closures to the credit reporting agencies. Instead, in many situations, creditors simply stop reporting any information about an account, creating uncertainty about the current status of the account. These "not currently reported" accounts constitute a significant portion of all accounts in the credit reporting agency's data.

For the discussion that follows, credit accounts are grouped according to their indicated status and whether they are currently reported. We define an account as being "currently reported" if either (1) its status had been reported to the credit reporting agencies within two months of the date the sample of credit records was drawn or (2) it was last reported (at any time) to be closed *and* had a zero balance at the date of last report. We treat all installment and mortgage accounts paid down to a zero balance as closed. With these definitions, we create four mutually exclusive groups:

- "Open" credit accounts are currently reported and not reported as closed. These include accounts that a consumer can use in the future to incur additional debt, such as an open-end revolving account, and accounts that are in the process of being paid down on a scheduled basis, such as a closed-end mortgage or installment loan.
- "Closed" credit accounts are currently reported and reported as closed. Closed accounts cannot be used to incur additional debt. Virtually all of these accounts have been fully

repaid and have a zero balance, although a positive balance remains on a small number of closed revolving accounts.

- A large number of accounts are not currently reported. For the analysis, these accounts are divided into two categories:
  - "Dormant" accounts are non-installment, non-mortgage accounts that were last reported as open with no outstanding balance but for which the time of last reporting was more than two months from the date the sample was drawn. These accounts appear to be inactive, but it is not possible to determine whether they are open or closed from the data.
  - All other accounts that are not currently reported are placed in the "unknown" accounts category. All of these accounts had a positive or unknown balance at their last date of reporting. These include installment, mortgage, and, to a small extent, revolving accounts that may have been paid off but lack a final record of disposition. It also includes accounts that were sold or transferred to another creditor or collection department or agency but not reported as closed by the selling or transferring institutions. Finally, it includes accounts on which reporting ceased following entry into collection or due to another state of serious delinquency.

Data quality issues may arise for any account, regardless of its status. Even for open accounts that are routinely updated, some information about the account may not be reported, such as a credit limit on a revolving account. Leaving that issue aside, currently reported (either open or closed) credit accounts are likely to have the fewest issues related to the timeliness of information shown in a credit record.

The current account status was known ("currently reported") for about 74 percent of the accounts in the credit reporting agency sample (table 2, derived from second memo item).[16] The

[16] The data used for this study represent the complete credit records of a nationally representative sample of individuals. However, raw account distributions in such data are not proper estimates of the distribution of characteristics of a representative sample of credit accounts. This disparity occurs because many accounts, including joint accounts or accounts with co-signers, are contained in the credit records of multiple individuals. An adjustment for such multiple reporting has to be made before statistics that are representative of all credit accounts can be computed.

The weight of jointly held accounts is inflated in the full credit reporting agency database because such accounts appear at least twice in the records. Like single-owner accounts, joint accounts should be counted only once. To do so requires weighting them to reflect the number of different individuals for which an account is

majority of these credit accounts (57 percent) were closed (derived from first memo item); the remainder were open. Because these accounts were currently being reported on at the time the sample was drawn, users of the data did not have to make assumptions about their current status.

The current status of the remaining credit accounts was not known ("not currently reported") and, thus, users of the data would have to make some assumption about their current status in order to use the data for credit evaluation or other purposes. Among the accounts that were not currently reported, 70 percent were dormant accounts (derived from memo item). For these accounts, the primary issue was whether the account could be used by a consumer. In other words, it is unclear from the credit records whether a consumer could begin to use the account or whether the creditor had terminated the account relationship without notifying the credit reporting agency.

The remaining accounts that were not currently reported (grouped in the "unknown" category), which comprised about 8 percent of all the credit accounts in the sample, can present a particularly vexing problem for users of the data because this category includes accounts that had a positive (or unknown) balance at the date of last report. This category includes accounts in good standing that may have been sold, transferred or paid off but not reported as such, as well as derogatory accounts that were transferred to a collection department or sold for which no subsequent report was obtained.

Accounts that are not currently reported are likely to be flagged by consumers as errors since the information is out of date and the balance shown in the record is likely to have changed because some payments were made, the account in question was transferred to someone else who is reporting the subsequent account, or a refinancing took place and the old account was never reported as closed. Like other not currently reported accounts, some dormant accounts may in

---

reported. The number of holders is not fully reported in the credit records, so approximate weights were used to reasonably represent the distribution of all U.S. credit accounts reported to credit reporting agencies. Our statistics on various measures of account activity and usage are calculated with these weights and, thus, they are estimates rather than actual quantities.

Single accounts were assigned a weight of 1 and joint accounts a weight of one-half. Co-signed and authorized user accounts were given weights of zero under the assumption that they were represented elsewhere by the main account user. A complicating factor was that for some accounts, "holder status" was unknown. Weights were assigned to these accounts on the basis of relative numbers of single and joint accounts in the data for accounts of a similar type.

It should be noted that a very small proportion of accounts (0.03 percent) indicates some disputed information due to a notice of dispute by the consumer under the FCRA. Disputes may arise over virtually any aspect of an account relationship including the exact amount owed, the timing of payments, or whether a particular purchase of a good or service was even made by the account holder. Since the true status of these accounts is under dispute, and because their proportion is small, we exclude them from the tables in this article.

fact be closed, but information on the status is missing, leaving it up to the judgment of the user of the data to determine how to evaluate such accounts. Dormant accounts seem less likely to be flagged as errors by consumers since the consumer may not have used the account since it was last reported. Like open revolving accounts, however, records on dormant accounts may be missing some pertinent information, particularly credit limits.

How much influence an account that is not currently being reported has on an individual's credit score will vary depending upon the specific details of the account information, for example, the size of any outstanding balances, but also importantly on any rules the credit score modeler has instituted to address accounts with such stale information. For example, a modeler may ignore accounts showing positive balances that were current in their payments when last reported but that have not been updated in the past 12 or more months.

In sum, credit reporting agency data provide a wide-ranging and comprehensive picture of accounts, reflecting individuals' experiences with credit. However, in some instances the data are not sufficiently up to date or complete to permit a clear understanding of an account's current status.

The following section, takes a more detailed look at the information in the credit reporting agency files focusing on issues of data limitations concerning those items most pertinent to credit evaluation: level of a consumer's indebtedness, payment history, and credit account characteristics.[17]

Level of Consumer Indebtedness

When conducting credit evaluations, creditors consider the type and amount of credit a consumer owes, as well as the proportion of available credit they have in use (their rate of credit utilization). For revolving accounts, credit utilization is measured as the proportion of available credit in use (outstanding balance divided by credit limit). For mortgage and installment accounts, credit utilization is generally measured by calculating the proportion of the original loan amount that is unpaid, referred to here as the paydown rate.

Fundamental to measuring consumer indebtedness is deciding which accounts to treat as "active"–that is, installment and mortgage accounts in the process of being repaid and revolving

[17] For a more detailed discussion of factors considered in credit evaluation, including the relative weights given to different factors, see the description provided by Fair Isaac and Company at their web site www.myfico.com. Also see, Avery et. al. "Credit Risk, Credit Scoring and the Performance of Home Mortgages."

accounts upon which consumers can draw. Clearly, credit evaluators would include currently reported open accounts as "active" in any calculations. The difficulty, however, is in determining how to treat accounts that are not currently reported; those in our dormant and unknown categories. The dormant account category likely includes many accounts that are not currently reported that can be further drawn upon by the consumer.[18]

Outstanding balances. With regard to the credit record data on outstanding balances, non-zero balances are virtually always shown for open mortgage and installment accounts as one would expect. Zero balances on such accounts likely indicate error. By contrast, many revolving accounts showed no outstanding balance even though they were by far the largest proportion of active accounts measured by number. The large share of revolving accounts that showed a zero balance at last report is not surprising. The use of credit cards varies greatly; some go unused for a period of time, while others are used regularly either as a convenient means of payment or a source of credit. From the credit records, it is not possible to draw conclusions about whether the particular balance level shown for either an active open-end or closed-end account is accurate.

Missing credit limits. Credit evaluators routinely consider credit utilization rates in assessing creditworthiness. To calculate a utilization rate for a revolving account, one must have information on both an account's outstanding balance and its credit limit. Unfortunately, a credit limit is not reported for many revolving accounts despite the fact that limits likely exist for the vast majority of these accounts. Approximately one-third of all active revolving accounts in the sample lacked credit limit information (table 2).[19] For these accounts, credit evaluators must use

[18] For the analysis of consumer indebtedness, we include currently reported open accounts as well as dormant revolving accounts that were last reported within a year from the date the sample was drawn in our definition of active. While we used one year as a cutoff, discussions with industry professionals indicate that there is no hard and fast rule regarding a single appropriate choice. The choice of the cutoff affects the number of accounts deemed to be active and the potential borrowing capacity of an individual but has no bearing on the amounts owed because all of these (dormant) accounts had zero balances at time of last report. We include no accounts from our unknown category, which for reasons discussed below, we have reason to believe are most likely closed.

[19] The incidence of missing credit limits is significantly lower in current credit reporting company data. Industry estimates indicate that credit limits are currently missing on about 13 percent of revolving accounts. The higher

other techniques to estimate a utilization rate. The most common approach in these circumstances is to use the highest balance ever reported on the account as a surrogate for the credit limit (referred to here as the "high balance") and calculate a utilization rate using this measure. This alternative approach creates very different profiles regarding the extent on which revolving accounts have been drawn. For mortgages and installment loans, the credit limit and the high balance (the original amount borrowed) are one and the same, and so the profiles will be identical.

The profile of credit limits using this approach differs notably between revolving accounts that had credit limits reported and those that only had a high balance reported. A much larger percentage of accounts that could only use the high balance amount as the credit limit had limits below $1,000 (compare the revolving account categories in tables 3a and 3b). Thus, use of the high balance measure for credit limits on revolving accounts for which limits are not reported likely understate the actual credit limits available on those accounts.

Utilization rates for revolving accounts

Combining information on outstanding balances and credit limits (or the high balance for revolving accounts if the credit limit was not reported) allows users of the data to calculate account utilization rates. Notable differences exist between accounts with credit limits reported and those with only a high balance reported (table 4, compare top and bottom panels). These differences stem from (1) the use of a different measure of credit limit, and (2) correlations between the propensity of a creditor to report a credit limit and the account characteristics. For example, observed differences in the share of accounts that had utilization rates of zero can only be caused by differences in the propensity to report credit limits.[20] However, differences in the proportion of active revolving accounts calculated to have either relatively low utilization rates (from 1 percent to 24 percent) or very high rates (95 percent or more) can be strongly influenced

incidence of missing limits in our sample may stem from a period of time when a few large creditors decided to suspend reporting of this item for competitive reasons. Pressure from financial institution regulators and the credit reporting companies appears to have convinced these creditors to resume reporting credit limits. See, Robert M. Hunt, "The Development and Regulation of Consumer Credit Reporting in America," Federal Reserve Bank of Philadelphia, Working Paper no. 02-21, November 2002.

[20] For the construction of tables 3a, 3b, and 4, we assumed dormant accounts last reported within one year of when the sample was drawn were still open to the consumer and could be used to borrow. We also reviewed how the patterns shown in these tables changed if a two-year rule was used instead. As might be expected, the main effect is to increase the proportion of revolving accounts showing a zero utilization, however, the effect is small--increasing the percentage by only a couple of percentage points.

by which measure of credit limit is used. Here substantial differences can be observed. For revolving accounts with reported credit limits, one-fifth had a utilization rate in the low range, while only 5 percent of accounts with only a high balance reported fell in this range. At the other extreme, only 6 percent of active revolving accounts with reported credit limits had a utilization rate of 95 percent or more, while 31 percent of revolving accounts that only had a high balance reported had utilization rates this high.

The discussion above highlights the different utilization profiles of accounts with and without credit limits reported. To detect systematic patterns in the reporting of credit limits we examined the relationship between the creditor (using an identification number within the data since the name of the creditor was not provided) and the likelihood that a credit limit was missing. Results suggest that most of the variation in the reporting of credit limits for active revolving accounts can be explained by which creditor reported the information. Restricting ourselves only to creditors that reported a large number of accounts, we split these creditors into three groups: those that reported credit limits for fewer than 5 percent of their customers; those that reported credit limits for more than 95 percent accounts; and all others.[21] Although only 12 percent of the creditors in the analysis were in the first group, they accounted for almost three quarters of the total accounts with missing credit limits and less than 0.03 percent of those with limits reported (data not shown in tables). At the other extreme, the second group, representing 68 percent of the creditors and 86 percent of the accounts for which limits were reported, accounted for less than 1 percent of the accounts with missing limits.

The group in the middle, representing 20 percent of the creditors, is also revealing. Here, creditors reported limits for some active revolving accounts but not for others. Concerns have been raised that some creditors report limits selectively and, in particular, they do not report limits for some subprime customers because they do not want these customers to be targeted for solicitation by other creditors. We find only mild support for this view. Overall, 51 percent of the active revolving accounts of subprime customers held by creditors in this middle group had their credit limit reported versus 53 percent of accounts of their prime customers.[22] However, for a subset of creditors in this middle group–about 5 percent of all creditors--all specializing (more

[21] For this analysis we used a threshold of seventy-five active revolving accounts reported in our sample to define a "large" creditor. Six hundred and seventy-four creditors met these criteria. These creditors accounted for 96 percent of all missing credit limits in the credit reporting agency files.

[22] We used a credit reporting company credit score supplied with the credit files to make a rough determination of prime and subprime borrowers.

than 50 percent of their accounts) in subprime lending, some degree of selective reporting was apparent. For prime customers of these creditors, credit limits were reported about 85 percent of the time versus 40 percent for subprime customers at these institutions.

Credit Payment History

The most important factor considered in credit evaluation is a consumer's history of repaying loans and any evidence of monetary-related public actions or non-credit-related collections. Such an analysis considers not only the frequency of any repayment problems, but also their severity (how late), recency, and dollar magnitude. In general, an individual with a record of a major derogatory will have a more difficult time qualifying for new credit, may face higher prices for credit received, or may have actions taken to limit his or her ability to further borrow on existing open accounts. Data ambiguities or other limitations regarding payment history are of particular concern because of the potential importance of these items in credit evaluation. In assessing the credit circumstances of an individual, credit evaluators ordinarily consider the current status of accounts, the consumer's recent payment experience on credit accounts, and his or her record of payments over a much longer time period.[23]

Payment status at last report

For currently reported accounts or accounts that are closed or dormant, the account status at the date of last reporting will be the "correct" current status in virtually all cases.[24] One exception to this occurs because of inconsistencies in the way creditors report account delinquencies. About 11 percent of active accounts were reported by creditors that did not report minor delinquencies for any accounts. An additional 12 percent were reported by creditors that did not report any 30-59-day delinquencies. Nonrevolving accounts were particularly likely to fall in these categories. There is no evidence that these creditors update information supplied to the credit reporting companies less frequently than other creditors. Instead, these creditors appear to be reporting accounts as non-delinquent until they reach a seriously delinquent status. Consequently, customers of these creditors will tend to show a lower incidence of minor

[23] The FCRA specifies that consumer credit reports can not include any adverse item of information that is more than seven years old unless it involves a bankruptcy (10 year limit), criminal conviction (no time limit), or one of a few other narrow exceptions**.**

[24] It is possible of course that the account status has changed since the credit information was last updated by the reporting entity.

delinquencies than customers of other creditors. For accounts in the "unknown" category, a much more serious question arises as to whether the account status at the date of last reporting is the same as the account's correct current status. For these accounts, the creditor has not updated the account information in at least three months (and often much longer), and they all show positive balances, raising the likelihood that the account status has changed since it was last reported. Evidence indicates that derogatory and non-derogatory accounts in the unknown category differ in their likelihood of a changed status; thus, we discuss them separately.

Unknown category accounts not in major derogatory status. The current status of non-derogatory accounts in the unknown category is likely to be different from that last reported in most circumstances. Since these accounts showed positive balances at the date of last reporting, one can reasonably infer that their status had changed by the time the sample was drawn: Either the account was closed or transferred or the account holder made payments (and thus changed his or her balance) or did not make payments, in which case the performance status worsened. The most notable exception is for records of some types of student loans where repayment may be deferred for some period of time. About 67 percent of all accounts in the unknown category were not in major derogatory status at the date of last reporting (from table 2). About two-thirds of these accounts were revolving or open nonrevolving accounts (not shown in tables). Most of these accounts require monthly payments, thus it seems highly unlikely that their status at last report reflects their current circumstances. Recognizing the high likelihood that many non-currently reported accounts have had a change in status, the credit reporting companies have adopted "stale account" rules. The credit reporting agency's rule in place at the time our sample was drawn was to define all revolving and nonrevolving accounts with positive balances and no major derogatories as stale if they had not been reported within six months. Stale accounts were treated as closed and assigned a zero balance. Our data reflect this rule. Sixty-one percent of the revolving and nonrevolving accounts in the unknown category had been reported within six months of the date the sample was drawn (and over 80 percent within a year). These accounts are likely candidates for the stale account rule, and the likelihood that they have been closed or transferred is significant. The remaining accounts, constituting about 3 percent of all nonclosed revolving and nonrevolving accounts, are ones that were exceptions to the stale account rule. The status of these accounts is less clear.

Stale account rules were not used for mortgage and installment accounts by the credit reporting agency that supplied our data.[25] As a consequence, a significantly higher percentage of these accounts is in the unknown category than is in revolving and nonrevolving accounts. Almost one-third (32.5 percent) of all non-major-derogatory mortgages last reported with a positive balance were in the unknown category. Only 33 percent of these had been reported within six months of the date the sample was drawn. One can infer that many, if not most of these accounts, had been closed or transferred. Specifically, for over one-half the mortgages in the unknown category, consumers had a new mortgage for approximately the same amount reported as opened within two months of the date of last reporting of the mortgage in the unknown category–a strong indicator that the unknown category mortgage was refinanced or that the servicing was sold.

Installment loans show a similar but less severe pattern. About one-fifth of the nonclosed non-major-derogatory installment accounts are in the unknown category; 33 percent of these were last reported within six months of the date the sample was drawn. One might cautiously infer that many of the loans may not have been outstanding at the time the sample was drawn. About 48 percent of non-major-derogatory installment accounts in the unknown category have one of two conditions–they are beyond the original due date at the time the sample was drawn or the gap between the date the sample was drawn and the last date they were verified is larger than any previous gap in their payment history.

Another inference that can be drawn is that many of the mortgage and installment non-major-derogatory accounts in the unknown category may no longer have been outstanding at the time the sample was drawn. More than one-half of loans in the unknown category for each account type were reported by creditors that had not reported any accounts in our sample within three months of the time the sample was drawn.[26] If these creditors no longer report to the credit reporting companies, these accounts could only have been updated as a result of actions by the consumer or some action by the credit reporting agency such as applying a stale account rule.

The consequence of accounts that have not been properly reported as closed or transferred to another entity will, in most cases, be that consumers will show higher aggregate account balances. The issue goes beyond the actual balances owed and includes uncertainty

[25] Discussions with the credit reporting company that supplied our data indicate that it is in the process of implementing stale account procedures for these types of accounts.
[26] Creditors had to have reported at least ten sample accounts to be included in this calculation.

about the extent of any payment problems as well. As shown in table 2, about 36 percent of all accounts that were last reported as minor delinquencies were in the unknown category. For 80 percent of the installment accounts and about two-thirds of the other accounts in the unknown category with minor delinquencies shown at the date of last report, the account had not been reported within six months of the date the sample was drawn. Thus, the status of these accounts had likely changed; yet because the information remains unchanged in the files, these accounts can have a disproportionate effect on the assessment of current minor delinquency.

Unknown category accounts last reported in major derogatory status. Unlike minor delinquent or nonderogatory accounts, the status of a major derogatory account can remain unchanged for a long period of time. Consumers may have stopped paying, and creditors may have stopped trying to collect on the accounts. Thus, an account's status could in fact remain the same and not require updating. The failure to update is reflected in the sample data. Fifty-nine percent of the accounts last reported as unpaid (positive balance) major derogatories were in the unknown category (table 2). Of these, more than one-quarter had not been updated for over four years.

Limited evidence implies that some of these accounts were likely paid off but that such information was not reported to the credit reporting agency. Specifically, another mortgage was reported as originated *after* the date the account was last reported for about 10 percent of the unknown category mortgages with major derogatories. Generally, creditors require that all major derogatories be paid off before a new mortgage is originated. Similarly, a mortgage was reported as originated after the date of last report for about 3 percent of other unknown category accounts with major derogatories.

Further evidence indicates that even if some of these major derogatories had been paid off, such a change in status might not have been reported. About 32 percent of the major derogatory accounts in the unknown category were reported by creditors that had not reported on any accounts within three months of the date the sample was drawn. If these creditors are no longer active reporters, then even paid off accounts are unlikely to be recorded as such. Here, the account may still have existed, but it may have been transferred or sold, and thus reported

twice. In these circumstances, if the consumer were to pay off the account, then only one of these duplicate records may be updated as paid.[27]

Further, almost 12 percent of the major derogatory accounts in the unknown category were reported by creditors that, in our sample, only reported derogatory accounts. Such reporting patterns are particularly true of nonrevolving accounts, for which the figure is about 35 percent. These creditors may simply not report when accounts are paid off. The fact that some creditors only report major derogatory accounts has another implication for the completeness of credit files. Satisfactorily performing accounts of these creditors are not included in the files and the extent of these nonreported accounts is not known. The failure to report accounts in good standing may affect the credit evaluation of consumers with such accounts. For example, if consumers have low utilization of these nonreported accounts, the failure to report may reduce their credit scores. For those consumers having nonreported accounts with high utilization, though, the failure to report may actually improve credit scores.

Public Records, Collections, and Inquiries

In addition to credit account information, credit reporting agency records include information derived from a variety of public records, reports from collection agencies, and creditor inquiries about a consumer's credit history. Credit evaluators consider these types of information in judging the credit quality of individuals. The following discussion examines the information in the credit reporting agency files on each of these categories highlighting areas where issues of missing or ambiguous information may complicate use of the data.

Public records. Credit reporting agency records include information derived from public records available from government entities, including records of bankruptcy filings, liens, judgments, and some foreclosures and lawsuits. The data regarding bankruptcy distinguish between the various types of personal bankruptcies. The two main types of consumer bankruptcies are Chapter 7 and Chapter 13, each named after the chapter in the U.S. bankruptcy

[27] To test this conjecture, we compared the percent of all accounts that had ever been reported as major derogatories but that were last reported satisfactory (paid off or making payments) for two groups of creditors: (1) those that had not reported any accounts within three months of the date the survey was drawn and (2) those that had. In each case, we restricted ourselves to accounts that were opened in the same three-year interval (1995 to 1997). Creditors that were currently reporting accounts had a satisfactorily performing incidence rate about 50 percent higher than creditors that were not currently reporting.

code that defines the nature of the proceedings.[28] The data also distinguish (albeit imperfectly) between federal, state, and local tax liens and other liens. Otherwise, unlike credit account data, the public record data do not provide a classification code for the type of creditor or plaintiff (for example, a provider of medical services or a utility company). By examining names of plaintiffs, it is possible, though, to distinguish among broad types of judgments and lawsuits, such as those related to unpaid bills for medical and utility services (again, imperfectly). Although public records include some details about the action, the information available is narrower in scope than that available on credit accounts.

Overall, about 12 percent of the individuals in the credit reporting agency data had at least one public record item (derived from table 1), and almost 37 percent of the individuals with a public record item had more than one item noted (percent not shown in tables). Judgments and liens were the two most common types of public record noted in our sample of credit reporting agency data, representing 40 percent and 34 percent of the public records respectively. Bankruptcies accounted for nearly all the remaining public records. Foreclosures and lawsuits accounted for small proportions of the public record actions included in the data. These small proportions are a consequence of the fact that credit reporting companies gather such information only in limited circumstances. For lawsuits, the reason underlying this decision is a belief that the simple filing of a lawsuit, which precedes any decision on its merits, is of only very limited value, particularly for credit evaluation. Moreover, as shown below, there is a large inconsistency in the likelihood that a lawsuit will be reported to the credit reporting companies. For foreclosures, the reason credit reporting companies generally do not gather such information is their belief that most foreclosures are already reported in conjunction with credit accounts. Thus, collecting foreclosures from public records would be redundant.

In some cases, more than one public record item for an individual appears to be associated with a single episode. Several public record items may result from a single episode for a variety of reasons. Failure to pay a bill may result in both a lawsuit and a judgment appearing on an individual's records; several public records related to unpaid medical bills may stem from the same injury or illness; or an appealed or a refiled judgment in a different court may result in more than one record of a judgment. In addition, the records for an individual may

[28] Other bankruptcy chapters available to individuals include Chapter 11 and Chapter 12 bankruptcies but these are rarely used. For more information on bankruptcy, see "Bankruptcy Basics," Administrative Office of the United States Courts, June 2000.

show a state or local tax lien that has not been paid and a separate record of a paid tax lien of the same type, but these may or may not refer to the same original lien.

To the extent that case identifiers (docket numbers) are available, credit reporting companies use these identifiers to update public record information. For example, if a tax lien is reported paid with the same docket number used for the original public record of the lien, the original record will be updated showing the status as paid, rather than as a new lien item added to the consumer's record. However, consistent case identifiers are not always available; for example, new docket numbers may be assigned when a judgment is appealed.

In such circumstances, two or more distinct records for the same episode may appear in the data. In general, it is difficult to determine with certainty whether distinct public record items pertain to the same episode. The inability to always distinguish distinct public records from simple updates of information about a public record already on file can led to an erroneous count of the number of public record items in an individual's credit record.

Patterns in the public records from our sample of credit reporting agency data suggest a degree of inconsistency in reporting across plaintiffs and geographic areas. For example, as discussed above, lawsuits are inconsistently captured in the credit reporting agency public records. This inconsistency is reflected in our sample as only three states (Maryland, New York, and Pennsylvania) accounted for two-thirds of all individuals with records of lawsuits. Inconsistencies also can arise because of the practices of specific plaintiffs. For example, some plaintiffs obtain separate judgments for individual unpaid billed items, whereas other plaintiffs in similar circumstances might have combined the bills.

Collection agency accounts. Information on non-credit-related bills in collection is reported to credit reporting companies by collection agencies. In some cases, collections on credit-related accounts also are reported directly by collection agencies and consequently are grouped with the collection actions on non-credit-related bills rather than with the credit account information. Overall, about 30 percent of the individuals with credit reporting agency records had at least one such collection action reported by a collection agency (derived from table 1). For about 11 percent of the individuals, the only record item in their credit reporting agency file was a collection agency action. Because collections are considered to be a type of major

derogatory, they can have an important effect on the ability of a consumer to obtain credit or the price paid for such credit.

Unlike the credit accounts, but like the public records, collection actions are reported without a code indicating the type of original creditor. However, the data do include information that can be used to infer the type of entity that originally sought the collection. By our estimates, most collection actions reported by collection agencies do not involve credit accounts. Only about 6 percent of collections reported by collection agencies are credit account related. Among the non-credit-related collection actions, the majority (about 55 percent) are associated with medical bills. The high incidence of collections related to medical bills is not surprising given the large number of individual consumers and families that do not carry health insurance coverage and the high cost of many medical services.[29] The second largest category of non-credited-related bills involved in collection actions are unpaid bills for utility services, which by our analysis, account for about 24 percent of the non-credit-related collections.

Most collection actions reported by collection agencies showed small balances owed when originally reported to the credit reporting agency. About 34 percent of all the collections involved an original amount owed of less than $100, and 82 percent involved an amount less than $500. The data also indicate that only about 11 percent of the reported collection items have been fully paid off, with collections filed by a governmental entity the most likely to have been paid off and those that were credit related the least likely to have been reported as fully paid.

As with the public records, individuals sometimes have more than one collection agency action reported. About 44 percent of the individuals with a collection agency record had more than one item noted (data not shown in tables). Like the public records, tracking collection agency accounts for the purpose of updating their status is not always possible, because of changes in account numbers, sometimes resulting from transfers of the account across collection agencies.

[29] According to the Federal Reserve's 2001 Survey of Consumer Finances, about 9 percent of households had no public or private health insurance coverage and nearly 17 percent had only partial coverage, meaning that one or more members of the household had no coverage. These proportions are little changed from those found in the 1998 Survey of Consumer Finances.

Inquiries. Credit reporting agency records include information about inquiries made about a consumer's credit history. Such inquiries are conducted to ensure that an applicant for credit, apartment rental, insurance, or employment has a background that meets the minimum standard the inquirer has established for providing the service. The data do not include inquiries made by creditors about existing accounts, those made in conjunction with screening for pre-approved credit, or inquiries made by consumers themselves to review their own credit records. The data include the date of the inquiry, the type of credit being considered (missing for most inquiries), and the subscriber code of the inquirer. In our sample of credit reporting agency records, inquiries were retained in the records for two years. This practice is consistent with the view that credit underwriters focus primarily on a consumer's recent efforts to obtain credit.[30]

Overall, about 58 percent of the individuals in the credit reporting agency sample had at least one inquiry noted in their files (derived from table 1). Inquiries are often bunched in time. About 26 percent of the inquiries were made within one week of another inquiry that appears in a given individual's credit file, and about 60 percent were made within one month of another inquiry in the file (data not shown in tables). These figures are consistent with the view that consumers often approach several creditors when seeking a new loan. For example, a consumer purchasing a car or home may approach more than one creditor while shopping for the best available terms to finance the purchase. However, because only a very small proportion (fewer than 2 percent) of the records of inquiries included information about the purpose of the inquiry, it is impossible to determine with certainty if bunched inquiries represent shopping for a single loan purpose or requests for different loan products (for example, a mortgage and a credit account to purchase household items). Nevertheless, credit evaluators use different techniques to differentiate these two types of inquiry circumstances. For example, the type of creditor (for example, banking institution) may be used as a proxy for loan type along with the timing of the inquiry to identify multiple inquiries arising from shopping for a single loan.

Empirical Strategy

In this section, we describe the empirical steps taken to evaluate the potential effects of credit record data accuracy and completeness on access to credit. We first describe the general

[30] For example, Fair Issac and Company only considers inquiries recorded over the past twelve months when calculating a FICO score for an individual. See, www.myFICO.com.

procedure followed to develop the credit scoring model used in the analysis. We then present some of details of this work. Finally, we describe the population category that is the focus of the analysis. The section that follows provides the results of our empirical investigation.

The general procedure

To investigate the potential importance of data accuracy and completeness on access to credit, it is necessary to know the details of the credit scoring model (both the underlying factors and associated risk weights given to these factors) used to derive an individual's credit history score. In addition, one must have information about how the underlying factors were created from the original credit records, referred to here as credit record "roll-ups." Unfortunately, we did not have access to such proprietary information. We did, however, have access to three types of information that were used to *estimate* a credit history scoring model. This information included: (1) a proprietary credit risk score assigned to each individual in our sample; (2) a large set of credit record roll-ups for each individual, a subset of which were the input elements of the proprietary credit scoring model; and (3) detailed account level information in each individual's credit record, which we believe contains all information needed to construct the roll-ups.[31] The first two items were used to approximate the proprietary credit history model by calculating risk weights for various roll-ups using a "reverse engineering" procedure. The second and third items were used to "reverse engineer" the credit record roll-ups included in our version of the credit scoring model to determine how these were likely calculated.

Because there are a large number of roll-ups, and many potential specifications and combinations of these variables, our final estimated model undoubtedly differs from the actual model used to generate the credit risk scores included with the sample data. However, we were able to reverse engineer most of the credit record roll-ups with a high degree of accuracy and fit the proprietary model quite well with our approximation. Thus, we feel reasonably comfortable using our approximation of the scoring process to estimate the potential effects of a change in a credit record item on an individual's credit score. Moreover, different model builders consider different credit risk factors in building their scoring models and employ different rules for constructing credit record roll-ups and, consequently, even with a perfect replication of the

[31] As an example, a credit record "roll-up" might be an individual's total amount of mortgage debt outstanding or overall credit utilization for revolving accounts.

process we could not necessarily generalize from our results with any more specificity than offered here.

Approximating the credit scoring model

We used a slightly adjusted version of the proprietary credit score included with our data as the basis for our analysis. Consistent with general industry practice we transformed the "raw" score to a score with a range of 300 to 850 with a higher score indicating a higher degree of credit worthiness. The median adjusted score was 739 with 74 percent of sample individuals above 660 and 14 percent below 600 (the score ranges used to define our primary analysis sample defined below).

To approximate the proprietary credit scoring model, we used standard regression procedures to predict the adjusted credit score assigned to an individual in the sample data using a large set of credit variable roll-ups. We restricted the sample for this analysis to individuals with credit scores and with at least one credit account in their credit records.[32] A total of 75 variables were used to construct the regression equation, including numbers of credit accounts of different types and various characterizations of payment history patterns including the number of accounts where all payments have been made on time or in various stages of delinquency or in major derogatory status. Also included were measures of outstanding balances, credit limits on credit accounts, age of credit accounts, account inquiry information, and variables derived from public records and collection agency accounts.

The regression equation was estimated separately for two subpopulations: those with one or more major derogatory credit accounts in their credit records and those with none. The estimation was conducted in this way because allowing the estimated coefficients to differ across these two groups provides a noticeably better fit within the subpopulation that is ultimately the focus of the study, that is, consumers with credit scores that place them in the marginal range of creditworthiness, as defined below. It is also consistent with the common industry practice of using different "score cards" for different subpopulations. The R^2 for each of the two regressions was about 0.90, with a combined R^2 for the full population of 0.95. Proprietary considerations constrain our ability to report details of the regression equation specification or coefficient

[32] Individuals with no record of a credit account in their files generally were not assigned a credit score by the credit reporting agency. Individuals with a record of bankruptcy were also excluded from the sample.

estimates. It should be noted that a few variables included in our final credit scoring model were not statistically significant sometimes without the expected sign. As a consequence, our simulations of the effects of changes in an individual's credit record, did, in a few instances, lead to an anomalous outcome in the sense that some credit scores moved in an unexpected direction.

Reverse engineering the roll-ups

We were able to reverse engineer the roll-ups used in our estimated model with a high degree of accuracy. Our reverse-engineered versions of the 75 variables used in our model fit the true values for all variables for 98.3 percent of the individuals in our sample. Those individuals that we could not fit exactly were dropped from the analysis.

The population of interest

To assess the effects of credit record data integrity issues on credit availability, our analysis focuses on the subgroup of the population that is mostly likely to be materially affected by such data quality problems. Although data integrity issues can influence the credit score of any individual, those near the margin of access to credit are the ones mostly likely to be materially affected by such problems. For this analysis, we focus on individuals whose credit score falls in the range 600-660. Generally, this is the credit score range that subdivides individuals between the subprime and prime markets for mortgage and credit. We refer to individuals in this credit score range as the "at risk" population. Other credit score ranges may be appropriate for other forms of credit, but we focus on mortgage credit because it involves the largest loan amounts. Approximately 10 percent of the individuals in our working sample (the sample after applying restrictions) had credit scores in the 600-660 range.[33] This resulted in a sample of 20,349 "at risk" individuals.

Evaluating the Effects of Data Limitations on Access to Credit

In this section, we empirically evaluate the potential importance of data accuracy and completeness on access to credit, focusing on individuals whose credit score leaves them most

[33] The working sample was limited to those individuals who had a reported credit score and a record of at least one credit account. The working sample excluded individuals with a record of a bankruptcy since we believe that the credit scores of such individuals may be derived using a credit scoring model that differs from the one used for the remainder of the population.

vulnerable to such concerns. The approach is three-fold. First, we identify specific data issues that can be addressed empirically. Second, we modify the credit records of the "at risk" individuals to simulate a change in their credit records consistent with a "correction" of a particular data issue of concern. Third, we calculate the effect of this simulated change in their credit records on their credit scores.

We present results for each data issue in several ways. First, we show the proportion of the "at risk" population that is *not* effected by the simulated change in their credit records. (For example, simulating the effect of closing a long dormant revolving account will not affect individuals that have no such accounts.) Second, we show the proportion of individuals that could be affected, but, that nonetheless, had no material change in their estimated credit scores. Third, we calculate the proportions of the "at risk" population whose estimated credit score either decreases (increases) by less than (more than) 10 points. This calculation provides additional insight into the proportion of the "at risk" population that may or may not face a material change in credit terms (either a higher or lower interest rate) or who might be unable to gain access to credit. Finally, we present the mean change in credit scores for "at risk" individuals that were materially affected in order to provide one basis for determining how much variation in credit scores may occur. Because a change a credit record may increase or decrease credit scores, we show separately the mean change for those with a decrease in scores and the mean change for those with an increase in scores. Results of this analysis are shown in table 5.

Data issues that are tested

Issues of data accuracy and completeness may touch on any aspect of an individual's credit record. Here, we focus on aspects of the data that we can assess and that our research suggests could have a potentially significant effect on credit evaluation. For purposes of presentation, the analysis is organized by the four major components of the credit records: records regarding credit accounts, public records, collection agency accounts, and inquiries. Specifically, we evaluate the effects of: (1) the failure to report as closed an account that has been paid-off as a result of completing scheduled payments, a refinancing, or as a consequence of transferring the account to another creditor, (2) the failure of some subprime creditors to report accounts with continuous satisfactory performance, (3) the failure of the largest student loan creditor to report any information on any of their accounts, (4) the failure of some creditors

to report credit limits on revolving accounts, (5) the failure of some creditors to report minor delinquencies, (6) duplications in public records, (7) inconsistent reporting of lawsuits in public records, (8) inconsistent reporting of collection agency accounts under $100 of any type or medical collections under $500, and (9) the failure to consolidate multiple inquiries for the same loan.

Empirical results

This section reports the results of our analysis regarding each of the data quality issues listed above. Our discussion includes a brief description of the process used to simulate modifications to "correct" each data issue and the results of imposing the simulation on the "at risk" population. The results are shown in table 5.

The failure to report as closed an account that has been paid-off. As described above, credit records frequently include accounts which are recorded as still open with a positive balance but have not been updated for an extended period of time. The credit records give information, such as payment performance and outstanding balance as of the last date of report. It is very likely that the status of these accounts has changed but has not been reported to the credit agency. Recognizing this lack of updating, most credit scoring models use "stale account rules." These rules retain the historic information on payment performance, but implicitly make the assumption that the account is no longer outstanding. Typically, these rules are *not* applied to accounts in major derogatory status, since it is much less clear in this instance that the account's status has changed. Originally, our strategy in simulating "corrections" to the stale account problem was to choose different time periods for treating an account as stale. However, in reverse engineering the roll-ups for our analysis, it became apparent that at the time our sample was drawn a one-year stale account rule was imposed by the credit reporting agency in the roll-ups for most calculations (also as stated earlier, the credit reporting agency employed stale account rules in the "raw" data to "correct" many non-currently reported revolving accounts). Thus, testing for the effect of rules for a longer time period would result in no impact on credit scores. Consequently, we examined the effect of a shorter stale account rule. Analysis of the patterns of verification showed that the vast majority of open accounts were verified every month or two (table 2). Thus, we chose a three-month rule to simulate what the effect would be

of a more aggressive stale account rule. We divide our simulation into two components—the imposition of a tighter rule on the treatment of satisfactory accounts and those last reported as minor delinquencies. Results are reported separately for mortgages, installments, and revolving accounts.

The first point to be made is that a substantial majority of non-currently reported accounts were last reported more than one year prior to the date the sample was drawn and consequently, were already "corrected" by the agency's stale account rule. Nevertheless, imposing a tighter rule still affects about 20 percent of individuals in the "at risk" sample. Very few of the affected accounts were delinquent at last report, effecting less than 2 percent of the individuals. The biggest effect is for satisfactory revolving accounts, effecting 18 percent of the "at risk" population. Interestingly, imposition of the tighter stale account rule resulted in some scores going up and some going down—even for delinquent accounts. This reflects the complexity of how roll-ups are calculated. On average, imposition of a tighter stale account rule hurts individuals with satisfactory mortgages and helps everyone else. Individuals with any minor delinquency and with a satisfactory revolving account tend to benefit the most, with a significant percentage having their credit scores increase more than 10 points.

The failure of some subprime creditors to report accounts. This issue is different from the others in that we are asking a hypothetical question about accounts that we do not observe rather than a question involving a change in treatment for accounts that we do observe. Our task is thus more difficult, and does not allow us to say anything about incidence. We do believe, though, that our simulation sheds light on the potential impact on the credit scores of an "at risk" individual if this individual were to have a satisfactory account with a subprime creditor who did not report it. We simulated this potential effect by choosing a random satisfactory (never delinquent), relatively new (less than four years), mortgage, installment, and revolving account at a finance company for each "at risk" individual and rescoring the individual as if this account had not been reported. We restricted the simulation to accounts at finance companies under the assumption that most subprime lenders who would fail to report are of this type.

Our simulation suggests that the failure to report a satisfactory subprime account is likely to have only a small impact on the credit scores of most "at risk" individuals. Only 4 percent of the individuals with a simulated mortgage change had a score change of more than 10 points.

The figures were somewhat higher for installment and revolving accounts. It is worth noting that the results suggest that the failure to report a satisfactory account can increase or decrease an individual's credit score. At first blush this appears to be surprising. However, it appears to stem from the characteristics of the basic scoring model. Although good payment performance information is weighed positively, accounts with high balances or utilization rates can negatively impact a score, sometimes by more than enough to offset the positive impact of a good performance history.

The failure of the largest student loan creditor to report any accounts. Recently, newspapers have reported that Sallie Mae, the nation's largest student lender, had stopped reporting information on any of their accounts to two of the three largest credit reporting agencies. Further, they asked that all historic information on their accounts be suppressed as well. (More recently, newspaper accounts indicate that Sallie Mae has reversed its initial decision.) Our sample of credit records does not include any information that would allow us to identify any specific creditor. Thus, to approximate the potential effect of the original decision by Sallie Mae on the credit scores of "at risk" individuals, we deleted information on the student loan accounts of lenders representing about 50 percent of all student loans in our sample (about the percentage currently held by Sallie Mae).

Results show that about 10 percent of "at risk" individuals would be affected. Further, results suggest that most affected "at risk" individuals would see an *increase* in their scores as a result of not reporting their student loan data. This somewhat counter-intuitive result is due to the fact that many (27 percent) of the reported student loans of "at risk" individuals have some record of a delinquency.

The failure of some creditors to report minor delinquencies. A small percentage of lenders do not report minor delinquencies. Since these delinquencies are not reported, it is impossible to know with certainty what the effect of reporting such items would be on the "at risk" population, or what proportion of the population would be affected. To simulate what the potential impact would be if an "at risk" individual were to have an account with a lender who did not report minor delinquencies, we randomly selected a currently-reported, non-major derogatory, seasoned (open at least two years) open revolving account with a lender who *did*

report minor delinquencies for each "at risk" individual and assigned "paying as agreed" performance status to each 30 and 60 day delinquency in their performance record. This replicates what the record would show for a lender who reports 30 and 60 day minor delinquencies as satisfactory.

Two-thirds of the "at risk" individuals would be unaffected if one of their lenders, selected randomly, had never reported minor delinquencies because at the time the sample was drawn they had not experienced any delinquencies on the account. As might be expected, when an individual is affected, the score almost always goes up, frequently by more than 10 points.

The failure of some creditors to report credit limits on revolving accounts. As noted, a relatively large proportion (about 30 percent) of revolving credit accounts, affecting about 70 percent of the individuals included in the Federal Reserve sample, were reported without information about credit limits. As a consequence, credit evaluators must use other means to derive credit utilization rates. As discussed above, the most common approach (and the one used for the roll-ups) is to substitute the high balance for the missing credit limit, with the typical result being higher calculated utilization rates.

We simulated the effect of the nonreporting of credit limits on "at risk" individuals, by creating an estimated credit limit for each revolving account missing a limit. Since we lack information on the true credit limit in these cases, our simulation in effect measures the impact of using scores based on our estimated limits versus the estimates used by the credit reporting agency. The primary difference between the two estimates, we believe, is that ours are constructed to be statistically unbiased. That is, our estimates reflect the "best guess" for the missing credit limit based on other information in the person's credit record. Specifically, we use samples of accounts of "at risk" individuals whose credit limits were reported to estimate a regression model that predicts the credit limit for revolving accounts for those missing limits. Independent factors included outstanding balances and high credit levels, the age and type of accounts, the types of lender, balances and limits on other accounts, and payment performance information. The resulting distribution of estimated credit limits and utilization for accounts with imputed limits was virtually identical to the distribution of accounts with reported limits within the "at risk" population, reflecting the fact that missing limits are primarily a function of the lender and almost unrelated to the characteristics of the account.

Over 55 percent of the "at risk" sample was affected by the use of our imputed credit limits. In almost all cases, this led to an increase in credit scores. Almost 30 percent of the effected individuals saw a rise of more than 10 points in their credit score. Only a small portion of the effected group saw a decrease in their score.

Duplications in public records. As discussed earlier, there is evidence that the reporting of some public record information may result in the addition of a new public record item rather than the appropriate updating of an existing item. This can occur because of changed docket numbers and other causes. To address the potential effects of this problem, we simulated the consolidation of likely duplicated public records into single items. Simulated duplicates were identified by matches on the recording date, amount owed, and creditor. In many instances, these involved the original filing of a judgment or lien followed by a record of a paid judgment or lien with all information identical to that in the first record except for the docket number. In other instances, there may have been technical refilings that ended up as duplicates with the same (or almost identical) information other than the docket number.

Very few individuals (less than 2 percent) in the "at risk" population were affected by this simulation. Further almost one-half of those that were effected experienced no material change in their credit score, none showed a change in score of more than 10 points and very few experienced a change of more than 1 point.[34]

Inconsistent reporting of lawsuits in public records. As described earlier, our analysis of credit records suggests lawsuits are inconsistently reported to the credit reporting agencies: Almost two-thirds of the individuals in our sample with a record of lawsuit resided in only three states. To simulate the potential effects of inconsistent reporting of these items, we removed all suits from the credit records of "at risk" individuals.

As with the duplicative public records, very few (1 percent) of the "at risk" individuals experienced any effect and one-third of the number who were impacted had no material change in their score.

[34] The coefficients for multiple public records in our approximation of the credit scoring model were statistically insignificant and, in some cases, had the wrong sign. This caused the somewhat unexpected result of some credit scores going down when a duplicate public record was eliminated. This also effected the simulation of the elimination of lawsuits from public records described below.

Inconsistent reporting of small collection agency accounts. Our analysis of collection accounts reveals that many are for very small amounts which may not be consistently reported. Recognizing this possibility, some credit evaluators chose not to include collection accounts involving small dollar amounts in their scoring models. Other lenders report that they remove small collection accounts involving medical services from credit evaluation as they often involve disputes with insurance companies over liability for the account. Unfortunately, this generally must be done with manual overrides since the collection credit record data do not include a code identifying claims associated with medical services. To test the potential effect of these restrictions we conducted two simulations. First, we eliminated all collection records involving items under $100. Second, we developed a code for the type of original creditor using the names of the plaintiffs which we used to identify collection items involving medical services. Our simulation involved eliminating all such items with amounts below $500.

Both simulations affect almost one-quarter of the "at risk" population. Moreover, about one-third of the affected group had a substantial increase (more than 10 points) in their credit score. This occurs because many "at risk" individuals have collection items in their credit records and more than one-half of these we estimate to involve medical services.

The failure to consolidate multiple inquiries for the same loan. Concerns have been raised that simple counts of inquiries in scoring models may unfairly penalize consumers who shop for credit. Credit evaluators have developed rules for mitigating this concern by consolidating multiple inquiries for loans of the same type within a short time period. Aside from the date, two pieces of information are available to model builders for this purpose—the type of loan for which the inquiry applies and the type of lender making the inquiry. Unfortunately, at least at the time our sample was drawn, the type of loan was missing for 98 percent of the inquiries, making use of this variable impractical. Lender-type was available and was used in the roll-ups by the credit reporting agency to distinguish inquiries from auto and real estate creditors from others, as proxies for mortgage and auto loan shopping. Unfortunately, analysis of existing mortgage and auto loans in our sample indicates that more than 90 percent of both types of loans were made by other types of lenders (primarily banks, auto finance companies, credit unions, and many mortgage companies).

It is not obvious how this problem can be corrected with existing data. To simulate the potential size of the effect of complete reporting of the type of inquiry we conducted two experiments. First, we identified all individuals in the "at risk" sample who had taken out a mortgage or auto loan in the two years prior to when the sample was drawn (inquiry information is kept for two years). For each loan, we consolidated inquiries in the two month (for autos) or four month (for mortgages) period preceding the date the loan was opened into a single inquiry. The second simulation was somewhat broader, and consolidated all inquiries within any two-week period into a single inquiry. This simulation, would take into account credit shopping which did not result in a loan.

The first simulation affected only 4 percent of the "at risk" population, primarily because most of them had not taken out a mortgage or auto loan during the period impacted by the simulation. However, almost all affected individuals showed some increase in their score, with most increases less than 10 points. The second simulation affected a larger (almost one-quarter) proportion of the population; but once again the effect on scores was relatively small (although almost always positive).

Conclusions

The information gathered by credit reporting agencies on the borrowing and payment experience of consumers is a cornerstone of the consumer credit system in this country and benefits both users of the data and consumers alike. Experience indicates that the information assembled and provided by credit reporting agencies enables domestic consumer credit and mortgage markets to function much more efficiently and at lower cost than would otherwise be possible. Despite the great benefits the current credit reporting system offers, concerns have been raised about the accuracy, timeliness, and completeness of credit records and the effects of such data limitations on access to credit and other services. Such concerns have grown as credit evaluation and decision-making in other areas, such as insurance underwriting; have become more reliant on credit history scores and less reliant on judgmental assessments based on manual reviews of credit reports.

Clearly, some minimum degree of accuracy and completeness of credit reports is required for the benefits of the credit reporting system to be realized. Moreover, the more accurate and complete the information assembled by credit reporting agencies, the greater the potential to

enhance efficiency in the credit granting process, reducing costs to the advantage of both consumers and creditors. Although inaccuracy can limit the potential efficiency benefits from a well-developed credit reporting system, as well as disadvantage (or advantage) individual consumers who have errors or incomplete information in their credit reports, there are limits to the degree of accuracy that can be obtained without undue costs. Some inadvertent errors and incomplete reporting inevitably will arise in a system comprising millions of account records and billions of transactions yearly. Moreover, in a purely voluntary system of reporting, competitive pressures on some creditors may limit their willingness to provide complete and timely information in all cases, or, in some instances, to even participate in the system.

From a public policy perspective, it is necessary to determine whether the incidence of errors and their impact on the actual credit granting process is significant enough to warrant consideration of new regulatory initiatives. In this paper, we have attempted to provide information that bears on both issues. First, we identify specific areas in which our research indicates data quality issues have arisen. Second, we identify the population in the marginal credit underwriting range where credit availability is most likely to be affected by a change in their credit records. Third, we devise simulations which approximate the effects of "corrections" to data quality issues for the "at risk" population identified in step two.

Our analysis suggests that many data quality issues have already been recognized and accounted for by credit score modelers, such as the developer of the score used in this paper, and, consequently, our corresponding simulations in many cases imply little change in the credit scores of "at risk" individuals. This is particularly true for stale account rules involving satisfactorily paid accounts. Other data quality issues, although potentially important for effected individuals, appear to be relatively rare. This is the case for public record duplication problems and stale account rules applied to accounts involving minor delinquencies.

Other data quality issues may have a more significant impact on credit availability either because they potentially affect a larger portion of the "at risk" population or because of the magnitude of their effects on the credit scores of the individuals impacted. Most prominent in this regard are data quality problems involving credit limits, collections, and inquiries. Changes in the treatment of inquiries could affect one-quarter of the "at risk" populations, although our analysis suggests that most individuals would see only a modest improvement in their scores with improved data. Discounting small collection items because of inconsistent reporting would

also effect about one-quarter of the "at risk" population, but here our simulations suggest that a fairly significant portion of the affected population would have a material improvement in their credit scores.

It appears that the failure to report credit limits has the largest effect of any data quality issue. More than one-half of the "at risk" population is affected by missing limits on revolving accounts and we estimate that about 30 percent of this group would see a material increase in their credit scores if all credit limits were reported. Finally, our experiment to simulate the effects of the nonreporting of satisfactory accounts by subprime lenders generated somewhat unexpected results. Although one might have expected that the failure to report satisfactory accounts would hurt consumers, our simulations suggest that this is not always the case. Further, this does not appear to stem from our model approximations, but rather it appears to stem from the characteristics of the basic scoring model which negatively weights accounts with high balances or utilization rates, thus offsetting the positive impact of a good performance history.

Our research has implications for the efficiency of the current credit reporting system error resolution process. The current system relies heavily on consumers to identify and take actions to dispute "incorrect" or missing items in their credit reports. One problem with this system is that consumers have no incentive to challenge information which is favorable to them, even if it is in error. Our research also indicates that even when data is incomplete or in error, it often has little or no bearing on an individual's credit score or their access to credit. Currently, consumers have access only to general information about the types of factors that are weighed in credit evaluation, or in the case of credit denials, the chief reasons for the adverse action. Lack of specific information may lead some consumers to believe that virtually any data quality issue is pertinent and should be disputed, causing the credit reporting agencies and reporters to incur unnecessary costs to correct or update files. Other consumers may be unaware of the potential importance of specific data issues, such as missing credit limits, and may not take appropriate action. Either situation is undesirable.

In evaluating these results, it is important to keep in mind that our sample data were drawn in mid-1999. There have been improvements in the data reporting system since that date. Most notably, industry sources report that credit limits are reported much more frequently than in the past. Mortgage account information is also reported more consistently. Further, consumers

have become more aware of the importance of the information in their credit records and are now more likely to check their credit reports for errors.

Before these results are taken as definitive estimates of the effect of data quality issues on credit availability, several important caveats need to be made. First, we have not investigated all potential sources of error. Most notably, we cannot say anything about the consequences of mistakenly including account records in an individual's file which do not belong to them. Second, we have used only one credit scoring model to simulate our results and have relied on our approximation to the model to quantify our results. Third, we have taken no account of manual reviews of credit records which are part of many underwriting systems. Many data quality issues are identified and addressed in such systems. Fourth, we have used data from only one credit reporting agency. It is common for creditors, particularly in the mortgage market, to obtain data from all three national credit reporting agencies for credit underwriting. Reconciling inconsistencies in data across the three agencies may lead to corrections of many of the data quality issues we have identified.

It is also important to bear in mind that in this paper we have only addressed the potential effects on credit scores of addressing data quality issues. We have said nothing about how such problems could be corrected, nor assessed the costs of doing so. Nor have we addressed the potential gains in efficiency that might result from developing models based on more complete and accurate data. If the current level of accuracy and completeness is not socially efficient reaching the optimal level may not be easy. Credit information has aspects of a classic public good. The parties that bear the costs of correcting errors or providing more timely and complete information may not receive much benefit from the improvement in accuracy. Further, remedies, such as imposing additional legal liability penalties may, in a system of voluntary reporting, lead to unintended consequences including less information reporting, and a less, rather than more, efficient system. Policy makers need to weigh all of these considerations in determining whether and how the current credit reporting system should be changed.

References

Avery, Robert B., Raphael W. Bostic, Paul S. Calem, and Glenn B. Canner. 2000. Credit Scoring: Issues and Evidence from Credit Bureau Files. Real Estate Economics 28: 523-547.

Avery, Robert B., Raphael W. Bostic, Paul S. Calem, and Glenn B. Canner. 2000. Credit Scoring, Credit Scoring and the Performance of Home Mortgages. Federal Reserve Bulletin 82: 21-48.

Avery, Robert B., Raphael W. Bostic, Paul S. Calem, and Glenn B. Canner. 2003. An Overview of Consumer Data and Credit Reporting. Federal Reserve Bulletin 89: 47-73.

Barron, James M. and Michael Staten. The Value of Comprehensive Credit Reports: Lessons from the U.S. Experience. Working paper. 1-30.

Bostic, Raphael B. and Paul S. Calem. 2003. *Privacy Restrictions and the Use of Data at Credit Repositories*. MIT Press, Cambridge, MA.

Cate, Fred H., Robert E. Litan, Michael Staten, and Peter Wallison. Financial Privacy, Consumer Prosperity, and the Public Good: Maintaining the Balance. Working paper. 1-33.

Credit Score Accuracy and Implications for Consumers. 2002. Study by the Consumer Federation of America and the National Credit Reporting Association. 1-48.

Golinger, Jon and Edmund Mierzwinski. 1998. Mistakes Do Happen: Credit Report Errors Mean Consumers Lose. Report by U.S. PIRG: 1-11.

Hillman, Richard J. 2003. Limited Information Exists On Extent of Credit Report Errors and Their Implications for Consumers. General Accounting Office (GAO-03-1036T): 1-18.

Hunt, Robert M. 2002. The Development and Regulation of Consumer Credit Reporting in America. Working paper 02-21. Federal Reserve Bank of Philadelphia.

Nott, Loretta and Angie A. Welborn. 2003. A Consumer's Access to a Free Credit Report: A Legal and Economic Analysis. Congressional Research Service (RL32008): 1-14.

Staten, Michael and Fred Cate. 2003. Does the Fair Credit Reporting Act Promote Accurate Credit Reporting. Paper presented at Building Assets, Building Credit: A Symposium on Improving Financial Services in Low-Income Communities, Joint Center for Housing Studies, November 18-19, 2003, Harvard University.

1. Distribution of individuals with credit reporting agency records, by status

Number and percent of individuals

Status	Number	Percent
Had credit account	216,202	87.2
Open and active account[1]	198,399	80.0
No active account	12,637	5.1
Authorized user only	5,166	2.1
MEMO		
Had credit account and:		
Public record	28,534	11.5
Collection account	46,559	18.8
Inquiry only[2]	139,584	55.9
No credit account	31,507	12.7
Collection	31,392	12.7
Authorized user	4,486	1.8
Public record	1,944	.8
Inquiry only[2]	3,321	1.3
None of the above	318	.1
All	248,027	100.0
MEMO		
Open and active account only[1]	59,188	23.9
Collection only	25,905	10.4
Public record only	42	*
Inquiry only[2]	55	*

* Less than 0.5 percent.

1. Active accounts are those used within one year of the date the sample was drawn.
2. Only includes inquiries made within two years of the date the sample was drawn.

2. All credit accounts and balances, grouped by status and distributed by account characteristic

Percent

Account characteristic	Account status									
	All		Currently reported				Not currently reported			
			Open		Closed		Dormant (zero balance)		Unknown (positive or unknown balance)	
	Percent	MEMO Percent of characteristic	Percent	MEMO Percent of characteristic	Percent	MEMO Percent of characteristic	Percent	MEMO Percent of characteristic	Percent	MEMO Percent of characteristic
Type										
Open-end										
Revolving										
Check credit	1.8	100.0	1.9	35.2	1.3	30.9	2.6	27.3	1.5	6.7
Bank credit	30.5	100.0	38.0	39.6	29.1	40.2	25.1	14.9	20.8	5.3
Finance company credit	4.7	100.0	4.4	29.3	3.1	27.5	9.6	36.7	3.9	6.4
Retail credit	23.8	100.0	24.8	33.2	10.1	17.9	53.8	41.1	23.7	7.7
Other	1.9	100.0	2.1	28.5	1.9	34.4	1.8	13.8	7.0	23.3
Total	62.7	100.0	71.2	36.1	44.3	29.9	95.4	27.6	51.5	6.4
Nonrevolving										
Total	4.7	100.0	4.1	27.9	4.0	36.4	4.6	18.0	10.7	17.8
Closed-end										
Installment										
Total	26.6	100.0	19.0	22.7	43.7	69.6	0.0	0.0	26.3	7.7
Mortgage										
Total	6.1	100.0	5.7	29.9	7.9	55.4	0.0	0.0	11.5	14.7
Total	100.0	100.0	100.0	31.8	100.0	42.3	100.0	18.2	100.0	7.8
MEMO										
Percent missing credit limit										
Revolving	31.6	100.0	32.8	52.2	0	0	38.9	33.9	24.2	13.8
Holder										
Single	78.9	100.0	80.0	32.3	74.8	40.2	85.3	19.6	81.0	8.0
Joint	21.1	100.0	20.0	30.1	25.2	50.4	14.7	12.6	19.0	7.0
Total	100	100.0	100	31.8	100	42.3	100	18.2	100	7.8
Creditor										
Banking institution	44.7	100.0	48.2	34.3	51.4	48.6	27.2	11.0	35.3	6.1
Finance company or credit union	19.8	100.0	14.9	24.0	26.9	57.7	10.2	9.4	22.9	9.0
Retail	24.8	100.0	25.0	32.1	12.1	20.7	54.1	39.7	24.2	7.6
Other	10.7	100.0	11.9	35.1	9.6	37.8	8.6	14.5	17.6	12.7
Date Opened										
# 1 year	8.1	100	19.6	77.0	1.9	10.0	3.2	7.2	6.1	5.8
1 to # 2 years	9.3	100	16.0	54.7	5.5	24.8	5.8	11.3	11.0	9.2
2 to 4 years	19.3	100	21.9	36.2	18.3	40.2	14.7	13.9	24.2	9.7
> 4 years	63.4	100	42.5	21.3	74.3	49.7	76.3	21.9	58.7	7.2
Total	100	100	100	31.8	100	42.3	100	18.2	100	7.8
Date Last Had Balance										
Current	31.0	100	67.1	68.7	4.6	6.3	0	0	100.0	25.0
# 1 year	13.8	100	17.3	39.8	13.6	41.6	14.2	18.6	0	0
1 to # 2 years	10.4	100	6.1	18.7	14.9	60.8	11.7	20.5	0	0
2 to 4 years	16.7	100	5.9	11.2	24.8	63.1	23.6	25.7	0	0
> 4 years	28.1	100	3.6	4.1	42.0	63.3	50.5	32.6	0	0
Total	100	100	100	31.8	100	42.3	100	18.2	100	0
Total	100.0	100.0	100.0	31.8	100.0	42.3	100.0	18.2	100.0	7.8

2. All credit accounts and balances, grouped by status and distributed by account characteristic (continued)

Percent

Account characteristic	Account status									
	All		Currently reported				Not currently reported			
			Open		Closed		Dormant (zero balance)		Unknown (positive or unknown balance)	
	Percent	MEMO Percent of charac-teristic	Percent	MEMO Percent of charac-teristic	Percent	MEMO Percent of charac-teristic	Percent	MEMO Percent of charac-teristic	Percent	MEMO Percent of characteristic
Date last reported										
2 months	39.8	100	100	80.0	18.8	20.0	0	0	0	0
2 month to 1 year	15.5	100	0	0	14.8	40.3	25.9	30.3	59.1	29.5
1 to 2 years	8.9	100	0	0	12.9	61.5	12.1	24.7	15.9	13.8
2 to 4 years	13.8	100	0	0	20.6	62.9	22.4	29.4	13.7	7.7
> 4 years	22.0	100	0	0	32.9	63.3	39.7	32.7	11.3	4.0
Total	100	100	100	31.8	100	42.3	100	18.2	100	7.8
Payment status[2]										
Historic worst										
Major derogatory	7.8	100	3.1	12.8	9.2	50.0	1.4	3.2	34.1	34.0
Minor derogatory	7.0	100	8.1	36.8	6.4	39.2	4.9	12.7	10.2	11.4
No derogatory	85.3	100	88.8	33.1	84.4	41.9	93.8	20.0	55.6	5.1
Total	100	100	100	31.8	100	42.3	100	18.2	100	7.8
At most recent report										
Has balance/balance unknown										
Major derogatory	4.3	100	2.1	15.1	2.7	26.3	0	0	32.5	58.5
Minor derogatory	1.0	100	1.6	50.7	.3	12.9	0	0	4.8	36.4
No derogatory	25.7	100	63.5	78.4	1.6	2.7	*	*	62.7	18.9
No balance	68.9	100	32.8	15.1	95.4	58.5	100	26.3	0	0
Total	100	100	100	31.8	100	42.3	100	18.2	100	7.8
MEMO[3]										
Number of accounts (millions)	1,428	100	454	31.4	604	42.7	259	18.2	111	7.8
Percent of dollars	100	100	. . .	71.8	. . .	1.2	. . .	0	. . .	27.0

NOTE. Here and in subsequent tables, data are a statistically representative sample of a national credit reporting agency's credit record data, as of June 30, 1999; distributions may not sum to 100 because of rounding.

1. A minor derogatory status is a payment delinquency of 30 days to 119 days. A major derogatory status is a delinquency of 120 days or more, a payment plan, repossession, charge-off, collection action, bankruptcy, foreclosure, or adverse judgment by a court.
2. National estimates based on the sample.

. . . Not applicable.

*Less than 0.05 percent.

SOURCE. Here and in subsequent tables, author calculations using statistically representative sample provided to the Federal Reserve Board by one of the three national credit reporting companies.

3a. Borrowing capacity on open accounts with credit limits reported, by type of account and creditor

Percent except as noted

Type of account and creditor	Accounts with credit limit reported			Distribution of accounts, by dollar size of credit limit						
	Percent of account type	Mean (dollars)	Median (dollars)	1-499	500-999	1,000-4,999	5,000-9,999	10,000-24,999	25,000 or more	Total
Open-end										
Revolving										
Check credit	84.3	12,002	3,500	6.1	12.2	35.6	15.5	15.5	15.1	100.0
Bank credit	60.1	7,036	6,000	3.1	5.4	27.8	39.5	22.4	1.8	100.0
Finance company credit	88.4	3,467	2,500	4.5	10.5	60.9	19.2	4.4	0.5	100.0
Retail credit	71.9	1,575	1,000	15.9	30.3	47.8	5.6	0.4	*	100.0
Other	74.5	2,808	2,500	3.2	11.3	71.6	13.0	1.0	*	100.0
Total	67.5	4,534	2,500	8.5	16.3	40.5	22.4	11.0	1.3	100
Closed-end										
Installment										
Total	99.5	11,152	7,060	2.6	4.3	33.9	18.5	32.6	8.3	100.0
Mortgages										
Total	99.6	92,797	75,400	*	*	0.3	0.9	7.7	91.1	100.0

NOTE. Excludes accounts in a major derogatory status (for definition, see table 1, note 1), or no reported credit limit, or in dispute.
* Less than 0.05 percent.

3b. Borrowing capacity on open accounts with credit limits not reported, by type of account and creditor

Percent except as noted

Type of account and creditor	Accounts with high balance reported			Distribution of accounts, by dollar size of high balance							MEMO	
	Percent of account type	Mean (dollars)	Median (dollars)	1-499	500-999	1,000-4,999	5,000-9,999	10,000-24,999	25,000 or more	Total	High balance not reported	High balance reported
Revolving												
Check credit	15.7	9,887	2,471	6.2	11.7	37.1	16.6	14.0	14.4	100.0	17.8	82.2
Bank credit	39.9	1,605	374	30.7	16.3	38.5	11.7	2.5	0.3	100.0	33.7	66.3
Finance company credit	11.6	3,396	1,520	14.6	13.8	51.1	11.4	7.5	1.6	100.0	9.6	90.4
Retail credit	28.1	484	310	64.7	22.8	12.3	0.2	*	0.0	100.0	10.0	90.0
Other	25.5	522	400	52.0	32.7	15.2	0.2	0.0	0.0	100.0	16.9	83.1
Total	32.5	1,351	353	43.8	19.2	27.9	6.9	1.8	0.4	100.0	24.7	75.3

NOTE. Excludes accounts in a major derogatory status (for definition, see table 1, note 1), or no reported credit limit, or high balance, or in dispute.

* Less than 0.05 percent.

4. Use of borrowing capacity on open accounts

Percent

Type of account and creditor	Distribution of accounts, by percent of limit used							MEMO: Percent of limit used, accounts with a balance	
	0	1-24	25-49	50-74	75-94	95 or more	Total	Mean	Median
	Credit limits reported								
Revolving	55.1	20.0	6.8	5.8	6.6	5.8	100.0	19.6	0.0
Check credit	51.2	9.4	7.4	8.9	12.8	10.2	100.0	30.6	0.0
Bank credit	41.1	26.5	7.9	7.0	9.1	8.4	100.0	26.0	2.3
Finance company credit	38.2	26.7	9.8	8.6	9.3	7.5	100.0	27.3	5.3
Retail credit	73.4	12.7	4.9	3.7	2.9	2.4	100.0	10.5	0.0
Other	64.9	15.4	7.4	5.1	4.0	3.2	100.0	14.4	0.0
Installment	0.4	8.5	13.0	20.8	25.2	32.1	100.0	72.7	81.7
Mortgage	*	2.9	4.7	11.7	31.6	49.1	100.0	86.2	94.7
	Credit limits not reported (highest-balance proxy used)								
Revolving	48.7	5.3	4.2	4.8	5.7	31.2	100.0	41.3	5.5
Check credit	51.3	6.4	6.2	8.6	12.2	15.4	100.0	34.3	0.0
Bank credit	40.0	2.2	1.9	3.5	5.8	46.7	100.0	54.8	85.4
Finance company credit	52.8	6.3	6.5	8.2	12.1	14.1	100.0	32.9	0.0
Retail credit	63.1	11.2	8.3	6.9	4.9	5.7	100.0	18.5	0.0
Other	69.2	4.7	5.3	5.8	4.9	10.1	100.0	20.3	0.0

NOTE. Excludes accounts in a major derogatory status (for definition, see table 1, note 1), no reported credit limit, or high balance or in dispute.
* Less than 0.05 percent.

5. The estimated effects of "correcting" data accuracy, completeness and timeliness issues on the credit scores of individuals whose credit scores fall in the range 600-660

Data issue	Individuals not affected (%)	Individuals affected						Memo:	
		No change in score	Credit score declines (%)		Credit score increases (%)		Total	Effect on credit score (change in points)	
		(%)	1-9 points	10 or more points	1-9 points	10 or more points	(%)	Mean for those with negative change	Mean for those with positive change
Failure to close satisfactory account:									
Mortgage	96.2	11.7	46.4	3.9	37.2	0.9	100.0	-3.6	2.5
Installment	89.5	10.1	28.9	7.6	49.6	3.8	100.0	-5.9	4.1
Revolving	82.7	6.9	16.1	6.1	49.4	21.6	100.0	-7.3	8.2
Failure to close open, minor delinquent account:									
Mortgage	99.7	1.7	15.5	8.6	50.0	24.1	100.0	-7.9	7.3
Installment	99.2	4.3	17.1	4.9	45.1	28.7	100.0	-6.9	9.5
Revolving	98.8	0.8	14.5	5.8	36.5	42.3	100.0	-9.2	13.5
Failure of mortgage or finance company to report an open, satisfactory credit accounts:									
Mortgage	Unknown	11.3	41.1	2.1	43.9	1.6	54.5	-3.2	2.9
Installment	Unknown	6.1	60.2	16.0	15.0	2.6	100.0	-7.1	4.9
Revolving	Unknown	12.4	40.3	8.7	33.0	5.7	100.0	-6.3	5.0
Failure of largest student loan creditor to report:	90.6	8.1	24.5	4.2	37.0	26.2	100.0	-5.2	13.7
Failure to report a minor delinquency:	Unknown	65.9	6.2	1.4	14.6	11.9	100.0	-6.4	15.1
Failure to report a credit limit:	43.9	18.2	4.3	0.3	47.8	29.4	100.0	-3.5	9.4
Reporting of duplicate public records:	98.4	49.1	36.5	0.0	14.4	0.0	100.0	-1.0	1.2
Effect of including a lawsuit	99.0	31.4	15.0	3.6	18.0	32.0	100.0	-6.8	24.4
Reporting of collection accounts under $100	78.4	30.1	0.1	0.1	33.7	36.0	100.0	-12.4	19.5
Reporting of medical collection accounts under $500	75.3	14.4	0.1	0.1	36.1	49.3	100.0	-6.8	23.3
Effect of not consolidating multiple inquiries for auto and mortgage loans	96.0	5.5	3.7	0.3	85.0	5.6	100.0	-4.1	3.8
Effect of not consolidating multiple inquiries	76.8	2.0	1.6	0.1	91.4	5.0	100.0	-3.1	3.6