Data Point: Credit Invisibles

The CFPB Office of Research

Consumer Financial Protection Bureau

May 2015

- Kenneth P. Brevoort
- Philipp Grimm
- Michelle Kambara

This is another in an occasional series of publications from the Consumer Financial Protection Bureau's Office of Research. These publications are intended to further the Bureau's objective of providing an evidence-based perspective on consumer financial markets, consumer behavior, and regulations to inform the public discourse.

Table of Contents

1. Introduction

Consumers with limited credit histories reflected in the credit records maintained by the three nationwide credit reporting agencies (NCRAs) face significant challenges in accessing most credit markets.[1] NCRA records are often used by lenders when making credit decisions. In particular, lenders often use credit scores, such as one of the FICO or VantageScore scores, that are derived entirely from NCRA records when deciding whether to approve a loan application or in setting a loan's interest rate. If a consumer does not have a credit record with one of the NCRAs or if the record contains insufficient information to assess her creditworthiness, lenders are much less likely to extend credit. As a result, consumers with limited credit histories can face substantially reduced access to credit.

In broad terms, consumers with limited credit histories can be placed into two groups. The first group is comprised of consumers without NCRA credit records. We refer to this group as "credit invisibles." The second group includes consumers who, while they have NCRA credit records, have records that are considered "unscorable," meaning they contain insufficient credit histories to generate a credit score. Generally speaking, a credit record may be considered unscorable for two reasons: (1) it contains insufficient information to generate a score, meaning the record either has too few accounts or has accounts that are too new to contain sufficient payment history to calculate a reliable credit score; or (2) it has become "stale" in that it contains no recently reported activity. The exact definition of what constitutes "insufficient" or "stale" information differs across credit scoring models, as each model uses its own proprietary definition. Our analysis is based on a commercially-available credit scoring model that we believe uses a relatively narrow definition of a "scorable" credit record, but one that we believe is consistent with most credit scores used today. We refer to these records as "unscored" rather

[1] The three NCRAs are Equifax, Experian, and TransUnion.

than “unscorable” to reflect the fact that other credit scoring models might generate scores for these records. Nevertheless, we believe our estimates of the population with unscored credit records accurately reflect the circumstances faced by consumers with limited credit histories.

The challenges that credit invisibles and consumers with unscored records face in accessing credit markets has generated considerable attention from researchers and industry participants. Several studies have explored the potential of various types of “alternative data” to supplement the information contained in the NCRA credit records and allow credit scores to be generated for these consumers.[2] Stakeholders have debated the implications of doing so for those with limited credit history as well as those with scorable files whose credit profiles might change with the addition of such data. Several industry participants have also developed scoring products that are aimed specifically at these populations.[3]

Despite all of this attention, very little is known about the number or characteristics of credit invisibles or consumers with unscored credit records. This Data Point documents the results of a research project undertaken by Staff in the Office of Research of the Consumer Financial Protection Bureau (CFPB) to better understand how many consumers are either credit invisible or have unscored credit records and what the demographic characteristics of such consumers are.

This analysis was conducted using the CFPB’s Consumer Credit Panel (CCP), a 1-in-48 longitudinal sample of de-identified credit records purchased from one of the NCRAs and representative of the population of consumers with credit records. This dataset contains information on almost 5 million consumer credit records. While these data contain no direct-identifying information (such as name, address, or Social Security Number), for each credit

---

[2] For example, see Turner, *et al.* (2006), Experian (2014), and Schneider and Schutte (2007) for utility payments, Experian RentBureau (2014) for rental payments, and CFPB (2014) for remittances.

[3] For example, FICO recently announced that it is launching a pilot project that extends the number of consumers whose records can be scored using alternative data on utility and telecommunication bill payments and property record data (FICO, 2015). LexisNexis has also introduced a credit scoring model, RiskView, that uses alternative data to expand the number of credit records that can be scored (Feinstein, 2013). The new version of the VantageScore, version 3.0, uses alternative data when it is available on a credit record to expand the number of consumers whose records can be scored (VantageScore, 2013). For other examples, see Jacob and Schneider (2006).

record we observe the consumer's census tract, year of birth, and a commercially-available credit score.

We use these data from December 2010 to estimate the number of credit invisibles in each tract by taking the difference between the number of adults living in the tract according to the 2010 Decennial Census and the number of credit records in each tract, as estimated from the CCP. Since the 2010 Census publishes data on population by age, and since the CCP contains year of birth, we estimate the number of credit invisibles in each tract for each of thirteen different age groups. For each of these age groups, we also calculate the number of consumers with unscored credit records in each tract using the CCP. Then using variation across census tracts in the racial and ethnic composition of the population and their household incomes, which we take from the 2008-2012 American Community Survey, we estimate how the incidence of being credit invisible or having an unscored credit record differs across these demographic characteristics.

Key findings of this report include:

- As of 2010, 26 million consumers in the United States were credit invisible, representing about 11 percent of the adult population. An additional 19 million consumers, or 8.3 percent of the adult population, had credit records that were treated as unscorable by a commercially-available credit scoring model. These records were about evenly split between those that were unscored because of an insufficient credit history (9.9 million) and because of a lack of recent history (9.6 million).
- There is a strong relationship between income and having a scored credit record. Almost 30 percent of consumers in low-income neighborhoods are credit invisible and an additional 15 percent have unscored records. These percentages are notably lower in higher-income neighborhoods. For example, in upper-income neighborhoods, only 4 percent of adults are credit invisible and another 5 percent have unscored credit records.
- Blacks and Hispanics are more likely than Whites or Asians to be credit invisible or to have unscored credit records. About 15 percent of Blacks and Hispanics are credit invisible (compared to 9 percent of Whites and Asians) and an additional 13 percent of Blacks and 12 percent of Hispanics have unscored records (compared to 7 percent of Whites). These differences are observed across all age groups, suggesting that these differences materialize early in the adult lives of these consumers and persist thereafter.

2. Data

2.1 Data Sources

The data used in this study come from three sources. The first is the CFPB's Consumer Credit Panel (CCP), a longitudinal sample of approximately 5 million de-identified credit records that is nationally representative of the credit records maintained by one of the NCRAs. This study primarily uses data from December 2010; however, as described below, we also use information for these same consumers from December 2014 in cleaning the data.

For each time period, the entire credit record is supplied in the CCP, excluding any direct-identifying personal information (such as name, address, or Social Security Number). In addition to the credit records, the CCP includes a commercially-available credit score, which we use to indicate which records were scored and which were not. For each unscored record, an "exclusion code" is provided indicating why the record could not be scored using the model for the commercially-available credit score.

Like most credit scoring models, the model that generated the scores in the CCP was built to predict future credit performance (that is, the likelihood, relative to other borrowers, that a consumer will become 90 or more days past due on a credit obligation in the following two years).[4] In some cases, the model builders will determine that a credit record does not contain enough information to make a suitably reliable prediction. During score development, these records are excluded and are unscored by the model going forward.

[4] This is a generic definition of "credit performance" used in credit scoring models. The exact definition used will vary from one credit scoring model to another. For more information on measures of performance in credit scoring models, see Board of Governors of the Federal Reserve System (2007).

There are two types of unscored records in the CCP.[5] The first, "insufficient unscored" records, do not contain enough information to generate the score, meaning either that the record contained too few reported accounts or accounts that did not have a sufficiently long credit history. The second type, "stale unscored" records, do not contain any recently reported information. Our analysis examines these two types of unscored credit records separately.

When available, a year of birth is included in the CCP for each record.[6] We use this information to calculate the age of each consumer at the end of 2010. This allows us to examine how the incidence of being credit invisible or having an unscored credit record varies with age. Though credit records in the CCP do not include address information, each consumer's census tract using 2010 census definitions is provided. This allows us to measure how credit records are distributed across the country.

The second source of data used in this study is the 2010 Decennial Census, conducted by the U.S. Census Bureau. The Decennial Census indicates the number of consumers in each census tract. It also provides information on the racial and ethnic mix of each tract. In our analysis, we focus on four different racial or ethnic groups: Hispanics or Latinos ("Hispanics"), Non-Hispanic Asians ("Asians"), Non-Hispanic Blacks or African Americans ("Blacks"), and non-Hispanic Whites ("Whites"). All other non-Hispanic racial groups, which include American Indians or Alaska Natives, Native Hawaiians or Other Pacific Islanders, or multi-racial individuals, are included in a category labelled "Other."

The third source of data comes from the 2008-2012 American Community Survey (ACS), which is also conducted by the U.S. Census Bureau. Among other information, the ACS includes the median household income in each tract, county, and Metropolitan Statistical Area (MSA). We use this information to calculate the "relative income" of each tract. Relative income is defined as the ratio between the median household income of the tract and the median household income of the surrounding area, which is the MSA for urban tracts or the county for rural tracts. Following the definitions used in the Community Reinvestment Act, we then characterize each

[5] Credit records in the CCP will also be unscored if the record belongs to a deceased consumer. Our analysis focuses on living consumers whose records will only be unscored for these two reasons.

[6] Though credit records also contain the month and day of birth for consumers, the CCP does not include this information to help maintain the privacy of the consumers in our sample.

tract as low, moderate, middle, or upper income, depending on whether the tract's relative income is below 50 percent, between 50 and 80 percent, between 80 and 120 percent, or above 120 percent.

2.2 Dataset Creation

Estimating the number of credit invisibles is complicated by the fact that almost no data exists specifically for this population. Some datasets that collect data on a representative sample of the entire population, like the Survey of Consumer Finances or the ACS, certainly include information on credit invisibles, but do not collect information that allows one to determine which sample observations are credit invisible. Datasets like the CCP generally have good information about consumers with credit records but by definition cannot include consumers without credit records.

Our approach is to estimate the number of credit invisibles by comparing the adult population in the U.S. from the 2010 Decennial Census with an estimate of the number of adults who have a credit record at the NCRAs. While this may seem straightforward, it is actually a complex undertaking. The reason is that many consumers have multiple credit records within the data of the NCRAs. As a result, comparing the number of credit records maintained by the NCRAs with the U.S. population would be misleading. For example, the CCP in 2010 contained 4.91 million credit records. Given the 1-in-48 sampling rate used by the CCP, this implies that there were about 236 million credit records at the NCRA, more than the 235 million adults in the U.S. according to the Census. By itself, this would suggest that there are no consumers without credit records.

The reason that some consumers have multiple credit records is the existence of "fragment files." These are credit records containing a portion of a consumer's credit history that exist outside of the consumer's primary file. Take for example a consumer with a credit record who opens a new credit card. When the lender or servicer first reports the account, the NCRA attempts to match it with the correct credit record using a proprietary algorithm. If, based on that algorithm, the NCRA is unable to find any credit records that match, or is unable to find a unique match, perhaps reflecting erroneous or incomplete information reported with the new account, then the newly reported credit card will be placed in its own credit record. Most of these fragment files are temporary. Over time, as more information comes in, the NCRA may determine that the credit record is a fragment and that the accounts in the record belong to a consumer with an existing credit record. When this happens, the information in the fragment file gets subsumed in the consumer's primary credit record and the fragment file ceases to exist.

Fragment files present an interesting challenge for estimating the number of consumers who have credit records at the NCRAs. An accurate measurement requires pruning from the data those records that are likely to be fragment files; otherwise, we will overestimate the number of consumers with a credit record and underestimate the number of credit invisibles. For example, as discussed above, without any pruning the CCP (or other data based on credit records) would imply that all Americans have credit records or, possibly, that there are more records than people.

Our process of cleaning the data involves the following exclusions. First, since we are comparing credit records to the U.S. population, we exclude credit records that indicate the consumer was living outside of the fifty states. Second, we exclude the credit records of consumers who appear to be deceased in December 2010.

We then use hindsight to identify fragment files. We discard any credit record from December 2010 that does not appear in the December 2014 data as well, suggesting that the record had been purged from the database or merged into another record during this time. Finally, we exclude any credit record that had no reported year of birth in either December 2010 or December 2014. Birth dates tend to be an important characteristic in matching accounts to credit records. Accounts that lack this information are less likely to be (uniquely) matched to an existing credit record and are more likely to be placed into a fragment file. Any CCP record that was missing year-of-birth information for four years should also have been missing date-of-birth information in the records maintained by the NCRA over this period, which suggests that these records are fragments containing accounts that could not be linked.[7] We discuss these exclusions in more detail in Appendix A.

Once we have removed the likely fragment files, we estimate the number of credit invisibles in each tract as the difference between the tract's adult population according to the 2010 Decennial

---

[7] This is further supported by the prevalence of authorized user accounts in these files. Authorized users are people who are permitted to use a revolving account, like a credit card, without being legally liable for any of the charges that are incurred. Lenders generally do not require a lot of detail on these consumers and, based on our conversations with industry participants, their accounts often end up in fragment files as a result. For more information on authorized user accounts and the issues involved, see Brevoort, Avery, and Canner (2012).

Census and our estimate of the number of consumers in the tract who have a credit record. Since the 2010 Decennial Census provides tract-level information on population by age, we are able to calculate the number of credit invisibles for each of thirteen different age groups: Eleven age groups are defined using five-year spans of ages from 20 through 74 (i.e, 20 to 24, 25 to 29) and the remaining two contain 18-to-19 year olds and those 75 or older. We also estimate the number of consumers with insufficient-unscored and stale-unscored credit records from the CCP for each tract at each of the 13 age groups.

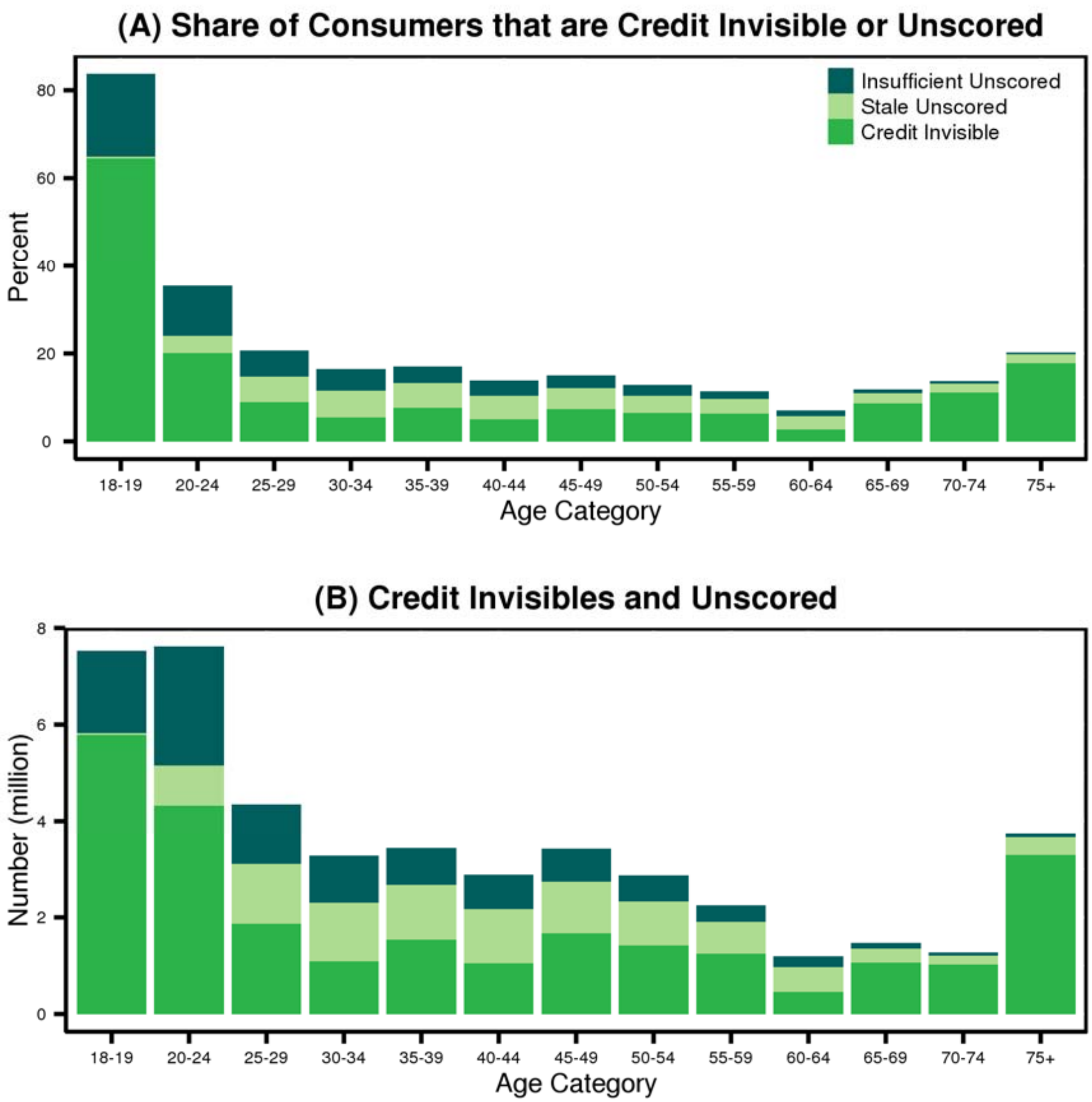
These estimates of the number of credit invisibles and consumers with unscored credit records depend crucially on the exclusions described earlier in this section. To the extent that some of the excluded credit records may have been the primary records of consumers, our estimates of the number of credit invisibles will be overstated and the number of consumers with a credit record (scored or unscored) will be understated. In contrast, if we have failed to exclude some credit records that are fragment files, then our estimates will tend to understate the number of credit invisibles and potentially overstate the number of consumers with credit records. One exclusion that we considered imposing, but decided against, was removing consumers whose only item on their credit record was a third-party debt collection or public record (such as a tax lien). While some of these are likely fragment files, our analysis suggested that removing these would likely exclude too many primary files. As a result, we believe that our estimate of the number of credit invisibles is likely low and our estimate of the number of consumers with unscored credit records likely overstated slightly since debt-collection-only or public-record-only credit records tend to be unscored. We provide additional detail on the consequences of each of these exclusions for our estimates of the number of credit invisibles and consumers with unscored records in Appendix A.

3. How Many Americans Have Limited Credit Histories?

Our estimates suggest that approximately 188.6 million Americans have credit records at one of the NCRAs that can be scored by the commercially-available model that informs our analysis. This represents over 80 percent of the adult population. An additional 19.4 million Americans, representing 8.3 percent of the adult population, have credit records that cannot be scored. These are almost evenly split between consumers with credit records that are insufficient unscored (9.9 million) and those that are stale unscored (9.6 million). The remaining 11 percent of adults, or about 26 million Americans, are credit invisible.

Credit history is something that consumers establish over the course of a lifetime. As a result, one would expect the problem of limited credit history to be more concentrated among the young. This pattern is observed in the data. Panel (A) of figure 1 shows the share of consumers in each age group that are credit invisible, have unscored records because of insufficient information, or have unscored records because of a lack of recently reported information. As shown, over 80 percent of 18 or 19 year olds are credit invisible or have unscored records. This percentage drops substantially for older consumers, falling below 40 percent in total for the 20 to 24 year old age group. After age 60, the number of consumers that are credit invisible or that have an unscored record increases with age. With our existing data, it is difficult to determine to what extent this reflects an age effect (a greater tendency of credit histories to shrink or become stale with age), a cohort effect (in which people born earlier than 1950 had thinner credit histories over the course of their lives, possibly reflecting less credit reporting during the periods of their lives when they were actively using credit), or some combination.

FIGURE 1: INCIDENCE AND NUMBER OF CONSUMERS THAT ARE CREDIT INVISIBLE OR HAVE RECORDS THAT ARE UNSCORED



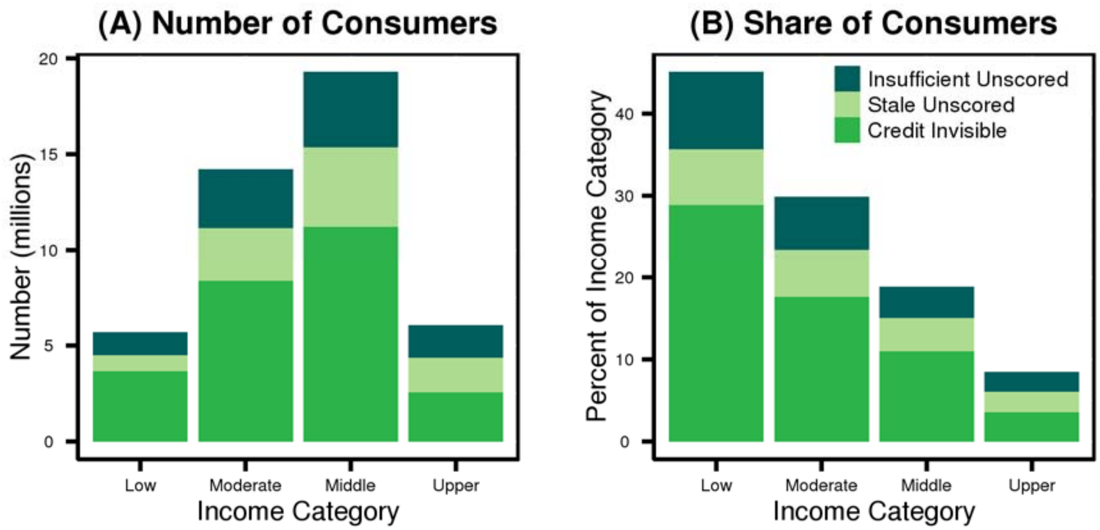
The data shown in panel (A) also indicate that the causes of an unscored credit record differ substantially by age. The share of consumers with an unscored credit record because of an insufficient credit history declines with age. Only a small percentage of consumers aged 65 or older have records that are unscored because of an insufficient history; instead, most of the unscored records for these older consumers are the result of a lack of recent information. Interestingly, having a stale-unscored credit record is not strongly related to age. In fact, the incidence of a stale unscored record is higher for consumers aged 25 to 49 than it is for consumers older than 50.

As this suggests, most consumers that are credit invisible or that have an unscored credit record are young. Panel (B) of figure 1 shows the distribution of the number of consumers who are credit invisible or have unscored records. Over 10 million of the estimated 26 million credit invisibles are younger than 25. Consumers in this age group also account for a disproportionate share of insufficient-unscored credit records. In contrast, most consumers with stale-unscored records are middle aged. Consumers aged between 25 and 50 account for over half of stale-unscored credit records.

Other characteristics besides age may also affect the likelihood of being credit invisible or having an unscored credit record. Among these is income. If higher-income consumers have an easier time qualifying for traditional credit, even without credit histories, then they may be more likely than lower-income consumers to open credit cards, auto loans, or other forms of credit that are frequently reported to the NCRAs. Relatedly, if lower-income consumers have a more difficult time qualifying for traditional credit and, as a result, rely on non-traditional sources like payday or auto-title lenders, then this will exacerbate the differences by income as these non-traditional sources of credit generally do not report information to the NCRAs.

Exploring the relationship between income and the incidence of being credit invisible or having an unscored record is complicated by the fact that credit records do not contain income information. As a result, we do not know the income levels of the consumers whose credit records are in the CCP and, thus, rely on the relative income of each census tract as an alternative measure. Panel (A) of figure 2 shows the number of consumers that are credit invisible or have an unscored credit record who live in census tracts with each of the four relative income levels: low, moderate, middle, or upper. As shown, middle-income tracts account for a larger portion of the credit invisible and unscored population than any of the three other income groups. Consumers from low- and upper-income neighborhoods, in particular, make up a notably smaller share of the credit invisible and unscored population.

FIGURE 2: NUMBER AND INCIDENCE OF CONSUMERS THAT ARE CREDIT INVISIBILE OR HAVE AN UNSCORED CREDIT RECORD BY CENSUS TRACT INCOME LEVEL



By and large, these numbers reflect varying population sizes in each income category. There are many more consumers in middle-income tracts than in low-income tracts, so it is not surprising that so many of these invisible and unscored consumers come from middle-income tracts. Instead, if we look at the share of consumers who are credit invisible or have an unscored credit record at each of these income levels, shown in panel (B), we see a very different pattern. Almost 50 percent of consumers in low-income tracts appear to either lack a credit record entirely or have an unscored credit record (mostly because of an insufficient credit history). At higher-income levels, this incidence falls sharply. In comparison, fewer than 10 percent of consumers in upper-income tracts are credit invisible or have unscored records. So while low-income tracts appear to comprise a relatively small share of the credit invisible or unscored population (about 5 million of the total 45 million consumers), this represents a significant share of the population in those tracts.

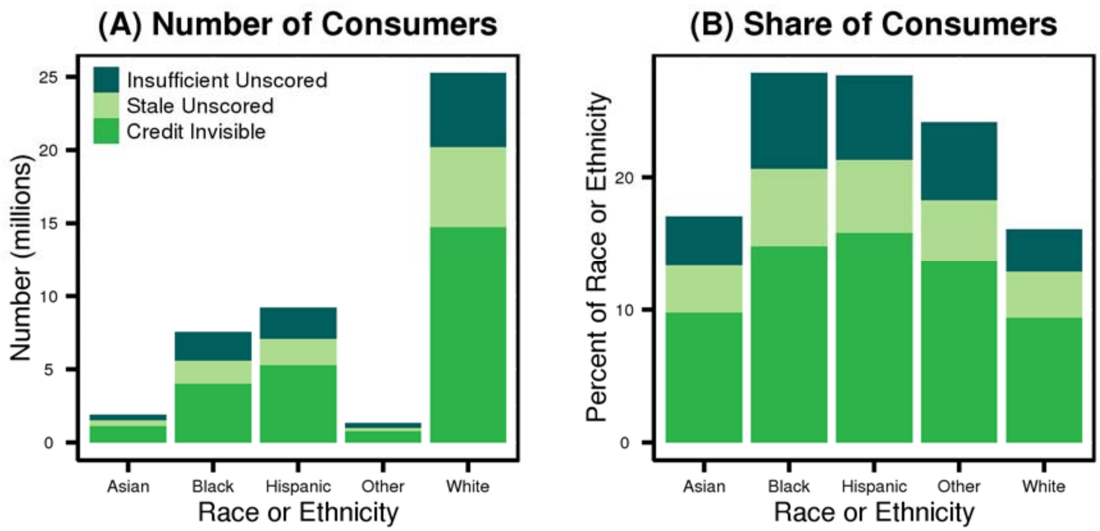
4. Patterns of Limited Credit History by Race and Ethnicity

4.1 Patterns by Race or Ethnicity

Another characteristic that has been mentioned in connection with consumers that are credit invisible or have unscored credit records is race or ethnicity. As with income, credit records do not contain any information about the race or ethnicity of the consumer. As a result, we do not observe this information for the consumers whose credit records are in the CCP and, unlike income, we cannot easily segment census tracts into different racial or ethnic groups. This analysis, therefore, requires a different approach than we used in the previous section.

To explore how the incidence of being credit invisible or having an unscored record varies with race or ethnicity, we examine cross-tract variation in the racial composition of census tracts and in the number of consumers who are credit invisible or have unscored records. Specifically, for each tract, we estimate the number of consumers in each of the thirteen age groups who are credit invisible. We then use the racial mix of the tract in each age group from the 2010 Decennial Census to estimate the racial or ethnic mix of credit invisibles, assuming for these purposes that the distribution of credit invisibles in any given tract is proportionate to the racial and ethnic composition of the tract (i.e., that members of each racial or ethnic group in a given tract have an equal chance of being credit invisible). For example, if we find that a tract has 100 credit invisibles in a given age group, and that tract's population in that age group is 15 percent Black, 10 percent Hispanic, 5 percent Asian, and 70 percent White, then we would assume that 15 of these credit invisibles were Black, 10 were Hispanic, 5 were Asian, and the remaining 70 were White. We make this calculation for each tract, at each age level, and aggregate the numbers nationally. We estimate the racial and ethnic mix of consumers with unscored records using the same method.

FIGURE 3: NUMBER AND INCIDENCE OF CONSUMERS THAT ARE CREDIT INVISIBILE OR HAVE AN UNSCORED CREDIT RECORD BY RACE OR ETHNICITY



The results of these aggregations are shown in figure 3. Panel (A) shows our estimate of the distribution of consumers who are credit invisible or have unscored records across the five different racial or ethnic groups used in this study. The patterns are largely consistent with the overall shares of these racial or ethnic groups in the population at large. Most consumers who are credit invisible or have unscored credit records are White. Minorities account for a smaller share of the population that is credit invisible or has an unscored record, largely reflecting the fact that minorities make up a smaller portion of the overall U.S. population.

Panel (B) shows the percentage of consumers from each of the five racial or ethnic groups who, using our estimates, are credit invisible or have an unscored credit record. Whites are the least likely racial or ethnic group to be credit invisible or to have an unscored credit record, though the rates for Asians are almost identical. Blacks and Hispanics, as well as those included in the "Other" racial category, are notably more likely to be credit invisible or to have an unscored record than Whites. Though Hispanics are slightly more likely than Blacks to be credit invisible, Blacks appear to be more likely than Hispanics to have unscored records.[8]

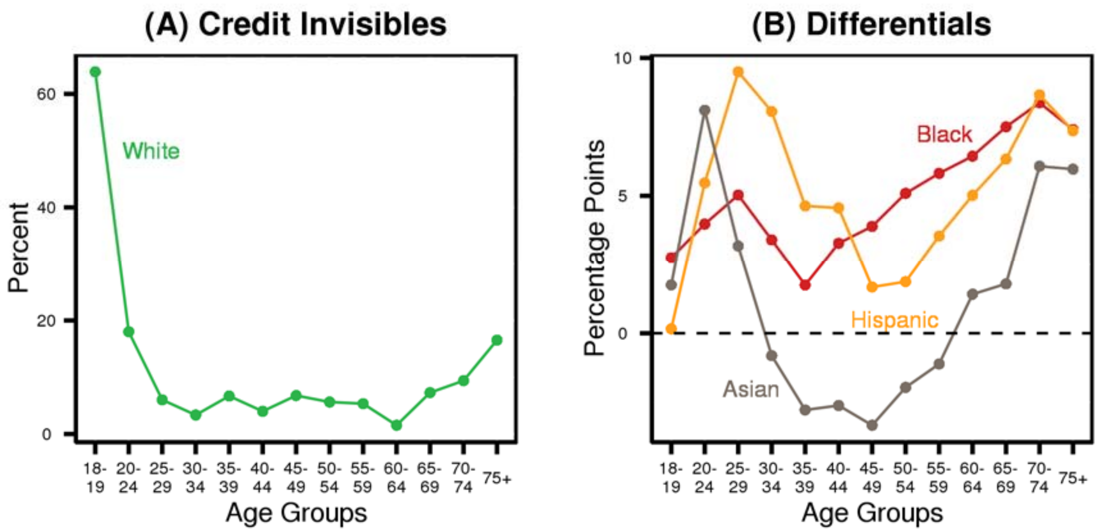
In calculating these estimates of the racial and ethnic mix of consumers with limited credit histories, we assumed that, within each census tract, consumers of each race or ethnicity had an equal likelihood of being credit invisible or having an unscored record. Our results suggest that consumers in census tracts with relatively more Blacks or Hispanics are more likely to be credit invisible or have an unscored credit record. Since we observe this pattern across tracts, it is likely that a similar pattern holds within tracts as well. If true, our estimate of the number of Blacks or Hispanics who are credit invisible or have an unscored credit record is likely underestimated and the number of Whites or Asians overestimated.

4.2 Racial or Ethnic Patterns by Age

The results by race or ethnicity suggest that minority populations, other than Asians, are generally more likely to be credit invisible or have unscored credit records. As shown in the previous section, the incidence of these forms of minimal credit history is strongly correlated with age. To better understand how these differences across racial or ethnic groups emerge over the course of a lifetime, we also compare the incidence of being credit invisible or having an unscored record by age across the different racial or ethnic groups. Because we do not observe how credit records change with age, we are unable to disentangle the effects of age from cohort effects, such as the different macroeconomic environments that consumers in different age groups have faced. Nevertheless, while not conclusive, the results of this analysis may provide some evidence about whether these differences emerge at young ages and, if so, whether they tend to dissipate with age.

[8] One factor that may distort these figures is the undercounting of minorities (and overcounting of Whites) in the 2010 Decennial Census. The Census Bureau's post-enumeration survey for the 2010 Census found that, while the 2010 Census was the most accurate to date, the White population may have been overcounted by 0.85 percent. Blacks may have been undercounted by 2.1 percent and Hispanics by 1.5 percent (Mule, 2012). These results would suggest that we are overcounting the number of White credit invisibles and undercounting the number of Black or Hispanic credit invisibles. These changes would not have had a notable effect on the number of consumers with unscored records, though the percentage of Blacks and Hispanics unscored records would be slightly smaller.

FIGURE 4: INCIDENCE OF BEING CREDIT INVISIBLE BY AGE AND RACE OR ETHNICITY



The share of consumers in each age group who are credit invisible is shown for four racial or ethnic groups in figure 4.[9] For most racial or ethnic groups, the age patterns are very similar, so we present the results slightly differently to sharpen the contrasts. The graph on the left, panel (A), shows the results for Whites, which we use as the baseline group. The results suggest that the incidence of being credit invisible is very high for 18-19 year olds, but then falls sharply. The share holds relatively steady after age 25, until it begins to increase with age after 60.

The graph on the right, panel (B), which is shown at a magnified scale, shows the results for Asians, Blacks, and Hispanics relative to the pattern for Whites. For example, among consumers who are aged 25-29, Blacks are 5 percentage points more likely than Whites in the same age group to be credit invisible. From the left graph, panel (A), we can see that about 6 percent of Whites are credit invisible, which means that 11 percent of Blacks are credit invisible at this age.

---

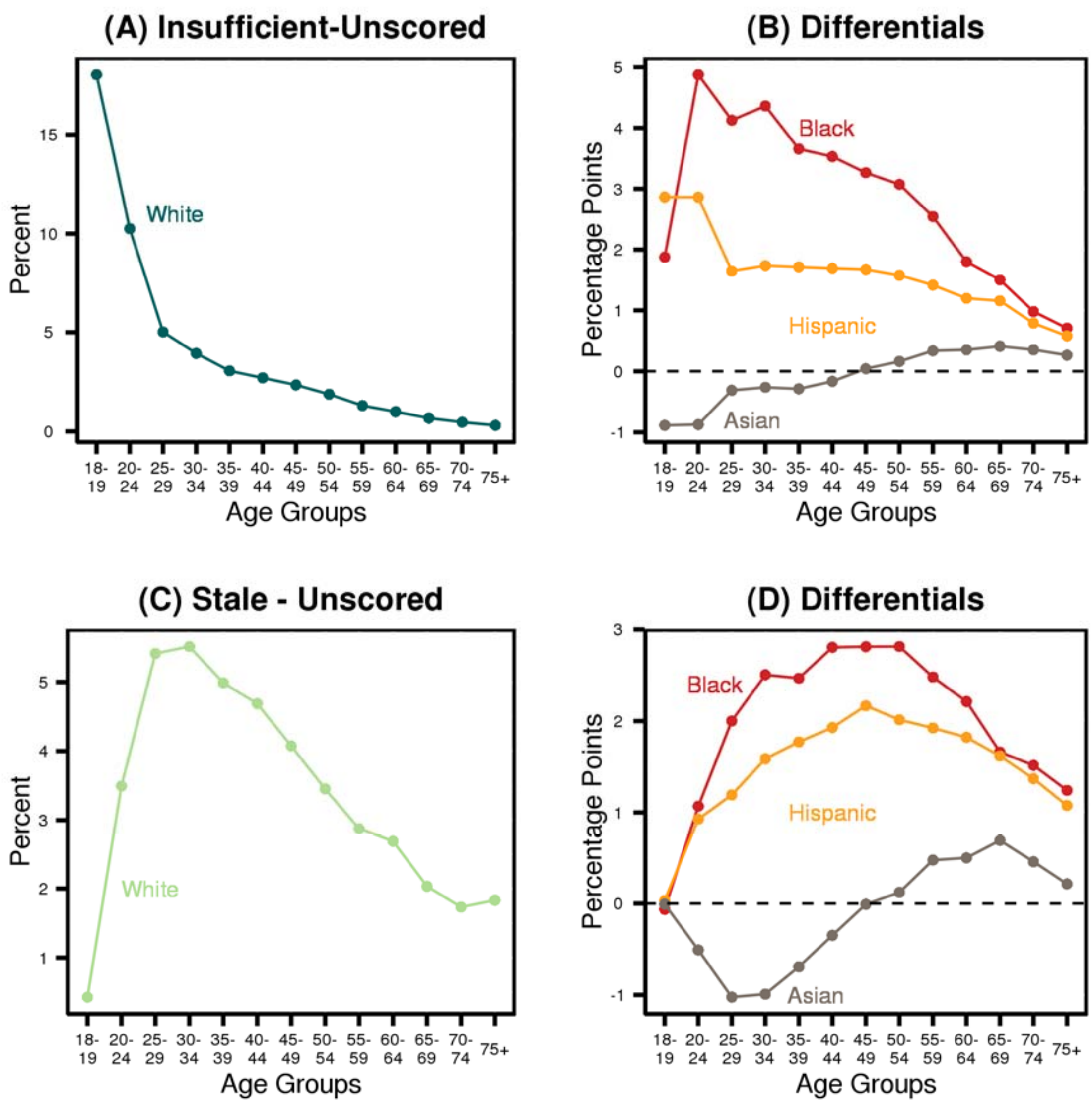
[9] Results for the "Other" racial group are omitted from the remaining graphs of this section to reduce the amount of clutter; however, they are provided in the tables in Appendix B.

The patterns in this graph indicate that Blacks and Hispanics are more likely than Whites to be credit invisible at almost every age. While we are unable to determine with our existing data whether these reflect age or cohort effects, these patterns suggest that the relatively higher incidences of being credit invisible for Blacks and Hispanics emerge at young ages and tend to persist over time. The difference between Whites and Asians is much less consistent across ages. Like Blacks and Hispanics, Asians younger than 30 or older than 60 are more likely to be credit invisible than are Whites; however, Asians aged 30 to 59 have a lower incidence of being credit invisible. This suggests that the relative equality between Whites and Asians in terms of the aggregate incidences of being credit invisible (shown earlier in table panel (b) of figure 3) conceals significant differences across ages.

Figure 5 shows a similar analysis for the incidence of unscored credit records. The left panels show the incidence for Whites of having an unscored record because of an insufficient credit history, panel (A), or a lack of recent history, panel (C). The right panels, (B) and (D), show the patterns by age for the other three racial or ethnic groups relative to Whites for these two types of unscored records, respectively.

The results indicate that the share of Whites with a credit record that is unscored because of an insufficient credit history declines steadily by age, as shown in panel (A). Blacks and Hispanics have consistently higher likelihoods of having an insufficient unscored credit record (panel (B)). The differences are largest at younger ages. While they decline for older consumers, Blacks and Hispanics of all ages are more likely than Whites to have an insufficient-unscored credit record. While young Asians are less likely, and older Asians more likely, than Whites to have an insufficient-unscored credit record, the gap between Asians and Whites remains less than one percentage point across all age groups.

FIGURE 5: INCIDENCE OF HAVING AN INSUFFICIENT-UNSCORED OR STALE-UNSCORED CREDIT RECORD BY AGE AND RACE OR ETHNICITY

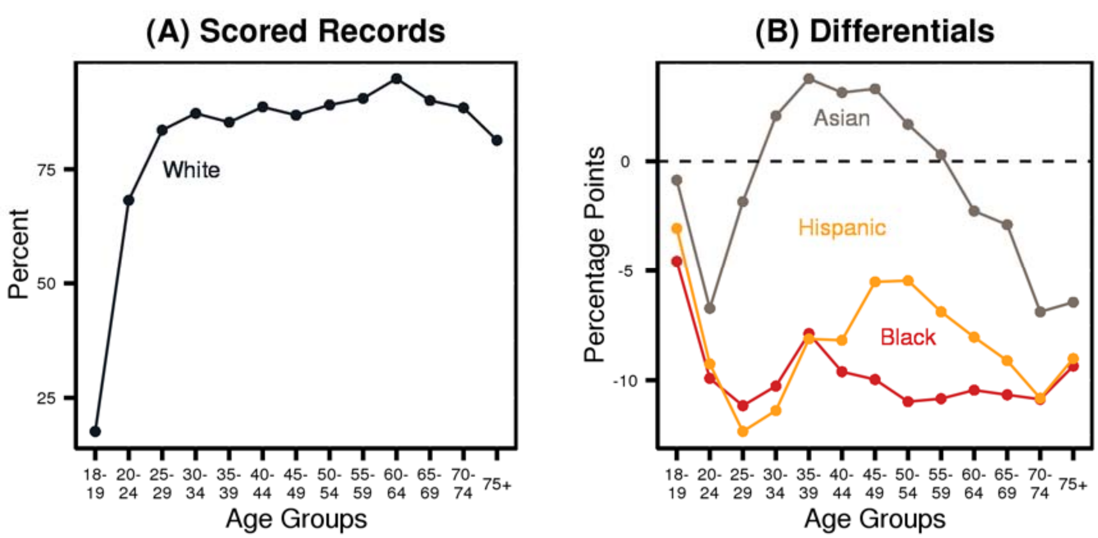


The patterns by age for records that are stale-unscored are somewhat different. For Whites, the likelihood of having a stale-unscored record increases with age until around 25-34 (panel (C)). It then declines with age thereafter. The share of Black or Hispanic consumers who have a stale-unscored record is consistently higher than Whites at almost all age levels (the exception being 18-19 years of age, when the likelihood of having a stale unscored record is near zero for all consumers). This gap increases with age until the mid-40s and declines thereafter. While the gap with Whites declines at older ages, Blacks and Hispanics appear to be consistently more likely to have a stale-unscored credit record.

Like the pattern observed for insufficient-unscored credit records, young Asians are less likely, and older Asians more likely, than Whites to have stale-unscored credit records. Again, however, the gap between Asians and Whites remains within 1 percentage point at all age levels, suggesting that the patterns for Asians and Whites are similar.

Taken together, these results suggest that Blacks and Hispanics are more likely to be credit invisible or to have unscored credit records. These differences are observed for all age groups, which suggests that these differences emerge at young ages and persist over the lifetimes of these consumers.

FIGURE 6: INCIDENCE OF HAVING A SCORED CREDIT RECORD BY AGE AND RACE OR ETHNICITY



The combined effects of being credit invisible or having an unscored credit record are shown in figure 6, which depicts the share of consumers at each age with a scored credit record. Again, the left panel shows the pattern by age for Whites and the right panel shows the relative patterns for Asians, Blacks, and Hispanics.

Panel (A) of this figure shows that the share of Whites with a scored credit record increases sharply with age up to around age 30 and then increases more gradually through age 60. At older ages, the incidence of having a scored credit record decreases somewhat with age. Blacks and Hispanics, shown in panel (B), are less likely than Whites to have a scored credit record at very early ages. This gap widens with age, becoming greater than 10 percentage points for ages 25-29 for both groups, and remains large thereafter, though it does narrow (particularly for Hispanics around 50 years of age, though this narrowing is not observed at older ages). As our earlier results would suggest, the pattern for Asians is somewhat different. At early ages, they are less likely to have scored credit records than are Whites; however, this gap shrinks during their 20s and disappears in their 30s to 50s, during which time they are more likely to have a scored credit record than Whites in the same age group. Asians older than 54, however, are less likely to have credit records than Whites.

Overall, these patterns suggest that the problem of limited credit history affects all racial or ethnic groups. Nevertheless, Blacks and Hispanics appear more likely to be credit invisible or have an unscored record. These differences are observed across all age levels.

5. Conclusions

The three NCRAs have traditionally been the sole source of information used to calculate credit scores like the scores produced by FICO or VantageScore. Consumers with limited credit histories established in the records of the three NCRAs generally have a harder time obtaining credit as a result because many lenders do not extend credit to consumers without a scored credit record or do so only in quite narrow circumstances. While there has been a lot of attention paid to the problem of limited credit history and to various forms of alternative data that might mitigate it, very little is known about the number of consumers who are affected and even less is known about their demographic characteristics.

This report uses data from the CFPB's Consumer Credit Panel and aggregate information from the 2010 Decennial Census and 2008-2012 American Community Survey, both conducted by the U.S. Census Bureau, to construct estimates of the number of consumers with limited credit histories. Our results suggest that there are 26 million adults in the United States without a credit record. This amounts to 11 percent of U.S. adults. Additionally, our results suggest that another 19 million adults (about 8 percent) have credit records that are considered "unscorable" by the commercially-available credit scoring model used in this analysis. These records are almost evenly split between those that are unscored because of an insufficient credit history (too few accounts) and those that are unscored because of a lack of recently reported credit history.

Our results also suggest that there is a strong relationship between income and having a credit record. Almost 30 percent of consumers in low-income neighborhoods are credit invisible and an additional 16 percent have unscored records. These percentages are notably lower in higher-income neighborhoods. For example, in upper-income neighborhoods, only 4 percent of the population is credit invisible and another 5 percent has an unscored record.

Additionally, our results suggest that there are significant differences in the incidence of having a limited credit history across racial and ethnic groups. While Whites and Asians are almost equally likely to be credit invisible or have an unscored record, the shares of Blacks and Hispanics with limited credit history are much larger. About 15 percent of Blacks and Hispanics are credit invisible (compared to 9 percent of Whites and Asians) and an additional 13 percent of Blacks and 12 percent of Hispanics have unscored records (compared to 7 percent of Whites). This elevated incidence of being credit invisible or having an unscored credit record is observed across ages, suggesting that these differences across racial and ethnic groups materialize early in the adult lives of these consumers and persist thereafter. These results suggest that the problems that accompany having a limited credit history are disproportionally borne by Blacks, Hispanics, and lower-income consumers.

References

Avery, Robert B., Paul S. Calem, Glenn B. Canner, and Raphael W. Bostic. 2003. "An Overview of Consumer Data and Credit Reporting," *Federal Reserve Bulletin* 89(2): 47-73.

Brevoort, Kenneth P., Robert B. Avery, and Glenn B. Canner. 2013. "Credit Where None is Due? Authorized-User Account Status and Piggybacking Credit," *Journal of Consumer Affairs* 47(3): 518-47.

Board of Governors of the Federal Reserve System. 2007. *Report to the Congress on Credit Scoring and Its Effects of the Availability and Affordability of Credit*. (Washington, DC: Federal Reserve Board).

Consumer Financial Protection Bureau. 2014. Report on the Use of Remittance Histories in Credit Scoring. Available online at http://files.consumerfinance.gov/f/201407_cfpb_report_remittance-history-and-credit-scoring.pdf.

Experian. 2014. "Let there be Light: The Impact of Positive Energy-Utility Reporting on Consumers," *Experian White Paper*.

Experian RentBureau. 2014. "Credit for Renting: The Impact of Positive Rent Reporting on Subsidized Housing Residents," *Experian RentBureau White Paper*.

Feinstein, Jeffrey. 2013. "Alternative Data and Fair Lending," LexisNexis White Paper. August. Available online at http://www.lexisnexis.com/risk/downloads/whitepaper/fair_lending.pdf (Last visited April 15, 2015).

FICO. 2015. "FICO, LexisNexis Risk Solutions & Equifax Joining to Generate Trusted Alternative Data Scores for Millions More Americans," FICO Press Release, April 2.

Available online at http://www.fico.com/en/newsroom/fico-lexisnexis-risk-solutions-and-equifax-joining-to-generate-trusted-alternative-data-scores-for-millions-more-americans-04-02-2015 (Last visited April 15, 2015).

Jacob, Katy, and Rachel Schneider. 2006. “Market Interest in Alternative Data Sources and Credit Scoring,” *The Center for Financial Services Innovation*, December.

Mule, Thomas. 2012. “Census Coverage Measurement Estimation Report: Summary of Estimates of Coverage for Persons in the United States,” May 22, 2012. Available online at http://www.census.gov/coverage_measurement/pdfs/g01.pdf (Last visited, April 15, 2015).

Schneider, Rachel and Arjan Schutte. 2007. “The Predictive Value of Alternative Credit Scores,” *Center for Financial Services Innovation Report*, November.

Turner, Michael A., Alyssa Stewart Lee, Ann Schnare, Robin Varghese, and Patrick D. Walker. 2006. “Give Credit Where Credit Is Due: Increasing Access to Affordable Mainstream Credit Using Alternative Data,” *Political and Economic Research Council and the Brookings Institution Urban Markets Initiative Report*.

VantageScore. 2013. “VantageScore 3.0: Better Predictive Ability Among Sought-After Borrowers.” VantageScore White Paper Series. December. Available online at http://www.vantagescore.com/images/resources/VantageScore3-0_WhitePaper.pdf (Last visited, April 15, 2015).

APPENDIX A:

Effect of Fragment File Exclusions

Estimates of the number of credit invisibles or consumers with unscored credit records depend crucially on the decisions about which credit records in the sample to identify as likely fragment files. In this Appendix, we provide detail about the exclusions that were applied in pruning the data of likely fragment files and the effects that these had on our results.

Before identifying likely fragment files, we excluded credit records for consumers living outside of the United States. Most of these were credit records for consumers living in Puerto Rico and other U.S. territories. About 44,000 records were excluded for this reason. We also excluded credit records that indicated the consumer was deceased in 2010. These exclusions were necessary to focus on the population of interest and make the credit record population as comparable to the Census data as possible.

Once these exclusions were made, this left a sample with 4.7 million credit records. From these we removed two groups of credit records that we believed were most likely fragment files. The first of these groups included credit records that disappeared between December 2010 and December 2014. In total, there were 242,727 records excluded for this reason. These records were split into two subgroups. The first subgroup included 138,152 credit records that were identified as having been consolidated into existing credit records, which is a direct identifier of a fragment file. The second subgroup included an additional 104,575 credit records that, while we could find no record of having been consolidated into an older credit record disappeared during the four years.

The second group included the credit records of consumers whose credit records were missing year-of-birth information in both December 2010 and December 2014. There were 153,152 records excluded for this reason. Date of birth is an important factor used by the NCRAs in matching reported account information to credit records. The fact that these files had no year-of-birth information in the CCP indicates that the credit record maintained by the NCRA did not have a date of birth and suggests that these files contain account information that the NCRA was unable to uniquely match to a primary file. Consistent with this, most of these credit records contain authorized user accounts.

In addition to these exclusions, our estimates were also affected by records that were not excluded, but that might have contained a large share of fragment files. In particular, we considered excluding those credit records that only contained information reported by third-party debt collectors or information from public records. Based on conversations with industry participants, we believe that NCRAs have a more difficult time finding unique matches for information from these sources. This suggests that these types of records may contain a large number of fragment files. Nevertheless, we believe that excluding all such records would have excluded too many primary credit records and chose to include these records in our estimates.

Table 1 shows the effect that each of these exclusions and inclusions had on our estimates of the number of consumers who are credit invisible or have an unscored credit record. The first line of the table shows our estimate presented in the body of this Data Point. The following lines show the effect that each exclusion or inclusion had on the overall estimate. For example, the decision to exclude credit records that were missing in 2014 because they likely were fragment files had the effect of decreasing our estimate of the number of scored records by 4.3 million and increasing our estimate of the number of credit invisibles by 11.7 million. Similarly, our decision to include credit records that contained only third-party debt collection accounts increased our estimate of the number of consumers with a scored credit record by 0.1 million and decreased our estimate of the number of credit invisibles by 2.1 million.

TABLE 1: EFFECTS OF SAMPLE EXCLUSIONS AND INCLUSIONS

	Scored Records	Credit Invisibles	Stale-Unscored	Insufficient-Unscored
Baseline Estimate	188.6	26	9.6	9.9
Exclusions:				
Missing in 2014 (Total)	-4.3	+11.7	-1.7	-5.7
Observed Merge	-3.0	+6.6	-1.2	-2.4
Disappeared	-1.2	+5.0	-0.5	-3.3
Missing Age	-6.5	+7.4	-0.4	-0.5
Inclusions:				
Debt Collection Only	+0.1	-2.1	+0.09	+1.9
Public Record Only	+0.01	-0.4	+0.01	+0.3

APPENDIX B:

Data Used in Figures

This Appendix provides the underlying data used to produce each of the figures in the text.

TABLE 2: SHARE OF CONSUMERS THAT ARE CREDIT INVISIBLE OR UNSCORED, DATA FOR FIGURE 1(A)

Age Group	Credit Invisible	Stale-Unscored	Insufficient-Unscored
18 to 19 years	64.5	0.4	18.9
20 to 24 years	20.2	3.8	11.5
25 to 29 years	8.9	5.9	5.9
30 to 34 years	5.5	6.1	4.9
35 to 39 years	7.6	5.6	3.9
40 to 44 years	5.1	5.4	3.4
45 to 49 years	7.4	4.7	3.0
50 to 54 years	6.4	4.0	2.4
55 to 59 years	6.3	3.4	1.7
60 to 64 years	2.7	3.1	1.3
65 to 69 years	8.6	2.3	0.9
70 to 74 years	11.1	2.0	0.6
75 years and over	17.8	2.0	0.4

TABLE 3: THE NUMBER OF CONSUMERS THAT ARE CREDIT INVISIBLE OR HAVE AN UNSCORED CREDIT RECORD BY AGE, DATA FOR FIGURE 1(B)

Age Group	Credit Invisible	Stale-Unscored	Insufficient-Unscored
18 to 19 years	5.8	0.0	1.7
20 to 24 years	4.3	0.8	2.5
25 to 29 years	1.9	1.2	1.2
30 to 34 years	1.1	1.2	1.0
35 to 39 years	1.5	1.1	0.8
40 to 44 years	1.1	1.1	0.7
45 to 49 years	1.7	1.1	0.7
50 to 54 years	1.4	0.9	0.5
55 to 59 years	1.2	0.7	0.3
60 to 64 years	0.5	0.5	0.2
65 to 69 years	1.1	0.3	0.1
70 to 74 years	1.0	0.2	0.1
75 years and over	3.3	0.4	0.1

TABLE 4: NUMBER OF CONSUMERS THAT ARE CREDIT INVISIBLE OR HAVE AN UNSCORED CREDIT RECORD BY CENSUS TRACT INCOME LEVEL, DATA FOR FIGURE 2(A)

Income Group	Credit Invisible	Stale-Unscored	Insufficient-Unscored
Low	3.7	0.9	1.2
Moderate	8.4	2.8	3.1
Middle	11.2	4.1	3.9
Upper	2.6	1.8	1.7

TABLE 5: INCIDENCE OF BEING CREDIT INVISIBLE OR HAVING AN UNSCORED CREDIT RECORD BY CENSUS TRACT INCOME LEVEL, DATA FOR FIGURE 2(B)

Income Group	Credit Invisible	Stale-Unscored	Insufficient-Unscored
Low	28.9	6.8	9.5
Moderate	17.6	5.8	6.5
Middle	11.0	4.1	3.8
Upper	3.6	2.5	2.4

TABLE 6: NUMBER OF CONSUMERS THAT ARE CREDIT INVISIBLE OR HAVE AN UNSCORED CREDIT RECORD BY RACE OR ETHNICITY, DATA FOR FIGURE 3(A)

Race or Ethnicity	Credit Invisible	Stale-Unscored	Insufficient-Unscored
Black	4.0	1.6	2.0
Hispanic	5.3	1.8	2.1
Asian	1.1	0.4	0.4
Other	0.7	0.3	0.3
White	14.7	5.5	5.1

TABLE 7: INCIDENCE OF BEING CREDIT INVISIBLE OR HAVING AN UNSCORED CREDIT RECORD BY RACE OR ETHNICITY, DATA FOR FIGURE 3(B)

Race or Ethnicity	Credit Invisible	Stale-Unscored	Insufficient-Unscored
Black	14.8	5.8	7.2
Hispanic	15.8	5.5	6.4
Asian	9.8	3.6	3.7
Other	13.7	4.6	5.9
White	9.4	3.5	3.2

TABLE 8: INCIDENCE OF BEING CREDIT INVISIBLE BY AGE AND RACE OR ETHNICITY, DATA FOR FIGURES 4(A) AND 4(B)

Age Group	White	Black	Hispanic	Asian	Other
18 to 19 years	63.9	66.6	64.0	65.6	65.3
20 to 24 years	18.0	22.0	23.5	26.1	21.0
25 to 29 years	6.0	11.1	15.5	9.2	9.3
30 to 34 years	3.3	6.7	11.4	2.5	5.5
35 to 39 years	6.7	8.4	11.3	3.9	7.3
40 to 44 years	4.0	7.3	8.6	1.4	5.6
45 to 49 years	6.8	10.7	8.5	3.4	9.5
50 to 54 years	5.6	10.7	7.5	3.7	7.9
55 to 59 years	5.3	11.2	8.9	4.2	8.1
60 to 64 years	1.5	8.0	6.6	3.0	4.6
65 to 69 years	7.3	14.8	13.6	9.1	10.2
70 to 74 years	9.4	17.8	18.0	15.5	13.4
75 years and over	16.5	23.9	23.9	22.5	20.0

TABLE 9: INCIDENCE OF HAVING AN INSUFFICIENT-UNSCORED CREDIT RECORD BY AGE AND RACE OR ETHNICITY, DATA FOR FIGURE 5(A) AND 5(B)

Age Group	White	Black	Hispanic	Asian	Other
18 to 19 years	18.0	19.9	20.9	17.1	18.9
20 to 24 years	10.3	15.1	13.1	9.4	12.4
25 to 29 years	5.0	9.1	6.7	4.7	6.3
30 to 34 years	3.9	8.3	5.7	3.7	5.2
35 to 39 years	3.1	6.7	4.8	2.8	4.4
40 to 44 years	2.7	6.2	4.4	2.5	4.1
45 to 49 years	2.3	5.6	4.0	2.4	3.5
50 to 54 years	1.9	4.9	3.4	2.0	3.0
55 to 59 years	1.3	3.8	2.7	1.6	2.2
60 to 64 years	1.0	2.8	2.2	1.3	1.8
65 to 69 years	0.7	2.2	1.8	1.1	1.4
70 to 74 years	0.5	1.4	1.3	0.8	1.0
75 years and over	0.3	1.0	0.9	0.6	0.8

TABLE 10: INCIDENCE OF HAVING A STALE-UNSCORED CREDIT RECORD BY AGE AND RACE OR ETHNICITY, DATA FOR FIGURE 5(C) AND 5(D)

Age Group	White	Black	Hispanic	Asian	Other
18 to 19 years	0.4	0.4	0.5	0.4	0.4
20 to 24 years	3.5	4.6	4.4	3.0	3.9
25 to 29 years	5.4	7.4	6.6	4.4	6.1
30 to 34 years	5.5	8.0	7.1	4.5	6.4
35 to 39 years	5.0	7.5	6.8	4.3	6.0
40 to 44 years	4.7	7.5	6.6	4.3	5.8
45 to 49 years	4.1	6.9	6.2	4.1	5.2
50 to 54 years	3.5	6.3	5.5	3.6	4.7
55 to 59 years	2.9	5.4	4.8	3.3	3.9
60 to 64 years	2.7	4.9	4.5	3.2	3.7
65 to 69 years	2.0	3.7	3.7	2.7	2.9
70 to 74 years	1.7	3.3	3.1	2.2	2.3
75 years and over	1.8	3.1	2.9	2.0	2.3

TABLE 11: INCIDENCE OF HAVING A SCORED CREDIT RECORD BY AGE AND RACE OR ETHNICITY, DATA FOR FIGURE 6(A) AND 6(B)

Age Group	White	Black	Hispanic	Asian	Other
18 to 19 years	17.7	13.1	14.6	16.8	15.4
20 to 24 years	68.2	58.3	59.0	61.5	62.7
25 to 29 years	83.6	72.4	71.2	81.7	78.3
30 to 34 years	87.2	76.9	75.8	89.3	82.8
35 to 39 years	85.3	77.4	77.2	89.0	82.3
40 to 44 years	88.6	79.0	80.4	91.7	84.5
45 to 49 years	86.8	76.8	81.3	90.1	81.8
50 to 54 years	89.0	78.1	83.6	90.7	84.4
55 to 59 years	90.5	79.6	83.6	90.8	85.8
60 to 64 years	94.8	84.3	86.7	92.5	89.9
65 to 69 years	90.0	79.3	80.9	87.1	85.5
70 to 74 years	88.4	77.5	77.6	81.5	83.3
75 years and over	81.3	72.0	72.3	74.9	77.0