MARCH 2017

Negative Data Suppression and Impacts on Credit Score Models

Contents

VantageScore

a higher level of confidence

Negative Data Suppression and Impacts on Credit Score Models

INTRODUCTION

Risk management processes utilize a diversity of consumer credit behavior information, often embedded within generic credit risk scoring models designed to assess the likelihood that a consumer will default on future debts. Conceptually, these models build empirical relationships between various categories of historical data in order to predict this future likelihood. Positive information from behaviors that demonstrate competency in credit management, such as paying monthly obligations on time, signal less likelihood to default. Conversely, negative information representing high-risk behaviors, such as failing to pay debts or spending up to the maximum credit limit, signal a higher likelihood to default. Other categories of negative information commonly used by credit scoring models include medical or non-medical agency collections, tax liens and civil judgments. Often, these data can have a substantial negative effect in the calculation of a consumer's likelihood to default. Many credit score models developed prior to 2009 include all of these data in their likelihood-to-default calculations.

BACKGROUND

In June 2016, Equifax, Experian and TransUnion announced a series of initiatives intended to enhance credit data reporting accuracy. Presented as the National Consumer Assistance Plan, (NCAP, www.nationalconsumerassistanceplan.com), key initiatives include:

- To require all data furnishers to use the most current reporting format.
- To monitor data furnishers for adherence to the announced reporting requirements and to take corrective actions for non-compliance by data furnishers.
- Regarding medical agency collections:
  - To prohibit medical debts from being reported on credit reports until after a 180-day waiting period has expired in order to allow insurance payments to be applied.
  - To remove from the credit report, any previous medical collections that have been paid by insurance or are being paid by insurance.
- Regarding all agency collections:
  - To require debt collectors to include the original creditor information with every account being reported for collection.
  - To require debt collectors to regularly update the status of unpaid debts, and to remove debts no longer being pursued for collection.

While a comprehensive analysis of the impact of these initiatives on data volumes is still underway, currently the known consequences to a consumer credit file are:

- Potential removal of all civil judgments.
- A substantial reduction in the number of tax liens.
- A reduction in medical-related agency collections, specifically those less than 180 days old.

VantageScore Solutions has published a research study and several presentations on the *impact to consumer credit scores* when these data (i.e., tax liens judgments and agency collections) are no longer available to a model that was originally developed to include these data. Generally speaking, the absence of these data causes consumers' scores to improve because they are no longer penalized by the inclusion of these negative events. A more comprehensive discussion can be found in the VantageScore white paper, *Impact to VantageScore 3.0 Credit Score Model from Revisions to Public Records Reporting*, at VantageScore.com/PublicRecordImpact.

The omission of such negative events, however, calls into question the ongoing predictive quality of credit score models given these data changes. This paper specifically considers two questions:

- Does the absence of these data irreparably reduce credit score model predictive performance or can other behaviors be used to similarly predict the risk of a consumer's future default?
- How do consumers score using models developed without the NCAP-related data? And are their underlying credit behaviors more or less attractive to lenders?

EXECUTIVE SUMMARY

This study, in analyzing the removal of certain negative data, concludes the following:

- Credit score models can certainly recover the predictive performance lost due to the absence of highly negative information such as tax liens, judgments and agency medical collections.
- The performance and composition of models developed without NCAP data clearly delivers equivalent predictive insight. Newer models may provide greater stability because they incorporate behaviors more closely related to the consumer's current and potential performance.
- Consumers scored using these new models who exceeded the score cut-off, reflected a more stable product mix and demonstrated superior credit management skills compared with those scored using older models that incorporate NCAP-related data.
- Given the regulatory focus on the NCAP data in addition to the analytic insights presented in this study, lenders should evaluate whether their incumbent scoring models satisfy this regulatory focus while continuing to deliver optimal risk assessment capabilities.

ANALYSIS APPROACH

- Two non-segmented credit score models were developed using two million anonymized consumer credit files randomly selected from one of the national credit bureau databases. One million consumer files were used for development and one million files were used as the hold-out validation sample. Credit files from the 2013-2015 timeframe were used.
  - The first model, All_Data, was built using the consumer's entire credit file *including* the NCAP-related data categories, tax liens, civil judgments, medical and non-medical agency collections. Only unpaid agency collections (medical and non-medical) were considered in this model. The model development process followed typical conventions used to develop commercially available credit score models.
  - A second model, Credit_Data, was built using the same credit file data with the identical development process. However, in this second model, NCAP-related data were excluded from consideration. Specifically, a worst-case scenario was assumed in which *all* judgments, liens and agency medical collections were excluded.
- Several analyses were conducted:
  - **Analysis 1:** Predictive performance (Gini) for the two models was determined for the U.S. population and for key subpopulations in which NCAP-related information was included in the consumers' file.
  - **Analysis 2:** Models were compared to identify compensatory behaviors that offset performance insights attributable to NCAP-related behaviors.
  - **Analysis 3:** Risk and score changes were identified for consumers with and without NCAP-related information on their credit files.
  - **Analysis 4:** Using a specific score cut-off, changes in the approve/decline volumes using the Credit_Data model compared with the All_Data model were determined. The underlying credit behaviors for the newly approved consumers were identified.

Analysis 1: How well do the All_Data and Credit_Data models predict credit risk?

Model Development
A logistic regression methodology was used to build the scoring models using VantageScore 3.0 leveled attributes. Attributes were built using the bureau credit file data, accounts, inquiries, public records and collections. A total of 900 attributes were considered for each model. All attributes met compliance and regulatory guidelines. Performance was assessed over a 24-month period where accounts were classified as 'good' if they had no delinquency greater than 30 days late. Accounts with delinquency of 90 days late or more were defined as 'bad'. Each model incorporated up to thirty attributes. The models were validated using a Gini statistic on a hold-out sample of one million consumers, representing the U.S. population, and also on sub-populations where NCAP-related behaviors were originally present.*

Predictive performance for both models was extremely strong. In fact, performance for both models was equivalent to VantageScore 3.0 performance (Figure 1).

For consumer sub-populations with public records (tax liens, civil judgments) or unpaid medical collections on their files, performance of both models was essentially equivalent (Figure 2). For several subpopulations, the Credit_Data model marginally outperformed the All_Data model. While the performance difference may not be substantial, this result suggests that using models built solely on information *directly* related to credit accounts, i.e., mortgage, auto, installment, credit cards, etc., provides a clearer signal of likelihood to default than information of a secondary nature, such as agency medical collections, where the obligation to pay may actually be the responsibility of the insurance company.

Analysis 2: What are the key attributes in the models and which new attributes in the Credit_Data model compensate for the loss of NCAP-related attributes?

As with typical commercially-available consumer credit scoring models, payment history and utilization-related information provide the greatest predictive insight (Figure 3: Top 5 Predictors). Additionally, the next 22 attributes are identical in both models. Of the 30 attributes in the models, 27 attributes were common to both models. While attributes were common, their coefficients within the respective models vary,

* *For a credit score, the gini coefficient compares the distribution of defaulting consumers with the distribution of non-defaulting consumers across the credit score range. The coefficient has a value of 0 to 100. A value of 0 indicates that defaulting consumers are equally distributed across the entire credit score range, in other words, the credit score fails to assign more defaulting consumers to lower credit scores. A coefficient value of 100 indicates that the credit score has successfully assigned all defaulting consumers to the lowest score possible. A gini coefficient above 45 is a good result.*

Figure 1: Predictive performance on U.S. population

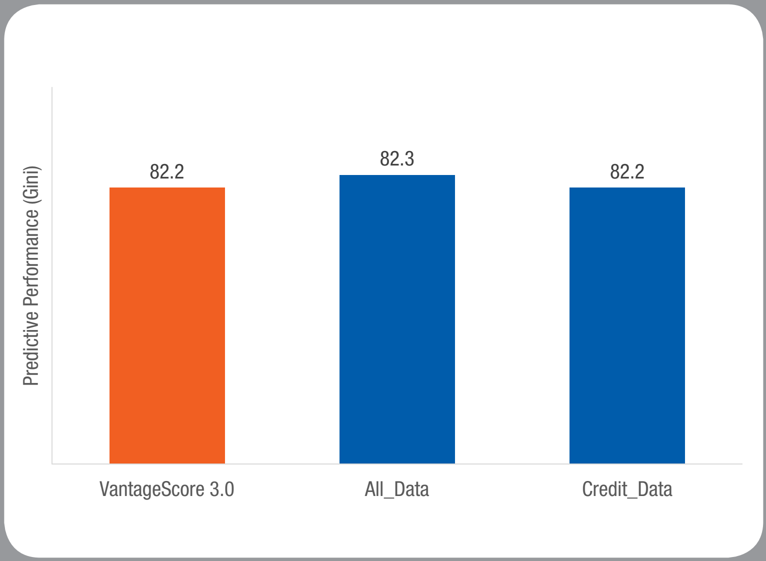


Figure 2: Predictive performance for consumers with public records and / or medical collections

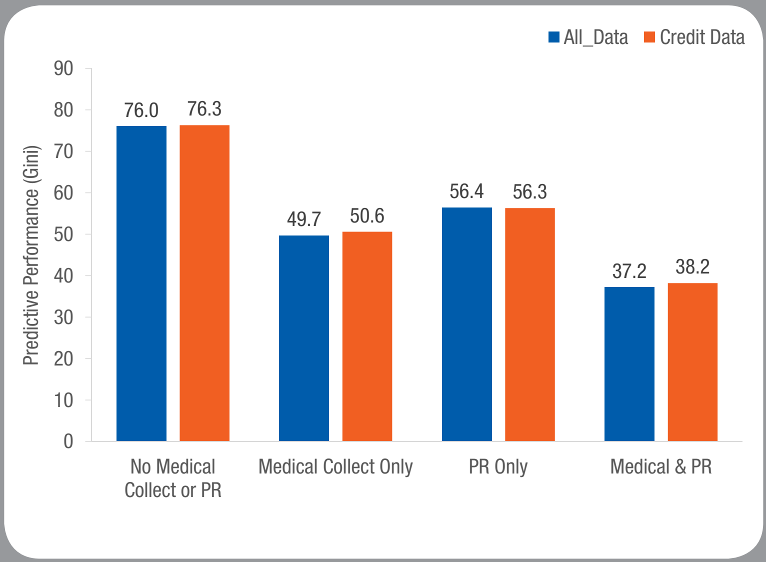


Figure 3: Key attributes in the All_Data and Credit_Data scoring models

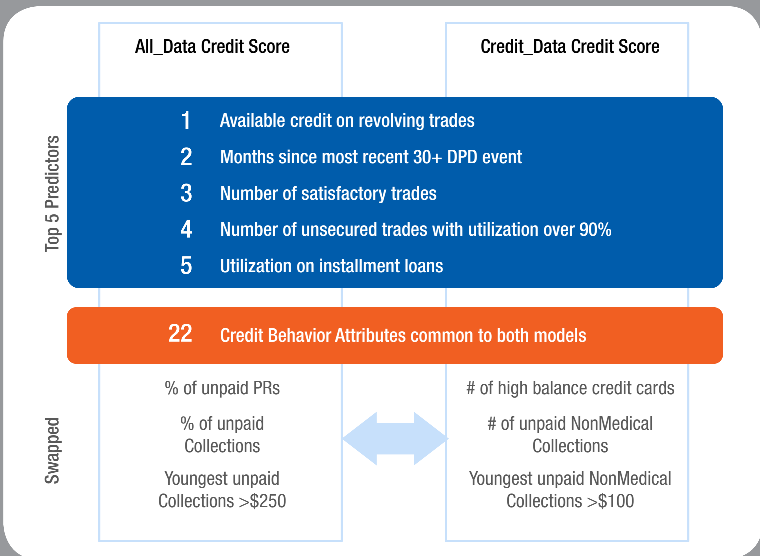
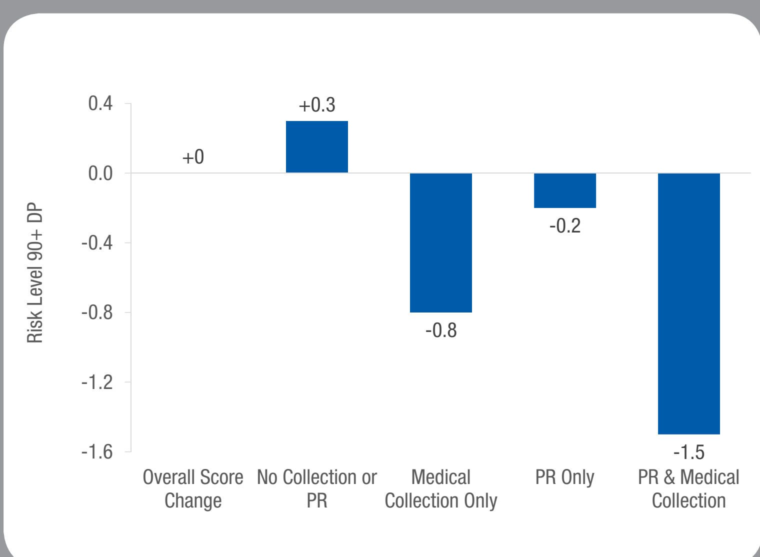


Figure 4: Risk change for consumers when scored under Credit_Data model



reflecting the adjusted importance, or weight, given to attributes in the Credit_Data model when NCAP-related attributes are unavailable.

Only three attributes differed between the two models and only one of these was substantially different in information content. Specifically one *'% of unpaid public records'* was replaced with *'number of high balance credit cards'* in the Credit_Data model. The remaining two attributes essentially substituted any unpaid collection-related information, medical and non-medical, for *only non-medical* unpaid collections information (Figure 3: see "Swapped").

While contributing some predictive value, NCAP-related attributes were not dominant drivers of predictive insight for credit score models. Furthermore, predictive performance was easily recovered using non-medical collections accounts and information related to high balance credit cards. Non-medical unpaid collection attributes in the Credit_Data model may represent a logical proxy for unpaid medical and non-medical collection attributes in the All_Data model. Similarly, the substitution of *'% of unpaid Public Records'* in the All_Data model with *'number of high balance credit cards'* in the Credit_Data model may provide greater ongoing predictive value in that the attribute reflects a *forward-looking* view of potential credit exposure rather than *historic* public record transactions. Certainly these findings suggest that greater scrutiny on predictive and stability dimensions of "future-oriented" versus "historically-derived" attributes is warranted.

Analysis 3: How do risk and consumer scores change using the Credit_Data only model as compared with the All_Data models?

Consumers were scored using both models and the resulting risk and score changes were compared.

On average, consumers with no collections or public records scored three points lower using the Credit_Data model than when scored using the All_Data model, representing an approximate 0.3 percent higher risk. Consumers who had medically-related unpaid collections on their credit files scored roughly eight points higher, indicating an approximately 0.8 percent lower risk (Figure 4).

As might be expected, those average score shifts are relatively small, in line with the fact that 90 percent of the model attributes were common to both models and that the same key five behaviors drove the primary predictive insight.

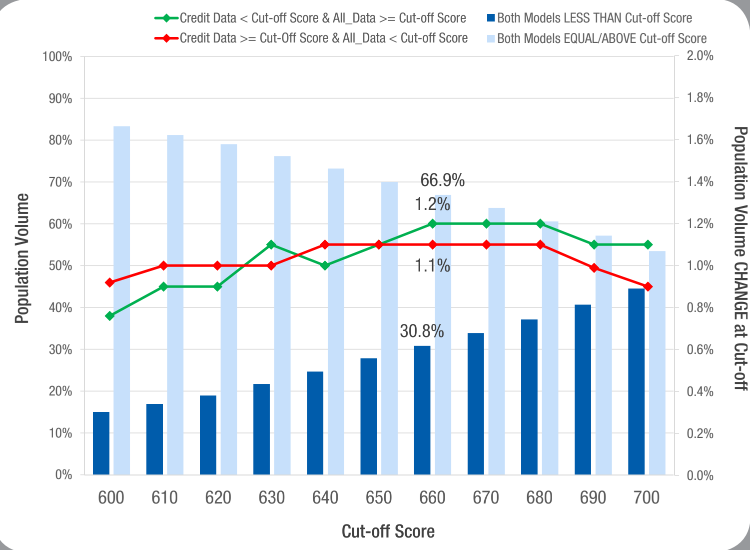
Equivalent predictive performance does not necessarily equate to consumers receiving the same scores from both models given they are scored on a slightly different attribute set with different weights in the models. It's reasonable to expect that consumers' scores will change, albeit only meaningfully for those consumers with collections and/or public records data on their files.

The consumer score migration table below (Figure 5) presents score migrations within the 600 to 750 score range when consumers are scored using Credit_Data model as compared with their original score from the All_Data model. Approximately 62 percent of consumers scored within the same risk tier using either model. An average of 23 percent of consumers receive score increases in the range of 15 to 30 points and an average of 14 percent receive score decreases of 15 to 30 points.

Figure 5: Consumer score and risk migration from the All_Data model to the Credit_Data model

Score Shift Using Credit_Data model				
Score band (All_Data)	Risk INCREASES by 15-30 points	Same Risk Tier	Risk REDUCES by 15-30 points	Risk REDUCES by more than 30 points
759-763	25.9%	65.9%	8.1%	0.1%
737-749	29.6%	57.3%	12.8%	0.2%
724-736	31.3%	51.4%	16.8%	0.5%
710-723	23.5%	57.2%	13.1%	6.2%
695-709	20.9%	66.0%	11.2%	1.8%
681-694	22.1%	64.7%	12.4%	0.8%
667-680	23.1%	61.7%	14.8%	0.4%
654-666	21.6%	61.9%	15.9%	0.6%
641-653	20.1%	63.0%	16.3%	0.6%
624-640	18.8%	65.1%	15.8%	0.3%
602-623	15.0%	67.3%	17.4%	0.2%

Figure 6: Swapset population at score cut-offs



Analysis 4: Assuming a strategy score cut-off of 660, what is the shift in the percentage of consumers who are now approved and declined using the Credit_Data model? What are the underlying credit behaviors for the newly approved consumers?

Using the example of a score-cut off of 660 in a strategy (Figure 6), 66.9 percent of consumers would be approved using either model. 30.8 percent of consumers would be declined using either model. 1.2 percent of consumers would be approved using the All_Data model but fail the score cut-off when scored by the Credit_Data model (Swap-out population). Finally, 1.1 percent of consumers failed the score cut-off when scored by the All_Data model but would be approved under the Credit_Data model (Swap-in population).

CONSUMER BEHAVIOR PROFILES

Beyond the condition of acceptable risk, do these newly approved consumers demonstrate attractive credit management practices to lenders?

Comparing the credit management practices of the Swap-in population compared to the Swap-out population, we observe the following:

- These consumers generally have a smaller overall credit footprint (Figure 7). Their account, balance and utilization mix reflects a concentration in mortgage loans, which are typically consumer's primary investment asset (Figure 7, 9 and 10) – reflecting a more stable credit profile.
- These consumers exhibit marginally lower delinquency levels on all accounts, representing higher quality credit management skills (Figure 8).
- These credit management practices typically represent lower risk behaviors in credit score models, suggesting that these consumers had failed the score-cut using the All_Data model off because of the NCAP-related behaviors.

In combination then, these newly approved consumers certainly represent an attractive population for lenders given the product and terms.

Figure 7: Average number of active accounts

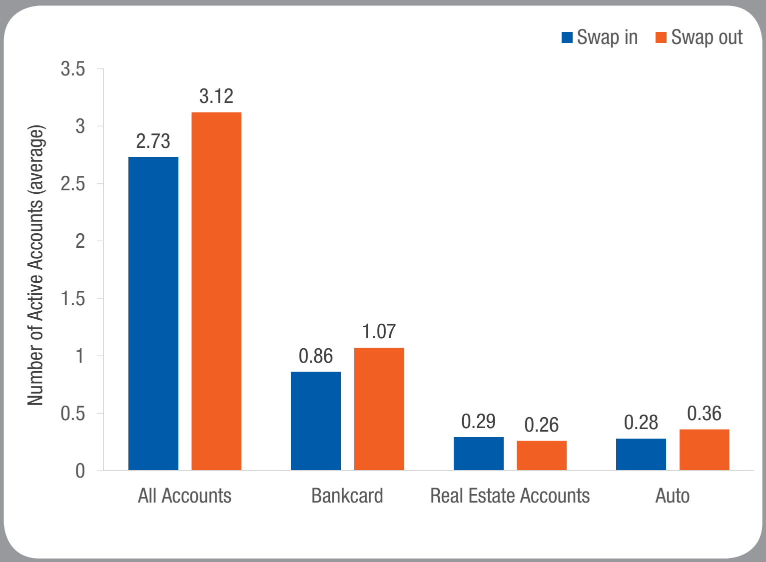


Figure 8: Average number of 30+ DPD accounts

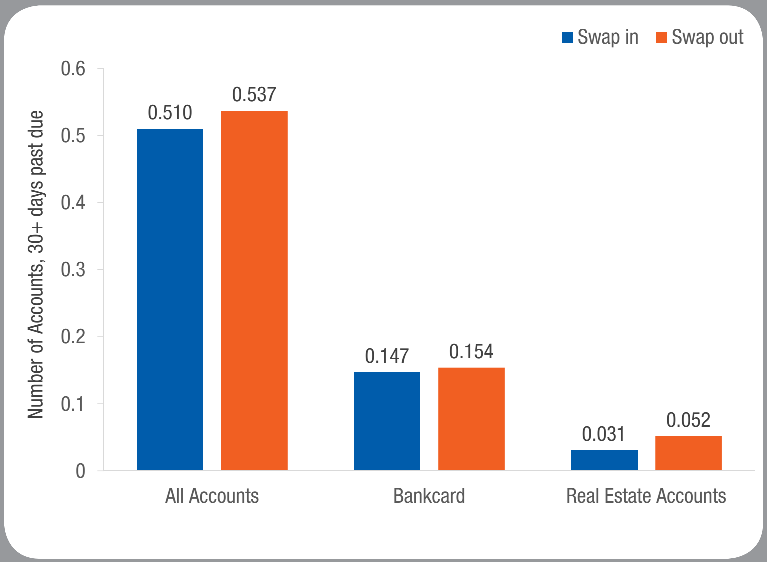


Figure 9: Average total balance ($)

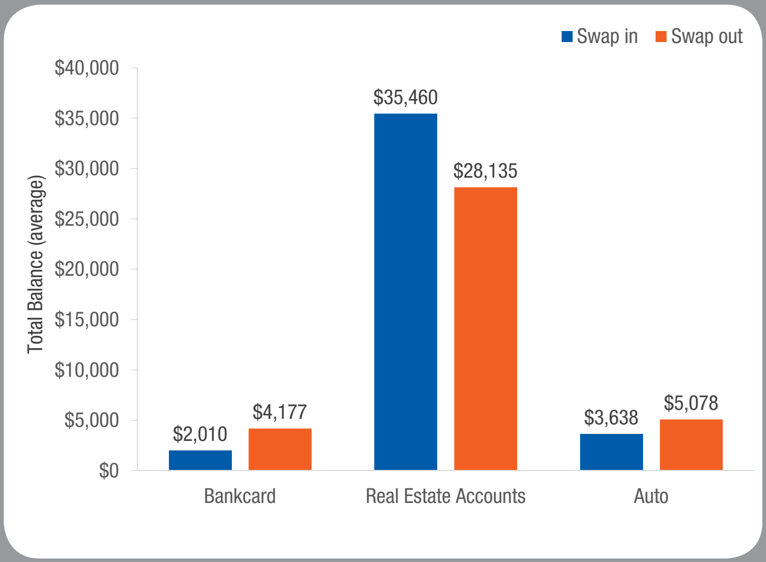
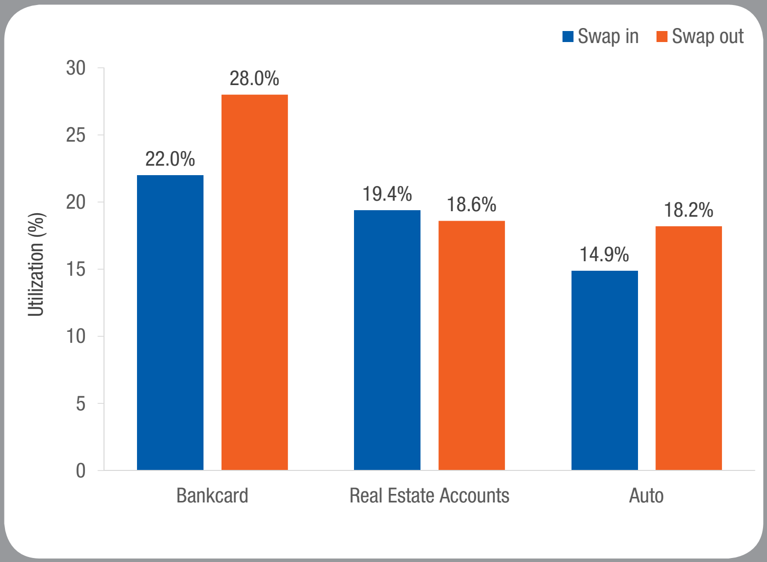


Figure 10: Average utilization (%)



CONCLUSION

The purpose of this paper was to explore the implications for credit score model quality from the removal of tax liens, judgments and agency medical collections. Rebuilding a scoring model without these data clearly shows that other behaviors can be used to compensate for the loss in predictive insight obtained from these data. In fact, models without these data may provide a cleaner risk signal given their primary and future-oriented relationship to consumer behavior.

Consumer scores should be expected to change under a new scoring model that addresses NCAP data exclusions. Typically, scores will improve or decline by 15-30 points. Given these score changes, some new consumers now will benefit from passing a lender's score cut-off. These newly-approved consumers demonstrate credit management qualities that are appealing to mainstream lenders from a risk-, asset stability- and capacity to expand-perspective.

This analysis has been conducted using the VantageScore credit scoring models and the leveled VantageScore 3.0 attributes. Lenders should evaluate how their incumbent models perform with and without NCAP data to determine whether there is detrimental impact on predictive performance and approved populations.

The VantageScore credit score models are sold and marketed only through individual licensing arrangements with the three major credit reporting companies (CRCs): Equifax, Experian and TransUnion. Lenders and other commercial entities interested in learning more about the VantageScore credit score models, including the VantageScore 3.0 credit score model, may contact one of the following CRCs listed for additional assistance:

EQUIFAX®

Call 1-888-202-4025

http://VantageScore.com/Equifax

experian™

Call 1-888-414-4025

http://VantageScore.com/Experian

TransUnion

Call 1-866-922-2100

http://VantageScore.com/TransUnion

VantageScore
March 2017

www.vantagescore.com