­

Acquisition Decision  
Support System (ADSS)

**Methodology Manual for Individual Accounts**

**Last updated Sep. 9, 2016**



# Document history

Table .1: History of material revisions to this document

|  |  |  |  |
| --- | --- | --- | --- |
| **Document version** | **Changes by** | **Description of changes** | **Date of changes** |
| 1 | Ellen Ramachandran, Anjali Dewan, Hao Zhou | Wrote to match new Decision Science document architecture, integrating content from multiple historical documents (e.g., CPS Operating Manual, OPEN Operating manual, New Acquisition Policy Manual) | Sep. 2016 |

# Document contents

[1. Document history 2](#_Toc512010361)

[2. Document contents 3](#_Toc512010362)

[3. Executive Summary 5](#_Toc512010363)

[4. Introduction 7](#_Toc512010364)

[4.1. Introduction to the ADSS model 7](#_Toc512010365)

[4.2. Business problem solved by the model 8](#_Toc512010366)

[4.3. Why the model is redeveloped 10](#_Toc512010367)

[4.4. History of the model at AXP 10](#_Toc512010368)

[4.5. Models encompassed by this manual 11](#_Toc512010369)

[4.6. Relationship between Methodology Manual and Model Documents 11](#_Toc512010370)

[4.7. Intended audience 11](#_Toc512010371)

[4.8. Organization of the manual 12](#_Toc512010372)

[5. Model Theory 13](#_Toc512010373)

[5.1. Dependent variable (model output) 13](#_Toc512010374)

[5.2. Functional form 14](#_Toc512010375)

[5.3. References 23](#_Toc512010376)

[5.4. Use of reject inference 24](#_Toc512010377)

[6. Model Data 25](#_Toc512010378)

[6.1. Modeling population 25](#_Toc512010379)

[6.2. Data sources 26](#_Toc512010380)

[6.3. Creating the modeling data set and sampling 31](#_Toc512010381)

[6.4. Quality assurance 32](#_Toc512010382)

[7. Building the Model (Model Approach) 33](#_Toc512010383)

[7.1. Summary of model building steps 33](#_Toc512010384)

[7.2. Segmentation (logistic regression only) 34](#_Toc512010385)

[7.3. Variable selection (logistic regression only) 37](#_Toc512010386)

[7.4. Final model form and interpretation 40](#_Toc512010387)

[8. Testing the Model (Outcome Analysis) 44](#_Toc512010388)

[8.1. Overview 45](#_Toc512010389)

[8.2. Out-of-sample testing 47](#_Toc512010390)

[8.3. Out-of-time testing 48](#_Toc512010391)

[8.4. Validation on recent vintages 48](#_Toc512010392)

[8.5. Sensitivity analysis 49](#_Toc512010393)

[9. Assumptions and Limitations 49](#_Toc512010394)

[9.1. Key assumptions 49](#_Toc512010395)

[9.2. Key limitations 51](#_Toc512010396)

[10. Governance, policies, and controls 53](#_Toc512010397)

[10.1. Guiding standards 53](#_Toc512010398)

[10.2. Governance framework 53](#_Toc512010399)

[10.3. Independent validation 54](#_Toc512010400)

[10.4. Oversight 54](#_Toc512010401)

[11. Implementing the Model 57](#_Toc512010402)

[11.1. Implementation platform 57](#_Toc512010403)

[11.2. Steps for implementing the model in the production environment 57](#_Toc512010404)

[12. Model Use and Business Impact 58](#_Toc512010405)

[12.1. Underwriting new accounts 58](#_Toc512010406)

[12.2. Assigning initial credit lines 60](#_Toc512010407)

[12.3. Determining pricing (APR) 60](#_Toc512010408)

[12.4. Activating the Lending on Charge (LOC) feature 61](#_Toc512010409)

[12.5. Managing ongoing cardholder risk 61](#_Toc512010410)

[12.6. Business impact assessment 61](#_Toc512010411)

[13. Tracking Model Performance 64](#_Toc512010412)

[14. Key Terms 69](#_Toc512010413)

[15. Appendix: bibliography 71](#_Toc512010414)

[16. Appendix: multivariate logistic regression 72](#_Toc512010415)

[16.1. Specification 72](#_Toc512010416)

[16.2. Maximum-likelihood estimation 73](#_Toc512010417)

[16.3. Diagnostics 73](#_Toc512010418)

[17. Appendix: model performance metrics 75](#_Toc512010419)

[17.1. Measures of discrimination and accuracy 75](#_Toc512010420)

[17.2. Population stability 77](#_Toc512010421)

[18. Appendix: Using MAS to impute or transform data 78](#_Toc512010422)

[19. Appendix: Cardholder Value (CMV) calculation 79](#_Toc512010423)

[20. Appendix: Overview of card products and portfolios 79](#_Toc512010424)

[21. Appendix: AXP application fields by market 80](#_Toc512010425)

# Executive Summary

At AXP, roughly 98 percent of decisions to approve or decline consumer card applications are automated.[[1]](#footnote-1) What enables this automated decision-making is the company’s underwriting algorithm, known as the Acquisition Decision Support System, or ADSS. For each new card applicant, ADSS produces a score, varying between zero and 99.9 percent, representing the likelihood the applicant will, if approved, default on AXP debt—defined as nonpayment for 90 days past billing—in the 12 months after the application.[[2]](#footnote-2) Thus an ADSS score of 4.5 represents AXP’s estimate that the applicant carries a 4.5 percent probability of default in the next 12 months, and so on, with higher scores indicating decreased creditworthiness.

ADSS is AXP’s main tool for underwriting and pricing new cardholders. Because the accuracy of these decisions determines the future credit quality of AXP’s portfolios, the effectiveness of ADSS is integral to the financial performance of the company. Put another way, much of the risk posed by the cardholder is fixed at the time of underwriting. Once the account is approved, risk management actions such as adjusting credit lines can control losses only to a limited extent, relative to having declined the cardholder at the outset.

The data used in ADSS come from three sources: (1) the applicant’s record at credit reporting bureaus; (2) data from the application itself (e.g., income and the card product sought); and (3) the applicant’s status with AXP’s co-brand partners (e.g., premium status in a frequent-flier program).

A team of roughly 20 data scientists builds ADSS models that are customized to geographies and products in the 22 markets where AXP operates. This manual presents the methodologies common across all ADSS models for consumer and small business portfolios. (The ADSS model for corporate underwriting is addressed in a separate manual.) This manual summarizes the overarching model design; typical data employed; standards steps for calibrating and testing; and protocols for approving, implementing, and tracking ADSS models. It does not present technical elements or results for particular ADSS models; for these, consult model-specific supporting documents.

ADSS models may be estimated in one of two ways. One variant is logistic regression, a technique AXP has used for decades to generate internal credit scores. A second, newer variant uses a machine learning algorithm known as Gradient Boosting Machine (GBM), which creates a collection of decision trees in a stage-wise fashion. GBM has empirically outperformed regression in discriminating risk, particularly for groups with fewer defaulters. Which methodology is used depends on the needs of the market and AXP’s progress in creating the data and technological foundations necessary to employ GBM. In the coming years, AXP expects all ADSS models to use GBM.

As discussed further in the Introduction, ADSS differs from AXP’s other internal scoring models in four ways. First is its use—it is used primarily for card originations, and only secondarily for the day-to-day management of new accounts. Second is its shorter performance window; it predicts default only for the coming year, since after this point it is phased out in preference to models with richer internal data.

A third difference is that ADSS relies heavily on external data, principally from credit reporting bureaus. Indeed, when AXP assesses which variables most greatly influence ADSS, it finds that the most impactful predictors—often 9 in the top 10—are sourced from credit bureaus. Typically, top drivers are credit inquiries (checks into the applicant’s record by prospective lenders) as well as the applicant’s credit score and delinquency history on external accounts.

A final difference between ADSS and AXP’s other models is that ADSS uses reject inference techniques to impute future defaults for declined applicants. This is necessary to avoid selection bias in the modeling data; since declined applicants are not in the cardholder pool, AXP cannot directly observe their outcomes.

After developing an ADSS model, modelers test how well it discriminates risk and predicts default on populations outside those used to develop the model in the first place. They also compare the model’s estimates with real-world defaults in later periods to confirm that its predictive power holds for recent data. The model is then formally reviewed and approved by AXP’s Modeling Strategy Committee (MSC), risk management executives, compliance, and AXP’s independent validators. Once the ADSS model is in production, AXP tracks its performance, typically quarterly, to ensure it continues to effectively predict defaults and rank order low- and high-risk applicants.

# Introduction

## Introduction to the ADSS model

The Acquisition Decision Support System (ADSS) is an internal credit rating model used by AXP to underwrite account originations. For each new applicant for an AXP card, ADSS produces a score, varying between zero and 99.9 percent, representing the likelihood the applicant will, if approved by AXP, default in the next 12 months. (Although ADSS is used for consumer, small business, and corporate card accounts, only the first two are addressed in this manual; the corporate model is discussed in its own manual on account of considerable differences in methodology and data.)

AXP maintains multiple internal credit risk models;[[3]](#footnote-3) how ADSS differs from the others can be viewed along four dimensions. First is use: ADSS is used to underwrite prospective cardholders at the point of application, meaning its primary objective is the credit evaluation of prospects, not current cardholders.[[4]](#footnote-4) (ADSS is used to underwrite only new applicants, not current AXP cardholders seeking an additional AXP card.[[5]](#footnote-5)) Although ADSS plays a supporting role in credit management during the account’s first year, this function is of subordinate importance. Specifically, during this first year, when AXP’s internal record on performance remains thin, ADSS score serves as an independent variable in AXP’s account management models (discussed in footnote 3).

A second difference between ADSS and other AXP risk models is data: because ADSS underwrites applicants who, by design, have no current AXP accounts, the model must rely heavily on information external to AXP. Beyond the self-reported data on the account application (e.g., income) and information about the application (e.g., the channel through which the applicant applied), all ADSS data is from external sources, whether co-brand partners or, more crucially, consumer and commercial credit bureaus.

A third factor differentiating ADSS from other risk models is its predictive period: ADSS estimates whether an account will default within 12 months, whereas AXP’s other internal models use a longer window (18 months). ADSS’s forecast window is 12 months because after this time, the accuracy of ADSS as a basis for business decisions is surpassed by other risk models that use the cardholder’s repayment and spending behavior with AXP. At this time, AXP phases out use of ADSS in preference for these other models.

A fourth unique feature of ADSS is its use of reject inference techniques to model the performance of declined applicants (that is, censored data). These techniques, detailed in section 5.4, are necessary to account for bias that may arise if modelers only use approved accounts in modeling data. Reject inference is a more important consideration for ADSS than for other models because the proportion of decline decisions (roughly a half) is orders of magnitude greater.[[6]](#footnote-6)

ADSS can be viewed not as a single model but as a family of models unified in their dependent variable (probability of default over 12 months) and varying in data and statistical methods. As of mid-2016, AXP maintains 22 ADSS models for individual applicants, each customized to differences in card portfolios (e.g., consumer versus small business) or markets (e.g., U.S. versus UK). By the end of 2016, AXP expects to reduce the total number of ADSS models to 11 by clustering continental Europe and, separately, the Asia-Pacific region.[[7]](#footnote-7)

Traditionally, ADSS algorithms used the statistical technique of logistic regression; increasingly, they use a machine learning method known as Gradient Boosting Machine (GBM). As discussed further in Section 5, GBM is an artificial intelligence algorithm that learns through experience to discern the relationship between applicant characteristics and default. Which technique AXP uses depends on the needs of the market.[[8]](#footnote-8)

All ADSS models contain two components. The first component, the primary statistical model, captures the majority of the information to predict the default probability. For some applicants, ADSS includes a second component, known as a post-model adjustment. This refines the base ADSS score to account for changes in credit conditions, the macroeconomy, or significant portfolio changes (e.g., new products entering the market or new data becoming available). These refinements take the form of a series of adjustments factors based on a tabular mapping – a non-parametric transformation of the first-stage model.

## Business problem solved by the model

The primary business problem solved by ADSS is automating the decision to approve or deny AXP card applicants. This is often envisioned as the third step in the cardholder lifecycle (Table 4.1).

Table 4.1: Typical steps in cardholder lifecycle (illustrative only)[[9]](#footnote-9)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Product design** |  | **Prospect targeting** |  | **Underwriting** |  | **Managing accounts** |
| Developing and testing economically viable cards that  appeal to target borrowers. |  | Identifying and attracting new card applicants. |  | Assessing the credit risk of target borrowers and deciding whether to offer them credit and on what terms. |  | Determining which charges to approve; posting charges to cardholder accounts; sending statements; collecting and processing payments; adjusting credit line limits; mitigating fraud; and so forth. |

At the time of prospect targeting, potential cardholders are grouped into categories of advance approval (Table 4.2). Later, these categories serve as segments in the ADSS model.

Table 4.2: Categories of advance approval [[10]](#footnote-10)

|  |  |  |
| --- | --- | --- |
| Category | Population | Description |
|  |  |  |
| Non-pre-approved (NPA) | Prospective AXP cardholders | Based on the information available at the time of targeting, AXP cannot project whether it will approve the applicant. Non-pre-approved prospects may or may not receive targeted AXP marketing. |
| Pre-screened (PS) | Prospective and current AXP cardholders | The prospect passes AXP exclusion and credit criteria at the time of targeting and receives customized AXP marketing. |

Prospects may be targeted through multiple marketing channels. Below are AXP’s main marketing mechanisms and the number of non-pre-approved U.S. consumer applicants who applied through them in the first two quarters of 2015.

Table 4.3: Channels for soliciting applicants (U.S. consumer non-pre-approved prospects, 2015 Q1-Q2)

| Marketing channel | The prospective cardholder… | Total applicants | Approved applicants | Approval rate |
| --- | --- | --- | --- | --- |
| Internet | Responds to a card offer on AXP’s website, a banner ad, an ad in search engine results, through a portal associated with an AXP co-brand, or a facilitator site promoting cards from various issuers. | 1,192,989 | 468,560 | 39.3% |
| Take One | Takes an application on a store countertop or at the point-of-sale. | 298,012 | 185,667 | 62.3% |
| Inbound Telemarketing | Calls AXP to apply for a card. | 90,616 | 50,818 | 56.1% |
| Direct Mailing | Receives a card offer in the mail. | 19,539 | 5,855 | 30.0% |
| Other | Receives an email; a telephone solicitation from a salesperson; is referred by member-to-member promotion; is solicited face-to-face by a sales agent; and more. | 1,069 | 562 | 52.6% |
| Total |  | 1,602,225 | 711,462 | 44.4% |

Once a prospect applies for a card, ADSS is used as a cutoff value to determine whether to approve the application and as an input in determining how much credit to extend and on what pricing terms.[[11]](#footnote-11) The specific cutoff value varies by portfolio and is determined by weighing the risk of default against cardholders’ expected profitability (net present value) over the life of their relationship with AXP. For an example, see Table 12.2: ADSS cutoff scores for accepting U.S. consumer applications (illustrative).

The table below summarizes the underwriting process and the role of ADSS in it.

Table 4.4: High-level schematic of how new card accounts are underwritten

|  |  | Begin application |  | Capture data |  | Make approval decision |  | Communicate the decision | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Applicant |  | Prospect submits application for AXP card |  |  |  |  |  | |  |  | |
| AXP’s system |  |  |  | * AXP captures application info. in its new accounts system * AXP retrieves data from external vendors |  | **AXP calculates ADSS score and uses it as a cutoff to decide whether to approve or decline  the application [[12]](#footnote-12)** | For lending products, AXP assigns the initial line and APR | |  |  | |
| AXP’s Global Collections Administration |  |  |  |  |  | Customer care specialists conduct manual reviews, as necessary |  |  | | |
| AXP customer care center |  |  |  |  |  |  |  | |  | * AXP sends an approval or decline letter to the applicant * For approved cardholders, AXP sends a card in the mail | |

## Why the model is redeveloped

AXP has used a variant of ADSS for more than 10 years, and has completed four substantive revisions, roughly every two to three years. These revisions are undertaken because the model’s performance deteriorates with changes in product terms and marketing, consumer behavior, and economic conditions.

Re-building ADSS also provides AXP with successive opportunities to enhance segments, refine variables (for example, introducing an index of bureau inquiries), automate the process for model-building, and make use of advances in technology. For example, as an increasing proportion of card applicants are submitted online, AXP is exploring new data sources that this change makes possible (for example, variables that leverage the IP address of the applicant).

In between major re-builds of ADSS, modelers conduct intermediate enhancements yearly. A key source of enhancements is AXP’s monthly “case reviews,” where modelers investigate real cardholder defaults to mine for earlier signals of default.

## History of the model at AXP

The corresponding section in the supporting model document summarizes material changes to particular ADSS models. Two changes that cut across most ADSS models are as follows. First, ADSS models historically used logistic regression. Enabled by faster and more economical computing power, AXP is now migrating ADSS to machine learning, which is more flexible and typically outperforms regression in predictive power.

A second cross-cutting change is that AXP traditionally developed ADSS models separately for each individual market. As noted in Section 4.1, in 2016 AXP created two cluster models for continental Europe and the Asia-Pacific (APA) region.

## Models encompassed by this manual

The table below presents the models encompassed by this manual. Text shading denotes how the company’s Enterprise Model Validation Group (EMVG) categorizes the model: **orange** denotes a model of critical importance; **blue**, significant; and no shading, moderate. (No ADSS models fall into EMVG’s fourth category, low importance.)[[13]](#footnote-13)

Table 4.5: ADSS models encompassed by this document

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Geographic market** |  | **Model owner as of Sep. 2016** |  | **Supervision of model operations** |
|  |  |  |  |  |
| U.S. |  |  |  |  |
| CPS (consumer) |  | Anjali Dewan, Vice President  [anjali.d.dewan@aexp.com](mailto:anjali.d.dewan@aexp.com) |  | Hao Zhou  [hao.zhou@aexp.com](mailto:hao.zhou@aexp.com) |
| OPEN (small business) |  |  |
|  |  |  |  |  |
| International |  |  |  |  |
| Australia  Canada  Japan  Mexico  UK |  | Anjali Dewan, Vice President  [anjali.d.dewan@aexp.com](mailto:anjali.d.dewan@aexp.com) |  | Sarthak Mishra  [sarthak.mishra@aexp.com](mailto:sarthak.mishra@aexp.com) |
| Argentina  Asia Pacific[[14]](#footnote-14)  Continental Europe[[15]](#footnote-15)  India | |  |  |

## Relationship between Methodology Manual and Model Documents

This manual describes modeling decisions that are universal across AXPs suite of ADSS models. Because its focus is commonalities, it does not describe exceptions or results for particular ADSS models. Readers seeking such specific information should consult the supporting model document for the model of interest in conjunction with this methodology manual.

## Intended audience

This document is intended to meet the needs of multiple sets of readers. These include the staff of regulatory agencies, including the Federal Reserve Bank (FRB), Federal Deposit Insurance Corporation (FDIC), Office of the Comptroller of the Currency (OCC), and the Utah Department of Financial Institutions. Other stakeholders for whom the document caters are model validators providing an independent review of the model’s conceptual soundness and auditors determining compliance with governance procedures and AXP policies.

In addition, the document is intended to serve the needs of (1) AXP management seeking to understand the model in sufficient detail to review its methodology and uses; (2) current AXP technical staff who wish to comprehend modeling choices; and (3) new AXP hires learning ADSS modeling.

## Organization of the manual

This document is organized as follows. Section 5 describes the theory behind the modeling approach, justifying the choices made and providing background on regulatory requirements and the academic literature. Section 6 provides information on the data used in the model calibrations. Section 7 details steps involved in building the model. Section 8 explains how the models are tested and benchmarked after initial development. Section 9 covers key assumptions and limitations; section 10 summarizes governance, policies, and controls. Section 11 describes AXP’s organizational and technological arrangements for model implementation. Section 12 provides information about how models are used and the impact on AXP’s business. Finally, a glossary of key terms and acronyms is provided, followed by appendices on technical issues.

# Model Theory

## Dependent variable (model output)

The ADSS dependent variable is a binary estimate of whether a cardholder will default or not default on outstanding debt. Default is defined as delinquency to 90 days past billing (60 days past due) or bankruptcy, both within 12 months.[[16]](#footnote-16) AXP takes the log-odds of the dependent variable to produce the final output, the probability of default on a scale from 0 and 99.9 percent.

To obtain the dependent variable for approved accounts, modelers check if, in each of the 12 future months, the account satisfies the default criteria. If so, the default indicator is set to one; otherwise, it is set to zero. See Section 5.4 for how AXP imputes the dependent variable for declined accounts.

As discussed in section 4.2, for the U.S. consumer market, AXP also uses a “front-end” ADSS model to narrow the universe of potential prospects into a pre-screened subset. The data for this front-end model is retrieved from credit bureaus (in soft inquiries) and co-brand partners. The “back end” ADSS used for credit underwriting is an identical model, except that it includes additional information from the application.[[17]](#footnote-17)

Within ADSS, accounts are categorized as defaulters only if the defaulted balance exceeds the thresholds presented in Table 5.1. These petty balance thresholds ensure the default is material and not triggered by membership or account fees or passive charges such as recurrent subscriptions.

Table 5.1: Petty balance thresholds for the dependent variable

|  |  |  |
| --- | --- | --- |
| **Market** |  | **Petty Balance** |
|  |  |  |
| U.S. |  | 500 USD; 1000 USD for premium products |
| Australia |  | 330 AUD |
| Canada |  | 300 CAD |
| UK |  | 100 GBP |
| Mexico |  | 2200 Peso |
| Japan |  | 16500 JPY |
| Italy |  | Lending: 75 Euro  Charge: 112 Euro |
| Germany |  | 190 Euro |
| France |  | 190 Euro |
| Netherlands |  | 190 Euro |
| Sweden |  | 700 TSEK |
| Austria |  | 190 Euro |
| Spain |  | 97 Euro |
| Finland |  | 97 Euro |
| India |  | 5000 INR |
| Hong Kong |  | 775 HKD |
| Singapore |  | 400 SGD |
| Taiwan |  | *no petty balance* |
| Thailand |  | 3500 Baht |
| New Zealand |  | 166 NZD |
| Argentina |  | 400 Argentine Peso |

**Table note:** This table presents the balance thresholds below which ADSS does not recognize a default. Premium refers to AXP products such as the Platinum® and Centurion® charge cards and the Delta Reserve lending product, which carry a higher annual fee and thus a proportionately higher balance threshold to ensure fees do not trigger default.

## Functional form

In the current implementation we use a specific implementation of a machine learning algorithm based on classification trees viz. Gradient Boosted Machine (GBM) algorithm for model building and this section describes relevant details as well as key methodological choices made.

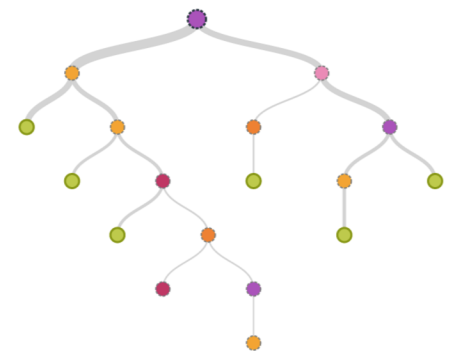
**Context**: Beginning in 2010, AXP began investing in its big data laboratory, a collection of technological investments to accommodate data of an increasing volume, velocity, and variety. These capabilities enabled the processing speeds necessary to apply machine learning algorithms for predictive analytics. These algorithms develop systems that can learn from data, rather than follow explicitly programmed instructions—an advantage that is especially attractive in for the ever-changing dynamics of fraud detection.

In 2012 and 2013, when AXP investigated alternative new methodologies for all risk modeling, On the basis of reviewing a bunch of relevant literature [1-7], it was found that gradient boosting machines (GBM) demonstrate exceptional performance in modeling binary variables.

**Gradient Boosted Machine (GBM)**

Gradient Boosted Machine (GBM) is a machine learning algorithm, where machine learning refers to algorithms that identify patterns in data without assuming a specific relationship in the functional form. Specifically, GBM builds a sequence of classification trees that iteratively reduce model error. A classification tree is a type of decision tree, in which each node or split point “tests” an observation (transaction) by some characteristic. (Illustrated in Fig 5.1)

Fig 5.1: Notional representation of a decision tree



For example, when predicting fraud for a set of transactions, a decision tree may first divide them by whether the account has an open fraud case related to it; secondly, by whether the plastic card has been recently cancelled; thirdly, whether the card member has a spike in charges at the merchant; and so forth. Each branch of the decision tree represents the result of the test; the terminus of each branch, called the leaf node, becomes the output. Continuing the example above, all transactions that satisfy all conditions on a given branch will be assigned the same probability of fraud. Whereas this example uses three split points, actual GBM results have upward of 100 split points in a single tree.

Since GBM is a machine learning algorithm, it learns from historical data the best sequence of split points, that is, the sequence that best explains variation in fraud rates such that all transactions in a leaf node are highly similar in their fraud probability.

A GBM model produces not a single tree but a collection of trees that together predict credit probabilities for a body of transactions with diverse risk characteristics. Methodologically, it builds a single tree at a time, stopping because the tree has reached the maximum number of split points (depth) according to rules pre-established by the modeler. It also may stop because each leaf node hits a constraint on the minimum number of required transactions. After building a single tree, the model builds the next tree to reduce the loss from the previous tree. It continues building new trees until it hits the upper limit number of trees established by the modeler.

The basis for GBM is the gradient descent algorithm, which is a technique to identify the local minimum of a function (here, the loss function describing the error of a prediction). Starting from any point on the function, the algorithm moves in the direction opposite to the gradient of the function in that point. The size of the movement at each step is a hyper-parameter which controls how large each step should be in the aforementioned direction.

**AXGBoost**

The point-of-sale fraud model uses the AXGBoost – built on an open source implementation of GBM algorithm, XGBoost. AXGBoost is a distributed implementation of the GBM algorithm and extension of open source XGBoost, whereas traditional GBM can’t leverage a distributed environment for building models on large data. AXGBoost enables us to run models on larger dataset under the time and computation infrastructure constraints. Also AXGBoost, extends open source XGBoost by adding specific features to support business requirements, e.g. variable importance (VI).

In the remaining part of this section, we describe mathematical details specific to AXGBoost viz. parameters, and objective function and optimization. We use XGBoost and AXGBoost interchangeably.

GBM like most machine learning algorithms are governed by parameters such as maximum depth of a tree, minimum child weight in each leaf node, maximum number of trees, etc. Additionally, AXGBoost needs other parameters. Table 5.2 lists key parameters in an AXGBoost model and their ranges tested by us.

**Table 5.2: Key parameters in an AXGBoost model**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Parameter** | **Notation** | **Definition** | **Discussion** | **Sample Values** |
| Depth of trees | D | Maximum number of split points (forks) in a branch | Too many produces overfitting to the training data; too few reduces the model’s ability to identify interactions among the data, which impairs model accuracy. | Between 5 and 9 |
| Learning rate (sometimes referred to as shrinkage) |  | How much each new tree changes the prediction of the model. | If learning rate is too high, estimates become unstable when the model produces additional trees. If it is too low, the model is inefficient and it takes too many steps to get to the stopping criterion. This means the model needs more trees for a comparably accurate output. | 10% |
| Number of trees | M | The number of trees (estimators) produced by the model. | Too few trees under fits the model to the training data; too many trees over fits it. | 500 to 2000 |
| Minimum child weight | cn | Minimum sum of instance weight (hessian hi) needed in a child | If the tree partition step results in a leaf node with the sum of instance weight less than min\_child\_weight then the building process will give up further partitioning It is a method to prevent too much specialization in the leaf nodes. A very high value can also lead to too much generalization. | 20 |
| Number of bins | b | The maximum number of cut points to look at for variable summary. | Only relevant for the approximate approach for tree building. Too many bins mean more granularity but slower model building. | 50, 100, 1000 bins |
| Lambda |  | Xg Boost regularization parameter | It helps control overfitting of the model. Higher values of lambda prevent overfitting more strictly by penalizing high weights for leaf values. | 0, 10, 100, 1000 |
| Gamma |  | Xg Boost regularization parameter | Gamma specifies the minimum loss reduction required to make a split. The splits with lower loss reduction are not favored in the trees. | Not Tested |

The tree is built optimizing the XG Boost Loss function at the tree illustrated in (4.1)

(5.1)

The summation is over the n training records, is the actual dependent (target) variable taking a value of 0 or 1 for observation and is its predicted value after t-1 trees have been built.   
  
The model begins with an initial prediction value of 0.5 for all observations while tree 1 is being built. The aim is to gradually align the predictions as close as possible with the actual labels in subsequent trees.

Each corresponds to an independent tree structure and leaf weights vector. For a given example, we will use the decision rules in the tree (given by) to classify it into the resulting leaf traversed by the observation.  
  
The additive termpenalizes for tree complexity.

(5.2)

Where is the number of leaves in tree, is the value of leaf and , are parameters.

To solve the equation we greedily determine the value of that most improves our model by minimizing the XGBoost loss measure described above in (4.1). To optimize the objective in the general setting, we approximate it using the second order Taylor expansion which leads to the equation below.

(5.3)

Where and are defined as the first and second order gradient statistics on the loss function with respect to prediction   
  
 (5.4)

In AXGBoost cross entropy has been used as the loss function for the algorithm,

(5.5)

For cross entropy the corresponding and terms can be calculated as shown below

(5.6)

Where are the label {0, 1} and prediction for an example and is the sigmoid function defined below

Defining as the instance set of leaf i.e. all the observations which pass through a tree to end in a leaf. The equation (5.3) can be simplified expanding the regularization term leading to the following generalized term as shown in equation below.

(5.7)

For a fixed structure we can compute the optimal weight of a leaf leading to a minima for equation (5.7)  
  
 (5.8)

and calculate the corresponding optimal objective function value by

(5.9)

Assuming and are the instance sets of left and right nodes after the split and , the instance set of all the examples at the parent node before the split then the loss reduction (gain) after a split is given by

(5.10)

This formula is used in practice for evaluating the split candidates. The model evaluates all the possible values existing in the available independent features greedily to obtain the best split leading to maximum gain.   
  
Post a split corresponding to the maximum gain the examples in the resulting left and right child nodes are further separately evaluated based on the above mentioned gain criteria till the entire tree structure is obtained.

The tree growing process is based on depth parameter provided by the user. Additionally during tree growth, the algorithm also checks for minimum gain criteria for a particular node. In the current implementation, only positive gains without considering are favored as the default value for minimum gain is set to 0.000001. Additionally the Minimum child weight provided by the user is also considered for stopping the tree growing. These are known as Pre-Stopping criteria. Hence the final resulting tree is a result of all the above discussed criteria being satisfied.

In AXGBoost after the tree building phase, a second post pruning stage is further considered to reduce overfitting through regularization using. Here we prune the tree nodes which have gain lower than provided by the user as an input parameter. In tree pruning, the nodes are pruned in bottom-up manner. It starts from the leaves and recursively prunes the nodes till a parent has no nodes as children with gain less than.

AXG Boost does not require external feature selection as the tree inherently captures variables as selected split points at nodes while it learns the classification simultaneously. The final model may not retain all the input variables. If certain variables do not produce optimal gain while their splits are evaluated they are dropped. This, in effect, leads to variable selection by the model. This method learns which features best contribute to the accuracy of the model while the model is being created. Also, AXGBoost uses and to introduce additional constraints into the optimization of the loss measure which further enhances the feature selection leading to a less complex model. The negative gains () are not considered for building further trees even if the maximum depth has not been reached.

Having calculated the value of the terminal leaves for the current tree, new prediction are calculated using (5.11) after updating values of as

(5.11)

Here is the prediction obtained recursively from the previous tree and is the prediction after the tree. These prediction in turn help in re-calculating the and for each observation to build the next tree and the entire process is repeated for every new tree built.   
  
Here is the shrinkage parameter or learning rate. It determines how much each new tree will change the prediction of the model. A lower value of is preferred for a stable model. If learning rate is too high, estimates become unstable when the model produces additional trees. If it is too low, the model is inefficient and it takes too many steps to get to the stopping criterion i.e. the model needs more trees for a comparably accurate output.

The final prediction for an observation is the sum of the initial prediction and the tree outputs of all the trees built.

(5.12)

Here is the initial prediction (0.5), is the learning rate or shrinkage, M is the total number of trees, and is the output of tree on observation. This output is the value of the leaf where ends as you traverse the tree beginning at the root based on the value of the independent variables for that observation.

Finally, the prediction is converted into a probability over the range 0 to 1 as follows.

P(y=1|x) = (5.13)

Some of the features of open source XGBoost have not been exposed in AXGboost for model building in the current implementation. Features like “Sparsity aware split finding” where the algorithm learns the best direction to handle missing values introduce some limitations when used for models built in American Express. For example, CPS and OPEN CDSS models have compliance requirements to always give a positive benefit to the customer in case some variables become missing. Like Customer Income for example. Today, we force missing imputations on the low risk side for these variables. This flexibility will not be there if missing is decided automatically. Other XGBoost features “Column Sub-Sampling” which prevents over-fitting has not been evaluated for exposure in existing package version.

### Methodological Choices

After describing the general theory of AXGBoost, in this section we justify two methodological choices we have made viz approximate split point determination using Weighted Quantile Approach and the choice of loss function.

### Approximate Weighted Quantile Approach

A common approach to find the best split point is the greedy approach wherein all unique points of the variable are evaluated after sorting the variable. The number of scans to find the best split point is proportional to the number of unique points in the variable. Hence, scanning all values when the data is large becomes computationally intensive especially in a single machine environment because of the memory limitations e.g. RAM (500GB).

To solve this problem (i.e. to make gbm scalable) Tianqi Chen and Carlos Guestrin, from the University of Washington, created scalable, distributed GBM library called XGBoost [4] which uses approximation to build a GBM model. AXGBoost can still use greedy approach for single machine.

In this approach, a quantile summary is defined for each variable which is representative of the actual data. For example, we can represent 1 billion data points with a 2000 points quantile summary. Instead of iterating over the full data set of 1 billion points in order to find the best split point, we iterate only over these 2000 points, making computation much faster.

Therefore, the tree building process will follow similar steps as the greedy approach. The only difference is that it finds split points from the quantile summary which comprise of a set of ‘proposed split points’. This method can be evaluated in both single machine (multiple data blocks, one processor, i.e. out of core computation) and distributed (multiple data blocks, multiple processors) setup.

In the case of distributed XGBoost, it compresses a large block of data through quantiles to minimize memory usage and ensure minimal loss of information (bounded by an epsilon-approximation) as found in the paper [4]. The quantile building process as in the paper is explained briefly as follows.

A large dataset Dbig is distributed over n machines M1, M2, M3, ... Mn for building the GBM model. Each machine has its own RAM and processor to work with. Each machine also takes a part of the dataset which we denote as D1, D2, D3, ... Dn.

In each boosting round, the machines look at the part of the dataset assigned to them to build weighted quantiles/summary for each feature. Quantiles are represented as Qk1, Qk2, Qk3, ... Qkn where Qki is the weighted quantile for kth variable in the ith machine. These quantiles contain appropriate information about the dataset so that there is very little (bounded by an epsilon-approximation) loss in data. These ‘local’ quantiles (i.e. quantile summary of the local data in the machine), are shared and combined with the help of AllReduce [6] to form a ‘global’ quantile (i.e. quantile summary of the complete data). These quantiles are a set of points called the ‘proposed split points’[4] which become the *only* points of interest for tree building.

Since the equations for finding the best split points require the sum of the residuals, the sums up to all the ‘proposed split points’ are calculated on the ‘local’ data in each machine. We then have a set of proposed split points with their corresponding sum of residuals (gradients and hessians) on the local data, which we call a ‘histogram’ [4]. This histogram is denoted as Hi for the ith machine. The histograms from all machines are combined with the help of AllReduce [6] to form a final histogram H.

After this, each machine builds a tree with the final histogram (H) and does not look at the local data it has in its RAM. Each machine builds a copy of the same tree (since the histogram is the same) and syncs with other machines. After the end of the whole boosting process each machine has the same ensemble of trees which it calls as the final model. Then any machine can write this model in the appropriate format to disk.

In case of XGBoost, the quantile creation happens for each boosting round and it depends on the weight [4], i.e. the *second order residuals (hessian)* of each record. It is a new technique used in XGBoost which is based on a widely accepted data summarization technique called GK summary [5].

An ε-Approximate Quantile Summary is made up of quantile and is defined by Q = {S,,, }, where for all y in the input space

.

A quantile in a quantile summary consists of a value the feature takes (the set S), the weight of the quantile (the function and an approximation of the rank of this value (the functions on the data set on which the quantile summary was built. A more detailed explanation can be found in the appendix of the XGBoost paper by Tianqui Chen and Carlos Guesterin [9].

### Choice of Loss Functions

As part of the review of GBM literature an assessment of different loss functions was conducted and the pros and cons were assessed to finalize Cross Entropy as the loss function of choice.

The following evaluation criteria were considered while choosing an appropriate loss function

1. Suitable for Classification: For models with binary dependent variables (e.g., fraud and non-fraud), the loss function chosen should be appropriate for classification
2. Ability to penalize negative margin: In classification, it is desirable to have positive margin as frequently as possible and it should also penalize negative margins more compared to positive ones. The margin is defined as , where is the dependent variable, the independent variables and is the model output. Let’s consider the classification rule based on the sign of , sign [, that is an example will be classified as 1 if is positive else it will be classified as -1. Correctly classified examples will have positive margin and incorrect ones will have negative margin. For example, consider a classification problem with label +1/-1, and the model output 0.9 for a transaction with true label as +1. The margin is then 1\*0.9, a positive quantity, indicating the output is closer to the true label. Margin is similar to residual, , in regression. A loss function assigns a cost (or penalty) for inaccurate prediction (or negative margin) in a classification problem. A good loss function is expected to penalize negative margin in a robust manner.
3. Should be differentiable: XGBoost, an efficient and scalable implementation of GBM, requires both gradients and hessians to be computed, hence differentiability of objective functions are desirable for implementing scalable and robust system.

Table 5.3.2 below discusses various loss functions against the above evaluation criteria

|  |  |  |  |
| --- | --- | --- | --- |
| **Loss Function** | **Suitable for Classification** | **Ability to Penalize Negative Margin** | **Differentiable** |
| Exponential | Yes | Penalizes negative margin more heavily than positive ones.  Penalty increases exponentially with large negative margins, hence too sensitive to noisy/mislabeled data. | Differentiable |
| Binomial Deviance / Cross Entropy | Yes | Penalizes negative margin more than positive ones.  Penalty increases evenly for large negative margins | Differentiable |
| Squared Error | Not optimal for classification. Not a monotone decreasing function of increasing margin. | Penalizes positive margin, sensitive to outliers | Differentiable |
| Absolute Error | Not optimal for classification. Not a monotone decreasing function of increasing margin. | Penalizes positive margin | Not differentiable. |
| Misclassification | Yes | Gives a unit penalty for negative margins and no penalty for positive margins | Not differentiable. |
| Support vector | Yes | Penalizes negative margins more than positive ones. | Not differentiable. |

The Fig. 5.3.2 below shows a plot of margin vs loss for various loss functions:

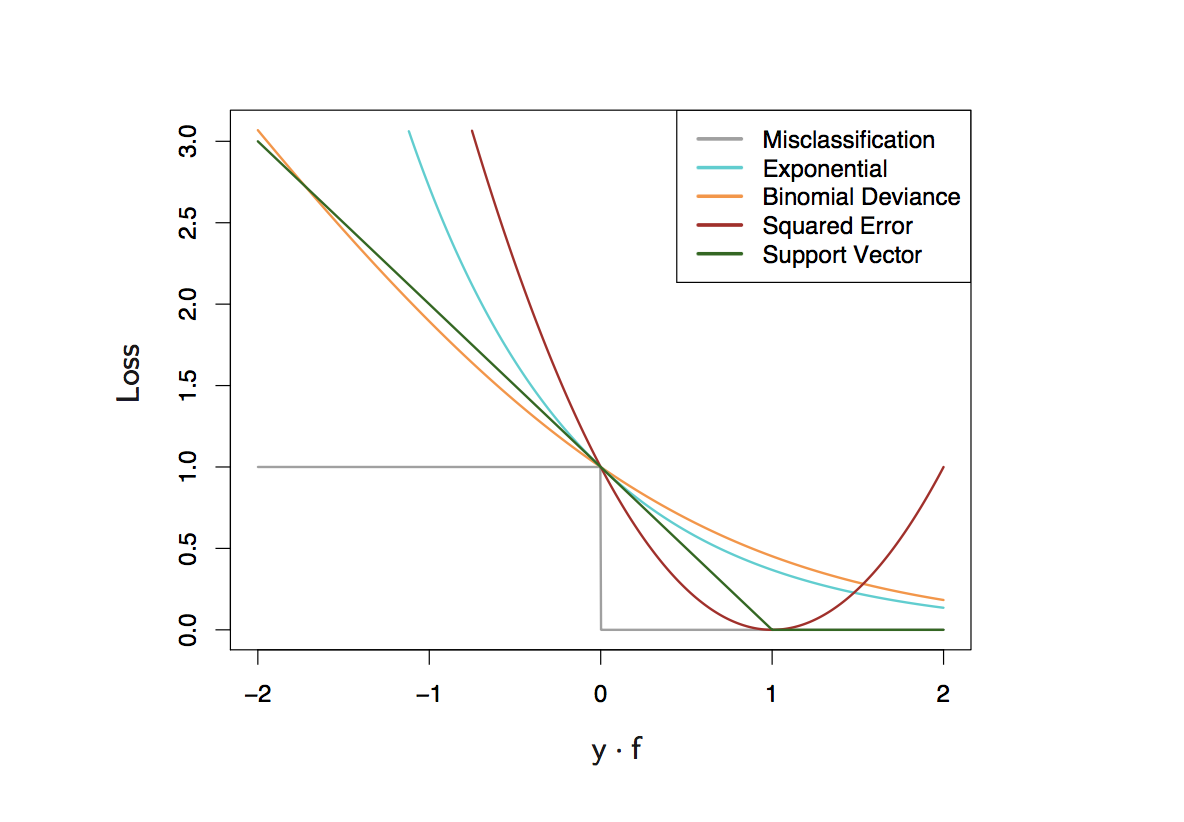


Fig. 5.3.2

Based on the table 5.3.2, Binomial Deviance and Exponential are the most suitable candidates for the loss function based on the evaluation criteria.

Further, as can be seen from Fig. 5.3.2., Binomial deviance penalizes negative margin in a robust manner without emphasizing heavily on large negative margins. However, exponential loss function penalizes negative margins more heavily than binomial deviance. In the presence of noisy data, an exponential loss function will over fit the noisy examples and create inferior models.

Binomial deviance loss function is also differentiable and suitable for GBM where trees are fitted to the negative gradients in an additive manner. It is also a monotone decreasing function of increasing margin, which ensures that the penalty for the loss function will decrease for correctly classified examples with prediction values greater than +1 or less than -1 with respective true labels as +1 and -1.

Based on the above discussion as well as literature, binomial deviance emerges as good choice for classification task. Cross entropy (5.5) which is mathematically similar to binomial deviance but can handle labels in the range {0, 1} for the dependent variable has been selected as the preferred loss function for the AXGBoost models. The properties for binomial deviance discussed above are equally applicable for cross entropy.

### Comparison with Other Machine Learning Algorithms

When AXP was investigating the value of alternate machine learning methods, it evaluated different learning methods such as Trees, Neural Networks, k-NN kernels etc. A comprehensive comparison of the pros and cons of different learning methods across various modeling characteristics is represented in Fig. 5.3.1.

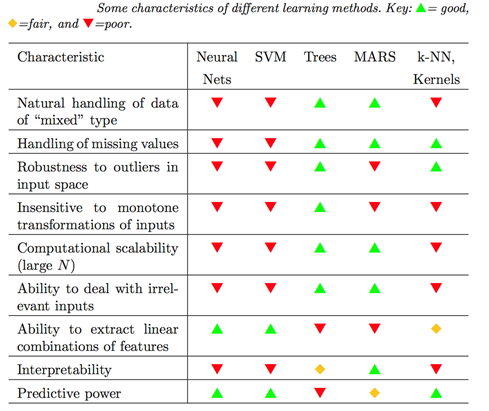


Fig. 5.3.1

As depicted in Fig. 5.3.1, the learning method ‘Trees’ have all desirable characteristics except predictive power or accuracy. Thus, ensembling/boosting of trees is required, as implemented in GBM algorithm, which helps to address this limitation of trees. Ensembling is a technique for combining many classifiers or base learners so as to obtain better predictive performance and model accuracy than could be obtained from any of the constituent classifier alone. In reference literatures the authors compare different ensemble methods, such as bagging and boosting over multiple datasets, and demonstrate that boosting performs better in most cases.

Modeling in industry applications comes with additional challenges of scalability, diverse mixture of numerical and categorical features, different scales of numerical data, missing values, etc. Given that, GBM (with trees as base learners) is considered as the ‘off-the-shelf’ algorithm for predictive modeling and data mining task, as one need not invest much time and efforts on preprocessing data, scaling inputs, tuning lots of parameters. In order to stay abreast of industry developments, a number of recent publications have been reviewed and contributed to improvements in the internal implementation of GBM [8-13].

Additionally, it has been reported in reference literature that XGBoost, a scalable version of GBM, won many machine learning competitions hosted bv Kaggle, KDDCup, Netflix prize, etc. proving that GBM and its variants are indeed ‘off-the-shelf’ algorithm for predictive modeling. Among 29 challenges winning solutions at Kaggle, considering top 3 teams, 17 solutions used XGBoost. The second most popular method was Deep Neural Networks at Kaggle. At KDDCup, XGBoost was used by every winning team in the top 10 solutions. In reference literature details of multiple notable competitions where XGBoost was the winning solution are provided. Thus, XGBoost is well established, highly recommended and recognized algorithm for predictive modeling in the industry.

AXP evaluated gradient boosting (GBM), bagging (random forest), KNN (K-nearest neighbor), and traditional regression on internal data and this exercise found GBM to be most discriminatory in predicting fraud (Table 5.3).

Table 5.: Comparison of discriminatory power (Gini) by statistical method (%)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| * Model | * Portfolio | * Logistic regression | * kNN | * Random forest | * GBM |
| * Fraud | * U.S. | * 92.3 | * 88.6 | * 94.9 | * 96.1 |

For the above evaluation, model was developed on May-Oct 2012 Data and validation was performed on Jan-Feb 2013 data.  
  
Beyond empirical performance, AXP selected GBM because of its theoretical advantages and operational convenience (Table 4.4).

**Table 5.4: Considerations in selecting the gradient boosting machine (GBM) method**

|  |  |  |
| --- | --- | --- |
| * Consideration | Theoretical | Operational/implementation |
|  |  |  |
| * GBM restricts the final outcome to between 0 and 1, with both points asymptotic. This is theoretically superior to using a simple linear regression, which may produce results beyond these bounds. | x |  |
| * Because it lacks pre-defined segments, GBM permits larger clusters of data, enabling AXP to derive patterns from a larger group of transactions. | x |  |
| * GBM allows AXP to retain correlated variables and thus extract more information from them. Suppose, for example, that when building a regression model, modelers must choose between two variables that are 60 percent correlated, since keeping both destabilizes the model. This means AXP loses the informational value from outside their overlap. GBM does not force this trade-off. | x |  |
| * Since GBM does not train coefficients like linear regression, but rather creates decision trees, it can capture non-linear trends. Thus AXP can capture changed slopes of variables or a trend reversal. | x |  |
| * GBM does not require that input variables be truncated to exclude outliers, since extremes will not bias the coefficient. Thus the model derives value from the full range of the variable. | x |  |
| * GBM does not require AXP to manually determine segmentation, since the algorithm iteratively selects optimal segmentation on its own. This simplifies model development. |  | X |
| * GBM models are straightforward to score in production for real-time decisions, since they have a short run time (computational execution). This enables timely decision-making to disrupt a fraudulent transaction. |  | X |

### Implementation and Governance

The Big Data Labs team (VP – Shourya Roy) is the owner of AXGBoost algorithm. The algorithm is implemented on the Machine Learning platform of American Express. The UAT for the distributed AXGBoost was performed by the Credit Decision Sciences team (VP - Radhakrishnan G G) in May 2017.

## References

1. Friedman, Jerome H. Greedy Function Approximation: A Gradient Boosting Machine. The Annals of Statistics 29, no. 5 (2001): 1189-232.
2. J. Friedman, T. Hastie, and R. Tibshirani. Additive logistic regression: A statistical view of boosting.
3. Hastie, Trevor J. and Tibshirani, Robert John and Friedman, Jerome H. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer Series in Statistics. New York, 2009.
4. Bagging, boosting, and C4.5, Quinlan (2006), Proceedings of the thirteenth national conference on Artificial intelligence - Volume 1
5. An Experimental Comparison of Three Methods for Constructing Ensembles of Decision Trees: Bagging, Boosting, and Randomization, Dietterich, T.G. Machine Learning (2000)
6. M. Greenwald and S. Khanna. Space-efficient online computation of quantile summaries. In Proceedings of the 2001 ACM SIGMOD International Conference on Management of Data, pages 58–66, 2001.
7. Ensemble methods in machine learning, Lecture Notes in Computer Science, Dietterich
8. R. Bekkerman. The present and the future of the kdd cup competition: an outsider’s perspective
9. Tianqi Chen and Carlos Guestrin. XGBoost: A Scalable Tree Boosting System. In 22nd SIGKDD Conference on Knowledge Discovery and Data Mining, 2016.
10. Tianqi Chen, Ignacio Cano and Tianyi Zhou. RABIT: A Reliable Allreduce and Broadcast Interface. <https://github.com/dmlc/rabit>.
11. XGBoost: A scalable Tree Boosting System, KDD, 2016
12. <https://github.com/dmlc/xgboost/blob/master/demo/README.md#machine-learning-challenge-winning-solutions>
13. <http://dmlc.cs.washington.edu/xgboost.html>

## Use of reject inference

If the ADSS data set encompassed only approved applicants, it would almost certainly be non-representative of the full applicant universe, since approved and declined applicants are likely to differ systematically in their risk characteristics. This selection bias is difficult to overcome because the data from declined applicants is censored. That is, when AXP declines an application, it is blinded to future observations about the applicant and so cannot establish a relationship between applicant characteristics and future loan performance. In an effort to construct a more representative data set, then, modelers use reject inference techniques to impute the performance of declined applicants.

### Modeling approach

Modelers use logistic regression to estimate the probability a declined applicant would have defaulted on AXP debt within 12 months if approved. This prediction is based on the applicant’s credit performance on external debts and trades. More specifically, modelers follow six steps.

1. Rank-order all applicants by actual ADSS score and identify those that are marginally inside the acceptance cutoff (for example, for the U.S. consumer charge portfolio, where the cutoff for acceptance is 18 percent, “marginally approved” applicants are those with ADSS scores between 12 and 17.99 percent).
2. Determine whether these “marginally approved” applicants defaulted on AXP debt in the 12 months after their application (i.e., assign a flag of 1 to indicate default, or 0, non-default).
3. Create a model for marginally approved applicants that uses only *external* loan performance at 12 months after the application to predict whether they default on AXP debt (the actual outcome recorded in step 2); this becomes the reject inference model.
4. Use the reject inference model to score the “true declines” to obtain their probability of default on AXP debt.
5. Assign default or non-default flags for the “true declines” such that the default rate of the set matches their average reject inference score.
6. Merge “true declines” (together with their imputed dependent variables) into the data set of applicants approved to obtain a final data set with both approved and declined applicants.

The reject inference model uses regression because the model is comparatively simple—with only three predictor variables—and so does not have the computational demands to justify machine learning.

# Model Data

## Modeling population

This section summarizes the data used for model building and the refinements the data undergoes before its use.

Because ADSS is used to approve or decline potential new cardholders, the population of interest is all applicants for AXP card products—both those approved and declined. Thus the first step in model building is identifying all card applicants for a window long enough ago to enable 12 months of forward-looking data, that is, a 12-month performance period. (By design, the performance period must match the forecast period of the dependent variable.) The selected window is typically a year in length to provide a statistically sufficient number of defaulters for each segment and to avoid the bias of seasonality.[[18]](#footnote-18) [[19]](#footnote-19) For example, for model building that occurs in 2016, the modeling window may be as follows:

Table 6.1: Data for model building (illustrative only)

|  |  |  |
| --- | --- | --- |
|  | **START OF PERFORMANCE PERIOD**  Month applicant characteristics are recorded | **END OF PERFORMANCE PERIOD**  Last month for tracking default |
| Development  data set | Jan. 2014 | Jan. 2015 |
| Feb. 2014 | Feb. 2015 |
| … | … |
| Nov. 2014 | Nov. 2015 |
| Dec. 2014 | Dec. 2015 |

The annual volume of card applications is tractable enough that AXP models on the entire population rather than a sample. A year of applicant data typically comprises roughly 4 million observations for the U.S. CPS portfolio; roughly 500,000 for OPEN; and roughly 1.4 million for international markets. After selecting the modeling window, AXP narrows this universe of card applicants to exclude those that may bias or otherwise impair the model. The table below summarizes typical exclusions.

Table 6.2: Criteria for excluding applicants from the ADSS modeling population

|  |  |
| --- | --- |
| **Exclusion criteria** | **Rationale** |
| Applicants whose applications were cancelled | The data set excludes applications that were cancelled because the applicant supplied incomplete information, even after multiple requests by AXP. Including cancelled applications would disproportionately introduce fraudulent cases, since fraudsters often discontinue applications after requests for further information. |
| Applicants with missing credit bureau information (declined or approved) | The data set excludes applicants with no consumer bureau record. These applicants may be approved for an AXP card, but will undergo manual income verification rather than automated underwriting through ADSS. |
| Applicants tagged as likely fraudulent (declined or approved) | The data set excludes applicants tagged as likely to be fraudulent, both those identified during initial underwriting and those identified after they become cardholders. They are excluded because fraud risk is estimated through separate AXP models; their inclusion would have the effect of double-counting the risk of default. |
| Approved applicants without 12 months of future AXP data | AXP may exclude applicants without 12 months of future AXP data (for example, an applicant who cancels his cards after 3 months), since AXP cannot relate characteristics during the application to future default. |
| Declined applicants   1. without past trades external to AXP or 2. without a record of a Vantage Score in the 12 months after the application[[20]](#footnote-20) | If an applicant has neither (1) past external trades, nor (2) a Vantage Score in the year after the application, AXP cannot relate the ADSS model to external credit performance. |

The corresponding section in the supporting model document quantifies the impact of these exclusions on the data set.

## Data sources

For all applicants in the final data set, AXP requires two types of information: (1) behavioral and demographic characteristics at the point of application; and (2) credit performance (i.e., repayment record) in the 12 months that follow the application. The first data source provides the basis for the independent or predictive variables; the second, the dependent variable.

### Behavioral variables used to construct independent variables

AXP obtains behavioral and demographic characteristics from three sources. From least to most important in terms of typical predictive power, they are (1) AXP’s co-brand partners; (2) the card application itself; and (3) consumer or commercial credit reporting bureaus.

Data from co-branded partners

ADSS may use the applicant’s rewards status in the loyalty program of a co-brand partner (e.g., elite status with Delta, Hilton, or Starwood). For Delta, AXP may also retrieve how recently and frequently the cardholder has flown; how frequently the cardholder flies business class; and the cardholder’s pattern of accumulating and spending frequent flier miles.

For Hilton and Starwood, AXP may retrieve the balance of membership rewards points; the history of earning and redeeming points; and the cardholder’s tenure in the rewards program.

For co-branded programs outside the U.S., e.g., with Kingfisher (India), Cathay Pacific (Hong Kong and Taiwan), KLM (Netherlands), or Singapore Airlines (Singapore), AXP may retrieve the cardholder’s program rewards tier, frequent flier number, or membership in an airline executive club.

For co-brand partners offering consumer financial services, such as Morgan Stanley or Charles Schwab,[[21]](#footnote-21) AXP may retrieve information on assets under management and debit spending.

Data from the AXP card application

ADSS also uses data from the application itself.[[22]](#footnote-22) As described in section 11, ADSS runs on a platform known as the company’s Global Decision Engine (GDE). GDE retains records from all past card applications, known as “offline tracking files.” These furnish the application data used in modeling.

For consumer cards, the AXP application generally solicits the following information:

* Name
* Home address
* Social security number (SSN) / national identification number (if available)
* Date of birth
* Phone
* Annual income
* Email (mandatory for online applications)

For small business cards, AXP solicits the same demographic details as above for the primary cardholder. In addition, it solicits business-related or “firmographic” details, as follows.

* Company name
* Company structure
* Tax identification number (if available)
* Company address
* Company phone
* Industry type
* Years in operation
* Company name on the card
* Number of employees
* Business revenue
* Estimated monthly spending

For both consumer and small business cards, the most predictive application field is typically annual income.

For a complete list of the application fields by market, see section 21.

Data from consumer or commercial credit bureaus

AXP purchases data from credit reporting bureaus on applicants’ history with other lenders, including the nature of the account (e.g., bank line or credit union); highest historical balance; revolving balance amounts; monthly minimum due; credit score; delinquency history; number and dates of credit inquiries; the presence of checking and savings accounts; and more.

For small business cards, AXP obtains both information about the applicant from consumer credit and information about the associated business from commercial credit bureaus. Commercial bureaus contain records of small business accounts with financial and non-financial creditors.

Table 6.3 summarizes the main sources of consumer and commercial bureau data across markets.

Table 6.3: Main sources of credit bureau data for ADSS

|  |  |  |
| --- | --- | --- |
| Commercial credit bureau | Dun and Bradstreet (D&B) | Small businesses’ credit history, sales, revenues, number of employees, industry, counterparty risk exposure, and more. |
| Small Business Financial Exchange (SBFE) | Aggregated data on small businesses, such as total balance, total number of trades, and total credit usage. |
| PH Bureau (France) | Small businesses’ credit score, sales, revenues, number of employees, industry, and more. |
| Consumer credit bureau | Experian (U.S.) | Cardholders’ open lines of credit, past payments, balances, inquiries from prospective lenders, delinquencies, balances, mortgages, and more. |
| TransUnion (U.S.) | Used less frequently than Experian, but provides similar data. |
| Bureau De Credito (Mexico) | Cardholders’ open lines of credit, past payments, balances, inquiries from prospective lenders, delinquencies, balances, mortgages, and more. |
| CIBIL (India) | Cardholders’ open lines of credit, past payments, balances, inquiries from prospective lenders, delinquencies, balances, mortgages, and more. |

Table 6.4 summarizes market-specific credit bureau scores used in ADSS models (alphabetically).

Table 6.4: Market-specific credit bureau score details

| **Market** | **Reporting bureau and score name** | **Score version** |
| --- | --- | --- |
|  |  |  |
| Argentina | VERAZ: FICO score | Generation 1 |
| Australia | VEDA: Relative Risk Index (RRI) | Generation 2 |
| Austria | KKE score | Generation 1 |
| Canada | TransUnion/Equifax: FICO score | Generation 1 |
| TransUnion/Equifax: HORIZON | Generation 1 |
| Finland | Asiakastieto: AT score |  |
| Germany | Schufa score |  |
| Hong Kong | TransUnion: CMS score | Generation 1 |
| India | CIBIL score | Generation 1 |
| Italy | Experian: Delphi | Generation 1 |
| CRIF1: Eurisc | Generation 1 |
| CRIF2: Eurisc | Generation 1 |
| Mexico | Buro de Credito: FICO score | Generation 2 |
| Netherlands | BKR score | Generation 1 |
| New Zealand | VEDA: Relative Risk Index | Generation 1 |
| Singapore | CBS score | Generation 2 |
| Spain | Experian: GEO score | Generation 1 |
| Sweden | UPC: Risk Value Score | Generation 1 |
| Taiwan | JCIC score | Generation 2 |
| UK | Experian: Delphi | Generation 8 |
| Experian: Consumer Indebtedness Index (CII) | Generation 2 |
| Equifax: Risk Navigator | Generation 3 |
| US | Experian, TransUnion, Equifax: FICOscore | FICO® Score 2008 |

### Categories of independent variables

The independent variables retrieved from AXP’s three data sources (partners, application, and bureaus) can be grouped into variable families or categories (Table 6.5). Viewing variables by category helps to determine whether the model has a balanced representation of variables that view the applicant from varied angles.

Table .5: Key variable categories for ADSS

| Variable category | Source | Definition |
| --- | --- | --- |
|  |  |  |
| Income | Application | Annual income from the card application. |
| Application channel | Application | How AXP received the application, for example, via the internet. See Table 4.3 for a complete list. |
| Application details | Application | Details from the card application, e.g., expatriate status or nature of bank account (checking or savings). |
| Product | Application | Nature of card sought; for example, applicant is seeking a no-fee proprietary lending card. |
| Trades | Bureaus | Number of open or closed "trades," that is, accounts reported to credit bureaus. These are typically financial debts, but may also include accounts for cable or utility subscriptions. |
| Utilization | Bureaus | Average credit usage on open or closed trades. Typically calculated as outstanding balance as a proportion of the credit line. |
| Delinquency | Bureaus | Number of accounts that have entered collections or number of months since the last reported delinquency. |
| Inquiry | Bureaus | Number of times the applicant sought to open a new account, typically in the last six months. |
| Size of wallet | Bureaus | The estimated annual discretionary income placed on plastic (a credit or charge card). |
| Tenure | Bureaus | The length of the applicant's history of accounts reported to credit bureaus (for accounts where the applicant is the primary cardholder). Sometimes calculated as average length of history on bank cards. |
| Credit score | Bureaus | A score indicating capacity to repay loans, e.g., FICO or Vantage score. |
| Open to buy | Bureaus | Sum of available (unused) credit on open trades; also calculated as maximum unused credit over a specified period. |
| Debt | Bureaus | Outstanding balances owed. Sometimes measured as debt to income ratio. |
| Credit limit | Bureaus | Average credit limit across external accounts. |
| AXP "supp" | Internal | Whether the applicant is a supplementary (secondary) cardholder on an existing AXP account. |
| Positive index | Partners | Whether the applicant exhibits a positive credit signal (for example, has a pre-existing corporate account; is quickly repaying a mortgage; or has a premium status with a partner airline). |

Each ADSS model may include many variables—as many as 70. Some categories will be represented multiple times, as illustrated in Table 6.6.

Table 6.6: Variables used in ADSS (illustrative only)

|  |  |  |
| --- | --- | --- |
| **ADSS Variable** | **Description** | **Category** |
|  |  |  |
| DALL001 | Number of open trades | Trades |
| DALL010 | Number of trades 30 days delinquent (reported within the last 6 months) | Delinquency |
| DALL024 | Number of months since the oldest trade was opened | Tenure |
| DALL033 | Number of trades opened in the last 24 months | Trades |
| DALL038 | Maximum available credit on an open trade with activity in the last 12 months | Credit limit |
| DALL043 | Number of months since the most recent 30, 60, or 90+ days-past-due delinquency | Delinquency |
| DALL044 | Average percentage of credit usage on an open trade opened in the last 24 months | Utilization |
| DALL053 | Number of months since the most recent trade was 60+ days delinquent | Delinquency |
| DALL344 | Number of accepted trades with credit usage over 75% (reported within the last 24 months) | Utilization |
| DAPPV228 | Annual income | Income |
| DFICO001 | FICO | Credit score |
| DINQV031 | High inquiry indicator (number of inquiries reported within the last 12 months ) | Inquiries |
| DMSC001 | Relationship as AXP supplementary cardholder | Supp |
| DAPPV078 | Application channel through which the cardholder applied | Application |

### Positive event variables

AXP aggregates standalone information that signals increased creditworthiness into a “positive event” index. This index is then used as an independent variable in the ADSS model.

All U.S. accounts (both consumer and OPEN) include a positive event index with the following components:

* Bank cards with high tenure: indicates high average tenure on active bank cards that are always current with low credit usage (utilization < 10%)
* Corporate linkage: indicates whether the applicant has an existing AXP corporate card
* Credit line of first bank card: indicates greater than $5,000 credit line on first bank card
* Deleveraging: indicates whether the applicant is quickly paying down a mortgage (high average speed of mortgage payment multiplied by log of mortgage balance)
* Golden eye trades: indicates the applicant has a flexible spending credit card with a competitor (a premium product with a revolving credit limit but no pre-set spending limit)
* Many current bank cards: indicates a high number of active bank cards that are always current with low credit usage (utilization < 10%)
* Membership miles: indicates presence of more than 50,000 Delta frequent flier miles
* Membership status: indicates basic or premium tier membership at Delta, Hilton, or Starwood

The U.S. OPEN ADSS model uses both the consumer positive event index, described above, as well as a commercial one with the following components:

* The application asks for five or more supplementary cardholders
* The applicant is employed in a low risk industry
* The number of employees is at least 30 *and* the business tenure is at least 10 years *or* the business ownership structure has not changed for at least 50 years
* The avg. monthly payment on commercial cards (as reported by the SBFE, defined in Table 6.3) exceeds $1,000

### Credit performance data used to construct the dependent variable

For approved applicants, information on whether the account is in default is sourced from internal databases on accounts under management. Table 6.7 summarizes.

Table 6.7: Tables used to create the dependent variable for approved applicants

|  |  |  |
| --- | --- | --- |
| **Portfolio / geography** | **Table (data set)** | **Description** |
| U.S. consumer (CPS) and small business (OPEN) accounts | IDN’s *apac\_monthly\_hist*[[23]](#footnote-23) | Data on credit performance |
| IDN’s *apac\_product\_codes* | Data on card products associated with the account |
| IDN’sFraud Application Detection System (FADS) files | Used in an intermediary step to retrieve account number (“cm11”), which is necessary to link application data (from GDE tracking files, defined in section 6.2.1) with default data (from *apac*, above) |
| International accounts | Level 2/glevel 2  files | Data on credit performance |

The dependent variable for declined applicants is imputed AXP default, created through the reject inference techniques described in section 5.4. Once calibrated, this reject inference model is used to score “true declines” and assign them a default/non-default flag based on their external credit performance 12 months after their AXP application.

The variables used to score these “true declines” are:

1. default on new external plastic cards (charge or lending cards);
2. default on pre-existing trades; and
3. Vantage score (see footnote 27 on page 23).

## Creating the modeling data set and sampling

### Identifying defaulters

**Identifying a defaulter flag for approved applicants**

If an approved applicant defaults within the 12-month performance period, the observation is marked to include it among defaulter data. If otherwise, it is flagged as a non-defaulter. These flags enable ADSS to relate applicant characteristics at the time of observation to the eventual default.

**Identifying a defaulter flag for declined applicants**

See section 5.4 for how AXP imputes the loan performance for declined applicants.

### Sampling

As noted in section 6.1, AXP models on the entire population of card applicants rather than a sample.

Modelers do employ random sampling to divide the modeling data set into in-sample and out-of-sample components, discussed further in section 6.3.4.

### Merging in independent variables

After modelers compile the master data set and flag defaulters, they merge into the data set a long list of independent variables related to creditworthiness, as discussed in section 6.2.1.

### Creating final sub-data sets

**Dividing the modeling data set**

After identifying defaulters and merging in independent variables, modelers divide the data set 70/30 at random such that 70 percent is used to develop the model (referred to as the development data set) and the remaining 30 percent is reserved to test it (referred to as the out-of-sample data set).[[24]](#footnote-24) The rationale for setting aside 30 percent as out-of-sample data is (1) ensuring each segment has a sufficient number of defaulters for statistically valid inferences; and (2) consistency with standard AXP practices for credit risk. Table 6.8 provides an example of the data used to build and test the model.

Table 6.8: Data for building and testing the model (illustrative only)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | **MODELING WINDOW**  Month observation is sampled and applicant characteristics are recorded |  | Month the performance window ends |  |
| Development data set |  | Jan. 2014 |  | Jan. 2015 | 70% of data are for developing the model  30% of data are reserved for testing the model (out-of-sample data) |
|  | Feb. 2014 |  | Feb. 2015 |
|  | … |  | … |
|  | Nov. 2014 |  | Nov. 2015 |
|  | Dec. 2014 |  | Dec. 2015 |
| Out-of-time data |  | Jan. 2016 |  |  |  |
|  | Feb. 2016 |  |  |  |
|  | Mar. 2016 |  |  |  |
| Early validation data |  | Apr. 2016 |  |  |  |
|  | … |  |  |  |
|  | Aug. 2016 |  |  |  |
| Model is put into production |  | Oct. 2016 |  |  |  |

As Table 6.8 suggests, modelers replicate the data collection for a period beyond the development data set. This becomes the out-of-time data set, which is used to observe how well the model forecasts defaults on data it has not yet seen. Finally, modelers create an early validation data set to test whether the model continues to effectively discriminate risk in more recent months.[[25]](#footnote-25)

Early-validation data is typically too recent to observe if an accounts meet the ADSS definition of default; for example, for a data snapshot 6 months back, modelers cannot know if an account would default in 12 months, since only 6 have elapsed. To surmount this challenge, modelers impute 12-month default trends based on the vintage’s early performance. For further details, see section 13.

## Quality assurance

Recognizing that the quality of input data determines the quality of ADSS scores, modelers review and reconcile internal and external data. Table 6.9 summarizes our quality assurance activities for internal data; Table 6.10, external data.

Table 6.9: Quality assurance checks for internal data

| **Quality  assurance area** | **Activity** |
| --- | --- |
| Data for the dependent variable | Checking that information on whether the applicant defaults is available for all applicants in the development data set. |
| Randomly sampling defaulter and non-defaulter data to manually check that the logic classifying applicants as defaulters functions as it is intended to. |
| Checking whether default rates are roughly consistent over time across the four data sets (the development, out-of-sample, out-of-time, and early validation data sets), or that variances are explained. |
| Data for independent variables | Performing univariate analysis, meaning a month-by-month comparison of the means, minima, maxima, number of missing values, number of zero values, and percentile distributions for each variable. Modelers conduct this analysis for all four data sets (listed immediately above). |
| Investigating irregularities such as unexplained or counterintuitive variations—and, if modelers discover any, determining their root cause and applying appropriate correctives. For example, if modelers find that the values of independent variables are unchanged across multiple months; modelers check if the data was mistakenly duplicated. |
| Code check | Checking logs of software code that retrieves internal data to search for errors. |

Table 6.10: Steps for processing independent variables from credit bureaus or other external sources

|  |  |
| --- | --- |
| **Initial processing** | **Description** |
|  |  |
| Initial processing | Confirming the data arrives in the expected format |
| Comparing the number of records and file size against expectations |
| Converting from a “.dat” to a “.sas7bdat” format |
| As necessary, clarifying with the bureau how compound variables were created |
| Checking whether the construction of variables is appropriately documented |
| Evaluation and cleansing | Confirming that the number of observations received matches the number of applicants in the sample |
| Evaluating variable means, minima, maxima, number of missing values, number of zero values, and percentile distributions to search for unexplained or counterintuitive variations |
| Looking for and removing duplicate entries |
| Comparing the rank-ordering of key risk scores with known instances of default |

# Building the Model (Model Approach)

## Summary of model building steps

Model development takes place broadly in five steps. These are summarized in Table 7.1 and elaborated in the sections that follow.

Table 7.1: Steps for building the model

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Step** |  | **Description** | **For regression** | **For GBM** |
| Data creation and sampling |  | The data is compiled as discussed in 6.3 and vetted for completeness and quality. Development and validation (testing) data sets are created. Modelers retrieve and cleanse independent variables and create the so-called “long list” of independent variables, which is a a comprehensive inventory of characteristics related to the applicant’s creditworthiness. | ✓ | ✓ |
| Segmentation |  | Modelers test alternate configurations of segments and determine which one maximizes risk discrimination. Applicants are then sorted into these segments. | ✓ |  |
| Variable selection | Single-factor analysis | Modelers narrow the long list of independent variables into a short list of more probable candidates for consideration in the final model based on the strength of their relationship with default. Modelers identify the expected relationship between each variable and default based on business intuition and correlation analysis. As needed, they transform variables so that their relationship to the dependent variable is linear. | ✓ |  |
| Missing data or zero values are imputed. | ✓ | ✓ |
| Multi-factor analysis | Modelers perform an initial regression using a stepwise procedure to identify the best subset of variables. They then combine permutations of independent variables to test their combined power and assess the contribution of each in the context of others. | ✓ |  |
| Tuning parameters |  | Modelers “tune” (adjust) parameters such as learning rate, minimum observations in leaves, tree depth, and number of trees to optimize model performance. |  | ✓ |
| Model adjustments |  | As needed, modelers adjust the base ADSS scores to reflect recent changes in the market or economy or to make use of data that changes too quickly for use in the model. | ✓ | ✓ |

Note: the table above presents the steps that are common to creating each model. Each of these steps is expanded in a separate section.

### Development environment

Models are built in SAS, a software package that retrieves and processes data and performs statistical analysis. For GBM models, modelers process data in SAS but build models in the GBM model environment in the company’s Big Data Laboratory.

## Segmentation (logistic regression only)

For logistic regression, modelers divide applicants into segments with shared risk characteristics—for example, length of history with the credit reporting bureaus. Segmenting differentiates risk, allows for changes in portfolio composition, and captures unique relationships between subpopulations and independent variables. These unique relationships arise because segments may respond to different risk drivers, or respond to the same risk drivers with different sensitivities or time lags.

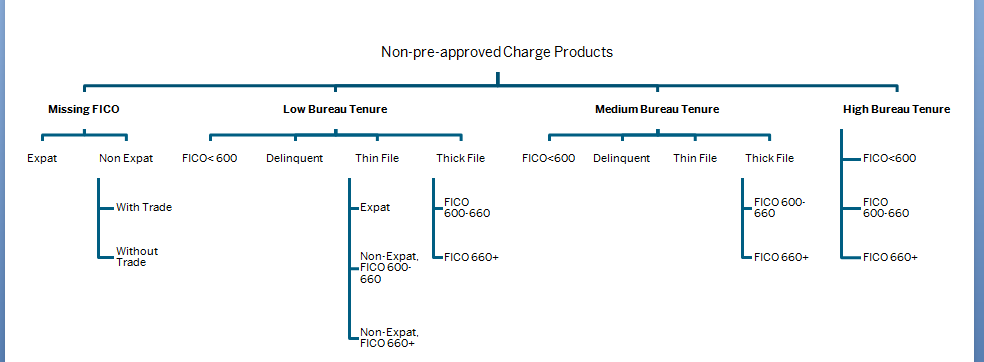
Note that modelers do not create segments for GBM models, since the GBM algorithm itself creates empirically-tested groupings that mimic segmentation.

Segments are order dependent and mutually exclusive. They are sequenced in an “if-then-else” condition, such that if an applicant is accurately described by the first segment, she is absorbed here; if not, she proceeds through each successive segment until one matches. After modelers divide the data set into segments, they perform each subsequent modeling step separately for each segment-level subsample.

### Principles for segmentation

The first level of segmentation is by the category of advance approval (non-pre-approved versus pre-screened), as discussed in Table 4.2. (Note that the pre-approved category, used only for cross-sell solicitations, does not apply for ADSS.) Applicants are then segmented by the type of product sought, since charge and lending have distinct characteristics and attract different cardholders. Below this grouping, applicants are grouped by the presence of a FICO score and the length of their credit history.

Fig. 7.1: Illustration of ADSS segmentation scheme



Note: This figure presents a typical sequence of segments (sub-populations treated as homogenous for the purpose of modeling). See the paragraphs that follow for definitions of thick- and thin-file segments.

These segments roughly fall in a continuum of most to least discriminatory, parsing increasingly fine gradations of risk. Their sequence also reflects the need to make maximum use of available data. For example, ADSScombines applicants without a FICO score to prevent their dispersion among later segments, so that these later segments are fully populated. The no-FICO segment is then subdivided between expatriates (for whom the absence of FICO most likely indicates a short U.S. credit history) and non-expatriates (where the absence has a less obvious basis). Non-expatriates are further divided according to whether they have a past trade recorded by the credit reporting bureau.

Applicants who have FICO scores are categorized by the length of their bureau history. This in effect captures the richness of available external data. Applicants may be further subdivided by FICO band; the presence of a past delinquency; and whether they are “thin file” or “thick file.” “Thin file” includes applicants who meet any one of three criteria: (1) fewer than 12 months of records at the credit bureau; (2) no plastic (card) trade accounts; or (3) fewer than three trades of any nature. “Thick” file” includes applicants who fail all of these three criteria.

To create this segmentation framework, modelers alternate segment hierarchies to determine which one maximizes risk discrimination, using t-tests to compare the difference between two provisional segments (see Section 7.2.3 below).[[26]](#footnote-26)

After modelers choose the segments, they divide applicant observations among them. Modelers then perform every subsequent model building step separately for each segment.

As noted in Table 7.1, modelers do not create segments for GBM models, since the GBM algorithm itself creates empirically-tested groupings that mimic segmentation.

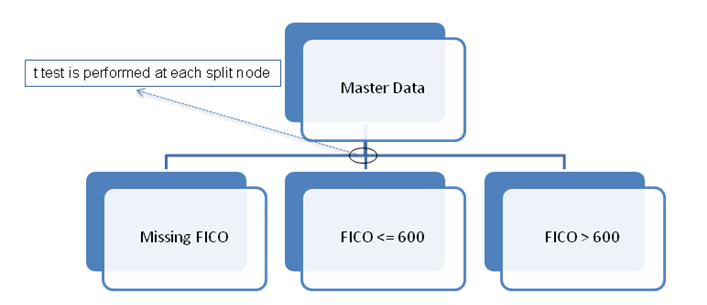
### Selected segments

In the supporting Model Document, this section identifies the final segments selected. It also provides a rationale for their selection, if this rationale is an addition or exception to those in Section 7.2.1.

### Statistical tests for segmentation

For regression models, modelers calculate Satterthwaite t-tests for proposed segments to determine whether their mean default rates differ (figure below). Modelers seek a p-value equal to or below 5 percent as evidence that the means vary (with a 95% confidence level) and the population should be split.[[27]](#footnote-27)

Fig. 7.2: Sample nodes to perform t-test



Note: This figure illustrates the process for testing whether proposed segments optimally divide one sub-population from another.

### Comparing alternative segmentation schemes

To benchmark the segmentation scheme, modelers compare it with the results of the Classification and Regression Tree (CART) algorithm, a segmentation tool in wide use throughout the financial services industry. CART is a decision learning tree technique that produces either classification or regression trees, depending on whether the dependent variable is categorical or numeric. CART first “grows” the full tree and then “prunes” it back. The pruning is done by examining the performance of the tree on a holdout data set and comparing it to the performance on a training data set. The tree is pruned until the performance is similar on both data sets. This tests the value not just of the particular binary split but of the entire configuration taken together.

AXP expects that the segmentation scheme selected should be substantively similar to that produced by CART. Significant deviations should have analytical or business support and be appropriately documented.

## Variable selection (logistic regression only)

Note: the only component of variable selection relevant for GBM is imputing missing values, if needed.

### Single factor analysis

In this step, modelers narrow down the long list of independent variables generated in data creation.

For regression, modelers submit the long list into AXP’s Modeling Automation Suite, or MAS, which makes recommendations based on criteria such as standalone significance and correlation. (See section 18 for a fuller description of MAS.) Beyond the variables recommended by MAS, modelers explore the value of variables that have a clear business relationship with the dependent variable or proven predictive power in the past. Examples are FICO, income, and inquiries from prospective lenders.

This step generates a short list of variables most likely to influence default, originating either from a MAS proposal or business judgment. This list includes representatives from all the variable categories listed in Table 6.5.

**Transforming variables**

MAS is also used to determine whether and how modelers should transform variables—that is, whether their measurement scale should be changed to increase their linear relationship to default.

**Excluding outliers**

Modelers also use MAS to test for outliers. Where they exist, modelers may apply floors (minimum allowed values) or caps (maximum allowed values) to prevent bias or exclude unreasonable results.[[28]](#footnote-28) Modelers typically review the limits proposed by MAS and then adjust them based on variable trends and business knowledge. For example, they may cap bureau tenure at 120 months, or 10 years, based on business knowledge that an even longer relationship provides no incremental informational value.

**Populating missing values**

Modelers may also populate missing values in this step, based on the risk profile of applicants, business knowledge, and past empirical findings. For cluster modeling within GBM, where some markets have values missing, modelers may apply a large negative transformation to signal an absent value to the algorithm.

### Expected variable signs

Modelers identify an *a priori* relationship between each variable and default, based on (1) business intuition, (2) findings from past models, and (3) portfolio-level correlations. For example, experience suggests that default rates rise with credit usage, so one may expect credit usage to carry a positive coefficient.

In the model-specific supporting document, these expectations are presented in the table of independent variables.

### Multi-factor analysis

**Initial regression model**

Next, modelers determine which combinations of short-listed variables together best predict default. They begin by creating a model for each segment through the SAS logistic regression procedure, using the event/trial method with stepwise selection, typically setting the p-value threshold to 0.01 percent.

Modelers then iteratively refine the model according to the following criteria.

* **Reasonableness of the coefficient sign**: One at a time, modelers remove variables whose coefficient sign contradicts *a priori* intuition, beginning with the least significant.[[29]](#footnote-29) Counterintuitive signs may indicate multi-collinearity or other interactions that may destabilize results.
* **Correlation among variables:** Modelers test for multi-collinearity and correct for it either by combining variables or by excluding the least influential one (i.e., lowest Wald Chi-square score).
* **Durable statistical significance:** After removing unreasonable or correlated variables, modelers iteratively re-estimate the model to check that the deletions do not reduce the significance of the remaining variables below the acceptance threshold (e.g., p-value > 0.001).
* **Weak contribution:** After cutting variables based on the criteria above, modelers discard variables if their removal drops the model Gini by less than 3 percent and does not eliminate a full category of variables (defined in Table 6.5). (See Section 8.1.2 for how Gini is calculated.)

Modelers then check the reasonableness of the collections of variables produced through this process—for example, checking that models include at least one variable from each variable family. This provides assurance that the model leverages the informational value of all available data sources.

Modelers follow a series of standard guidelines on acceptance thresholds. Although these vary by market and portfolio, general thresholds are as follows.

Table 7.2: General guidelines for including variables for logistic regression

|  |  |  |
| --- | --- | --- |
| Metric | Typical value | Rationale |
| P-value | <0.001 | A low p-value makes the model parsimonious and eliminates superfluous variables. |
| Variance inflation factor (VIF) | <3 | Lower VIFs ensure that the multicollinearity among variables is low, preventing over-fitting and making the model more robust. |

Note: This table presents guidelines for including variables for logistic regression.

Variance inflation factor (VIF) quantifies multicollinearity in ordinary least squares regression. If the variance inflation factor of a predictor variable were 5.27 (√5.27 = 2.3), this means that the standard error for the coefficient of that predictor variable is 2.3 times as large as it would be if that predictor variable were uncorrelated with the other predictor variables. For more details on the calculation of VIF, see Section 16.3: Diagnostics.

**Variable selection within GBM**

In a GBM approach, the algorithm selects the final modeling variables through recursive elimination, overlaid by modelers’ business intuition. Modelers begin with the “long list” of independent variables generated earlier. Next, they iteratively select final independent variables by testing their influence on the discriminatory power of the model. Specifically, modelers build an initial model with the long list of variables, and then recursively drop the variables with the lowest contribution (measuring model performance on testing data at each iteration). The optimal set of variables is that whose absence most greatly impairs model performance.

In summary, the steps in GBM variable selection are:

1. Build a model on the long list of independent variables
2. Measure model performance (Gini index) and variable contribution
3. Drop the variables with the least contribution
4. Build a next-iteration model with the new, shorter list of variables and compare its performance within the previous iteration
5. Repeat the iteration process until the drop in model performance is insignificant
6. Choose the shortest list of variables that does not impair model performance
7. Build the final model

**Addition Steps for GBM**

One additional step for GBM is selecting tuning parameters that optimize model performance. The combination of optimal parameters differs with the data size, the dependent variable event rate, the number of independent variables, and the power of information in the independent variables. Poorly chosen parameters can over-fit the model to development data and reduce model stability over time. Thus it is necessary to experiment and arrive at the right combination. Below are the guidelines for parameter tuning.

* **Learning rate:** Learning rate regularizes the predictions across trees, limiting over-fitting. Learning rate needs to be selected among the values 0.02, 0.05, 0.1, and 0.2. The lower the learning rate, the higher the rate of regularization, but more trees need to be built to obtain the best performance.
* **Minimum observation in leaves**: This parameter again guards against over-fitting. A very low number could over fit the model to development data; a very large one could impair performance through incomplete use of information as trees are shortened. Based on the size of the data, this parameter should be chosen by experimenting with the values of 50, 100, 200, and 500.
* **Depth:** GBM is generally built on trees of shallow depth; the prescribed depth is 3 to 5 for ADSS.
* **Number of trees:** The number of trees is inversely proportional to the depth of the tree in the model. At greater depths, GBM needs fewer trees to yield the same model performance. Modelers build the model with a large number of trees and then observe the point at which performance in the validation (testing) data begins to fall. This is the point at which the model begins to over-fit to development data. This can be viewed as a saturation point beyond which further trees undermine stability over time.

## Final model form and interpretation

### Model results by segment

In the model-specific supporting document, this section presents the final independent variables selected by segment (if the functional form is logistic regression) or the variables that contribute most to the model’s performance (if GBM).

### Variable contribution by category

Modelers assess the importance of each family of independent variables by testing how much variance it reduces. This analysis is complicated by the fact that many predictor variables are cross-correlated.

As a first step, modelers compute the variance reduction caused by a variable during segmentation. The variance and R-squared value of any data set are calculated as follows:

Table 7.3 illustrates how AXP quantifies variance reduction attributable to a variable (here, FICO) for a given segment. First, to obtain the variance of the data set, modelers assign each observation a score equal to the average default rate of the population, which is 0.6 percent.

Table 7.3: Quantifying variance (illustrative only)

|  |  |  |  |
| --- | --- | --- | --- |
| Observations | 544,992 | Defaulters | 3,394 |
| Default rate | 0.6% | Non-defaulters | 541,598 |
| Mean predicted score | 0.6% | Variance | 3,373 |

This yields a variance of

Modelers then measure the variance reduction attributable to each variable as follows.

Table 7.4: Quantifying the impact of FICO for a given segment (illustrative only)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | Observations in segment | 544,992 | Defaulters in   segment | 3,394 |  |  |
|  |  | Average default rate | 0.6% | Non-defaulters   in segment | 541,598 |  |  |
|  |  | Average ADSS score | 0.6% | Variance | 3,373 |  |  |
|  |  | **FICO < 650 or** | **FICO > 650 or** |
| Observations | 55,687 | Defaulters | 1,522 | Observations | 489,305 | Defaulters | 1,872 |
| Average default rate | 2.7% | Non-defaulters | 54,165 | Average default rate | 0.4% | Non-defaulters | 487,433 |
| Average ADSS score | 2.7% | Variance | 1,480 | Average ADSS  score | 0.4% | Variance | 1,865 |

Again, modelers give every observation an ADSS score equal to the average default rate of the population. In this example, the variance reduction attributable to the FICO during segmentation is [3373 – (1480+1865) =] 28.

Next, modelers calculate the variance reduction caused by the model and then distribute it among the variables that constitute the model. Modelers do this distribution through a technique called Relative Weight Analysis (RWA). RWA consists of four steps:

1. Convert original correlated predictors into orthogonal (uncorrelated) predictors using singular value decomposition.
2. Obtain the contribution of each original predictor in the orthogonal predictors through regression.
3. Obtain the contribution of each orthogonal predictor in Y through regression.
4. Use the information obtained in the first two steps to obtain the contribution of the original predictors in Y.

Following is an illustrative example. Carrying forward from the segmentation above, suppose modelers build a three-variable logistic regression model for the left node. The three variables are *FICO*, *tenure*, and *exposure* (referred to as original predictors or Xs).

In step 1, modelers use singular value decomposition to create new orthogonal (uncorrelated) variables (Zs), which are made from the original predictors (Xs). The number of Zs created equals the number of Xs. In this case, three new orthogonal variables are created.

Z1 = 0.5 \*FICO + 0.2 \* Tenure – 0.7 \* Exposure

Z2 = -2.2 \*FICO – 0.8 \* Tenure + 2.1 \* Exposure

Z3 = 0.8 \* FICO – 0.9 \* Tenure + 0.2 \* Exposure

In Step 2, modelers reverse regress each of the original variables with the new orthogonal variables. The reason is to make use of the statistical property that the coefficients of the regression equation when the independent predictors are orthogonal are the correlation between the respective independent predictors and the dependent variable. In other words, if modelers regress the Xs with Zs, the coefficients (As) of the variables will be the correlation between these variables; thus the square of the coefficients will be the variance accounted for by the Xs in the Zs.

Fig. .: Illustration of reverse regressing original and orthogonal variables (first stage)

FICO

Tenure

Exposure

Z1

Z2

Z3

**A11**

**A12**

**A13**

Original Predictors

Orthogonal Predictors

Tenure = A21 \* Z1 + A22 \* Z2 + A23 \* Z3

Exposure = A31 \* Z1 + A32 \* Z2 + A33 \* Z3

FICO = A11 \* Z1 + A12 \* Z2 + A13 \* Z3

Note: This figure illustrates the first stage of reverse regressing original and orthogonal variables in the process of Relative Weight Analysis (RWA). Note that exposure is not pertinent for ADSS; a more relevant variable is inquiries.

FICO = 0.8 \* Z1 – 0.4 \* Z2 + 0.45 \* Z3

Tenure = 0.3 \* Z1 – 0.7 \* Z2 – 0.65 \* Z3

Exposure = -0.8 \* Z1 + 0.5 \* Z2 + 0.34 \* Z3

On the basis on this equation, the correlation between FICO and Z1 is 0.8, so FICO explains 0.64 (0.82) of Z1.

Now, in step 3, modelers regress the dependent variable (Y) with the orthogonal predictors and obtain the regression equation. As in the previous step, the coefficients (Bs) represent the correlations between the Zs and Y.

Fig. .: Illustration of reverse regressing original and orthogonal variables (second stage)

Z1

Z2

Z3

**A11**

**A12**

**A13**

Y

**B1**

**B2**

**B3**

Original Predictors

Orthogonal Predictors

Dependent Variable

Y = B1 \* Z1 + B2 \* Z2 + B3 \* Z3

FICO

Tenure

Exposure

Note: This figure illustrates the second stage of reverse regression in the process of Relative Weight Analysis (RWA).

Y = 0.5 \* Z1 + 0.3 \* Z2 + 0.4 \* Z3

On the basis on this equation, modelers conclude that the correlation between Z1 and Y is 0.5, so Z1 explains 0.25 (0.52) of Y.

In the final step, modelers use the information from steps 2 and 3 to calculate the contribution of FICO in Y.

Table .5: Illustration of reverse regressing original and orthogonal variables (third stage)

Z1

Z2

Z3

Y

**B1**

**B2**

**B3**

Original Predictors

Orthogonal Predictors

FICO

**A11**

**A12**

**A13**

Note: This figure illustrates the third stage of reverse regression in the process of Relative Weight Analysis (RWA).

Extending the logic from the previous steps, the contribution of X1 in Y is given by

A112 \* B12 + A122 \* B22 + A132 \* B32

In this example, the contribution of *FICO* in Y is 0.82 \* 0.52 + (-0.4)2\*0.32 + (0.45)2\*0.42 = 0.2. Similarly, the contribution of *tenure* and *exposure* in Y is 0.1 and 0.2, respectively.

If the total variance reduction by the model is 100, then the contribution of FICO will be 40:

\*100 = 40

Modelers measure the contribution of a variable at the portfolio level by taking the ratio of the variance reduction attributable to it divided by the overall variance reduction achieved by segmentation and modeling, as follows.

where *i* is a variable. In this example, the overall contribution of FICO is as follows.

= 53%

As illustration, Table 7.5 presents the variable contributions for the legacy and new ADSS models for the U.S. consumer lending cards. This testing was conducted in Sep. 2015 when AXP assessed the value of replacing ADSS v1.1 with v2.0 for U.S. consumer cards.

Table .5: Contribution to ADSS by variable category, U.S. consumer lending cards (Sep. 2015)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Rank | | Old ADSS model  (relationship to default is in parentheses) | | % contribution | | New ADSS model  (relationship to default is in parentheses) | | % contribution | |
|  |  | |  | |  | |  | |
| 1 | | FICO (-) | | 15.3% | | FICO (-) | | 14.3% | |
| 2 | | Bureau tenure (-) | | 5.8 | | Bureau tenure (-) | | 5.3 | |
| 3 | | Inquiry index (+) | | 5.3 | | Inquiry index (+) | | 4.0 | |
| 4 | | Expatriate (-) | | 4.1 | | No. of plastic trades with >100% credit usage (+) | | 3.2 | |
| 5 | | Credit usage on plastic cards (+) | | 3.6 | | No. of trades 30 days delinquent (+) | | 3.2 | |
| 6 | | Bank cards opened in last 3 months (+) | | 3.6 | | Expatriate (-) | | 3.0 | |
| 7 | | No. of plastic trades with >100% credit usage (+) | | 3.1 | | Tenure on oldest plastic card (-) | | 3.0 | |
| 8 | | Minimum credit limit on bank cards (-) | | 2.7 | | No. of trades with positive balance (+) | | 2.7 | |
| 9 | | No. of collections (+) | | 2.7 | | Maximum open to buy (unused line) (-) | | 2.6 | |
| 10 | | Checking/savings (-) | | 2.6 | | No. of months since 60+ day delinquency (-) | | 2.6 | |

# Testing the Model (Outcome Analysis)

## Overview

After developing the model, modelers:

* Test how well it forecasts defaults on data it has not seen through
  + out-of-sample testing and
  + out-of-time testing;
* Conduct sensitivity tests to gauge how much it varies in response to changing inputs; and
* Compare its estimates against those from other models to gain confident that results approximate those from alternative approaches

The key measures for out-of-sample and out-of-time testing are below.

### Accuracy index

This index measures the agreement between the model’s output (ADSS scores) and the real default rates that AXP observes. Consider a vintage of applicants where the sum of average ADSS scores in Dec. 2014 is 7.58 percent, but where the sum of actual default rates in Dec. 2015—12 months later—is 6.38 percent (Table 8.1). Here ADSS is over-predicting default. In this scenario, the vintage’s accuracy index is 81.99 percent. (Note that accuracy index is the inverse of accuracy deviation)

Table .1: Illustrative calculation of accuracy index

|  |  |  |  |
| --- | --- | --- | --- |
| Decile of  ADSS score | Mean ADSS score  (Dec. 2014) | Mean actual default rate in early validation period  (Dec. 2015) | Divergence  (absolute value) |
|  |  |  |  |
| 1 | 6.08% | 5.24% | 0.84% |
| 2 | 0.75 | 0.71 | 0.04 |
| 3 | 0.26 | 0.24 | 0.03 |
| 4 | 0.13 | 0.09 | 0.05 |
| 5 | 0.09 | 0.04 | 0.05 |
| 6 | 0.07 | 0.02 | 0.04 |
| 7 | 0.05 | 0.02 | 0.03 |
| 8 | 0.05 | 0.01 | 0.04 |
| 9 | 0.04 | 0.00 | 0.04 |
| 10 | 0.04 | 0.00 | 0.03 |
| **Sum** | 7.58 | **6.38** | **1.20** |
|  |  |  |  |
| Accuracy deviation: **1.20** / **6.38%** = 18.81 | | | |
| Accuracy index: (1 - 18.01)= **81.19** | | | |

See section 17 for further details on accuracy index.

### Gini index

The Gini index measures how well the ADSS model rank-orders applicants by risk. A model with a perfect Gini index would order applicants in a series such that all defaulters have higher ADSS scores than all non-defaulters.

AXP expects that all new ADSS models improve Gini across all segments over the predecessor model. If not, modelers return to model building and adjust variables or variable treatments until Gini improves.

To calculate Gini, modelers first create a Lorenz curve to plot the cumulative percentage of applicants (rank-ordered from high to low by ADSS score) against the cumulative percentage of defaulters. Fig. 8.1 illustrates.

**Fig. 8.1: Illustration of Lorenz curve**



Note: This figure illustrates the Lorenz curve, the cumulative percentage of applicants (rank-ordered from bad to good by ADSS score) against the cumulative percentage of defaulters.

In Fig. 8.1, the straight line (y = x) represents a model with a random distribution—each percentile of ADSS score has an equal number of defaulters. The blue curve represents a perfect model, where one percentile of ADSS scores contains all defaulters.[[30]](#footnote-30) In practice, most models fall between these two extremes. Thus in Fig. 8.1, the tested model—the green curve—lies in the middle. The Gini of the tested model will fall to 0 percent (flatten out) as it approaches the random model; conversely, it will rise to 100 percent (bow out) as it approaches the perfect model.

Modelers calculate the Gini index by measuring the area between the tested model and the random model (straight line) and dividing by the area between the random model and the perfect model (blue line):

Gini = A/(A+B)

Thus if defaulters are evenly distributed among ADSS percentiles, the Lorenz curve hugs the 45-degree line and the Gini becomes zero [0/(0+B)]. At the other extreme, if all defaulters are in the highest ADSS percentile, Gini rises to one hundred percent [A/(A+0)]. AXP’s ADSS models often carry a Gini index of around 90 percent.

While Gini is valuable in assessing which of several techniques is preferable for a given data set, comparing across portfolios is not straightforward.[[31]](#footnote-31)

Details on modified Gini indices are presented in section 17.

### Tenure-neutral benchmarking

For both out-of-sample and out-time-time testing, modelers recalibrate the predecessor ADSS model on the development data set and compare its estimates with the new ADSS model. This is referred to as “tenure-neutral benchmarking.” This decomposes the impact of the new model from the impact of the new data—which tells AXP whether it is better to keep the older methodology in place and merely refresh the data, or whether there is indeed value in updating the methodology.

## Out-of-sample testing

Here modelers run both the new and predecessor ADSS models on the out-of-sample data withheld from the full modeling dataset. As noted earlier, this is typically a 30 percent holdout of applicants, selected at random. Comparing the new version of the model to the prior version tells modelers whether their modifications have in fact improved performance. Performance is tested by calculating accuracy index and Gini index, described immediately above.

In general, modelers conduct these tests at the overall portfolio level (e.g., consumer or OPEN); by product type (e.g., OPEN charge or OPEN lending); and by risk segment (e.g., within OPEN lending, missing FICO with a trade or missing FICO without a trade).

For out-of-sample testing, AXP generally expects the new ADSS model to outperform its predecessor in accuracy. By RIM policy, a segment must undergo post-model adjustments if (1) its accuracy index is below 50 percent and (2) it contains more than 5,000 defaulters. For segments with fewer defaulters or an accuracy index between 50 and 90 percent, the choice to adjust is at the discretion of team leadership. Where the accuracy index exceeds 80 percent (as it just does in Table 8.1) no adjustment is needed.

As an example of the output of out-of-sample testing, Table 8.2 presents results for the U.S. consumer ADSS model, calculated at the portfolio level. This testing was conducted in Sep. 2015 when AXP assessed the value of replacing ADSS v1.1 with v2.0 for U.S. consumer cards. Testing was conducted on a 30 percent holdout data set randomly selected from the larger development data set, which extended from 2012 Q3 through 2013 Q2.

Table .2: Result of out-of-sample testing (U.S. consumer ADSS model, 30% holdout sample)

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | Production model | Production model (tenure neutral) | Production to tenure-neutral ∆  (result of data change) |  | New GBM model | Production to GBM ∆  (result of methodology plus data change) |  | New logistic | Production to new logistic ∆  (result of methodology plus data change) |
| Lending | Gini index |  | 76.4% | 79.3% | + 2.9% |  | 82.9% | + 3.6% |  | 81.1% | + 4.7% |
| Accuracy |  | 61.0\* | 97.0 | + 36.0\* |  | 98.0 | + 37.0\* |  | 99.0 | 38.0\* |
|  |  |  |  |  |  |  |  |  |  |  |  |
| Charge | Gini index |  | 65.3 | 67.6 | + 2.3 |  | 71.8 | + 4.2 |  | 69.7 | + 4.4 |
|  | Accuracy |  | 87.0 | 99.0 | + 12.0 |  | 99.0 | 0.0 |  | 99.0 | 12.0 |

\*The accuracy index for the model in production is low because of a recent change in the definition of default.

## Out-of-time testing

Out-of-time testing is similar to out-of-sample testing, except that new and predecessor ADSS models are compared using the out-of-time data described in Table 6.8. Because this out-of-time data is later than the development data set, it reflects on the model’s stability over time. For example, it could indicate the new model was over-fitted for the earlier period.

In this exercise, modelers calculate Gini and accuracy index. Gini is considered more important because accuracy may be refined through post-model adjustments.

If the newer model produces a worse (lower) Gini index than the predecessor, modelers look for changes in the relationship between any independent variables and default. For example, they may find that in the development data set, the higher the delinquency index, the higher the default rate—but the relationship disappears in the out-of-time data set. In this example, modelers will remove the variable producing the instability.

As illustration, Table 8.3 provides portfolio-level results of out-of-time testing for the U.S. consumer ADSS model. This testing was conducted in Sep. 2015 when AXP assessed the value of replacing ADSS v1.1 with v2.0 for U.S. consumer cards.

Table .: Result of out-of-time testing (U.S. consumer ADSS model, data from 2012 Q1-Q2)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | Production model | Production model (tenure neutral) | Production » tenure-neutral  (result of data change) | New  model | Tenure-neutral » new (result of methodology change) | Production » new  (result of methodology plus data change) |
|  |  |  |  |  |  |  |  |
| Lending cards | Gini index | 76.5% | 79.4% | +2.9% | 81.0% | +1.6% | +4.5% |
| Accuracy index | 67\* | 94.0 | +27.0\* | 94.0 | 0.0 | +27.0\* |
|  |  |  |  |  |  |  |  |
| Charge cards | Gini index | 66.2 | 68.5 | +2.3 | 70.4 | +1.9 | +4.2 |
| Accuracy index | 85.0 | 99.0 | +14.0 | 99.0 | 0.0 | +14.0 |

## Validation on recent vintages

Because the modeling data requires 12 months of future performance, it is somewhat dated by the time the model enters production. To address this challenge, modelers test how well the model performs on a more recent vintage of applicants. Specifically, they select the most recent month which still has six months of future data (e.g., if it is Jul. 2016, they select Jan. 2016). They record the actual default rates as of this current month (Jul. 2016) and then extrapolate the rates out to 12 months (e.g., to Dec. 2016) based on the portfolio’s past performance. Finally, they calculate the accuracy index for the Jan. 2016 vintage by comparing its ADSS scores with the default rates extrapolated out to Dec. 2016.[[32]](#footnote-32)

## Sensitivity analysis

Sensitivity analyses gauge how much ADSS scores vary in response to changes in inputs. Modelers use them to check that this variance is reasonable and proportionate.[[33]](#footnote-33) Modelers gauge sensitivity through five steps (Table 8.4).

Table 8.4: Steps in sensitivity analysis

|  |  |  |
| --- | --- | --- |
| **No.** | **Step** | **Description** |
| 1 | Change variable | * Change one input variable by a defined increment (+5%, -5%, +10%, or -10%) |
| 2 | Cap and/or floor | * Apply a cap or floor the variable if it exceeds the valid range |
| 3 | Transformation | * Apply the final variable transformation |
| 4 | Run model to re-score the data | * Calculate the final model equation to obtain the new ADSS score |
| 5 | Compare the relative difference | * Compare the difference between the new and original score;  differences are typically presented as a comparison of means |

Note: This table presents standard steps for assessing the sensitivity of the ADSS model to input variables.

# Assumptions and Limitations

## Key assumptions

The list below identifies the assumptions that must hold for the ADSS model to produce valid results.

| **Modeling assumption** | **How it may impact model performance** |
| --- | --- |
| 1. **Model misspecification:** AXP assumes the regression model is properly specified and includes all key drivers and interactions. | **Category:** Methodology.  **Applicability:** Logistic regression.  **Explanation/basis:** Stepwise selection of regression structure is widely used for multi-factor modeling.  **Risk:** The model performs sub-optimally on out-of-sample and/or out-of-time data because it violates underlying assumptions or over- or under-fits.  **Development testing**: Modelers use out-of-sample and out-of-time testing to ensure the model performs acceptably. Modelers use variance inflation factor (VIF) analysis to identify multi-collinearity (see Section 16.3: Diagnostics). They benchmark against substitute approaches to test performance against alternatives.  **Future monitoring**: Modelers monitor model performance to promptly identify and rectify deterioration (see model tracking thresholds in section 13). |
| 1. **Logistic transformation:** AXP assumes the logit function is appropriate to model the linear dependence of default on explanatory variables. | **Category:** Methodology.  **Applicability:** Logistic regression.  **Explanation/basis:** Logistic regression is widely used to model categorical response. Modelers transform explanatory variables to ensure they are linearly related to default.  **Risk:** Model has sub-optimal performance.  **Development testing**: Modelers use the Modeling Automation Suite (MAS) to test the linearity assumption. They may use square root, logarithmic, or square functions to convert non-linear relationships into linear relationships.  **Future monitoring:** This technique is periodically benchmarked against alternative methodologies, e.g., machine learning techniques. |
| 1. **Overlays:** AXP assumes model overlays are appropriate for the interval between re-evaluations. | **Category:** Methodology.  **Applicability:** All (both logistic regression and GBM).  **Explanation/basis**: Modelers adjust the model after building it for multiple reasons: (1) because in the interval between the data sample and model production, the market environment may change, and adjustments allow AXP to bring performance into line with recent trends; and (2) because ongoing adjustments give AXP opportunities to make use of insights from studying newer defaults.  **Risk:** If model performance deteriorates too quickly, modelers may incur financial losses from increased defaults before modelers can implement corrective action.  **Developmental testing**: Modelers assess whether the adjustments match actual observed defaults.  **Future monitoring:** Although modelers measure accuracy and Gini only quarterly, modelers track week-over-week average default rates by portfolio, which provides an early indication of rising defaults. |
| 1. **Macroeconomy**: Modelers assume model performance is not materially affected by macroeconomic shifts. | **Category:** Data.  **Applicability:** All (both logistic regression and GBM).  **Explanation/basis:** The model purpose is grading portfolio risk for regular business conditions rather than under stress. As such, it does not explicitly model the affect of the macroeconomy.  **Risk:** Defaults could rise faster than expected in a downturn.  **Development testing:** While building the model, modelers test its performance on multiple time periods. While none were pronounced recessions, modelers confirmed that model performance (as measured in Gini) remained robust even while defaults vary. This provides confidence that the model will continue to perform well in heterogeneous conditions.  **Future monitoring**: Modelers track the model’s performance quarterly and adjust where performance falls unacceptably (see model tracking thresholds in section 13). During a downturn, modelers may make restrictive post-model adjustments to reflect heightened caution. |
| 1. **Acquisition and underwriting strategy:** AXP assumes the historical strategies for managing the business are consistent with current strategies, or that differences between them do not materially compromise the model. | **Category:** Data.  **Applicability:** All (both logistic regression and GBM).  **Explanation/basis:** To use a maximum of available historical data, modelers extract data from periods where acquisition and underwriting strategies may differ from those prevailing today.  **Risk**: Modelers calibrate models on historical data corresponding to different business practices (account terms, the use of balance transfers, underwriting standards, collection practices, etc.).  **Development testing**: Modelers test on out-of-sample and out-of-time data to ensure model outputs are stable across populations and through time.  **Future monitoring**: Modelers monitor model performance quarterly to promptly identify deterioration on existing products. |
| 1. **Product stability:** AXP assumes future product features are similar to those in the development data set. | **Category:** Data.  **Applicability:** All (both logistic regression and GBM).  **Explanation/basis:** To use a maximum of available data, modelers use card data across historical periods. These may differ in features or lending terms offered to cardholders (e.g., the introduction of no-preset spending limit on the SimplyCash product).  **Risk:** Calibration of models on irrelevant historical data.  **Development testing**: Modelers test on out-of-sample and out-of-time data to ensure model outputs are stable across populations and through time.  **Future monitoring:** Same as immediately above (we monitor model performance to promptly identify deterioration on existing products). |
| 1. **Vendor data:** AXP assumes data sourced from external vendors is reliable, consistent with the development data set, and appropriate for future use. | **Category:** Data.  **Applicability:** All (both logistic regression and GBM).  **Explanation/basis:** To use a maximum of available historical data, modelers use vendor-sourced data both historically and currently.  **Risk:** Biased model performance due to calibration on inconsistent or irrelevant historical data.  **Development testing**: Modelers test on out-of-sample and out-of-time data to ensure model outputs are stable across populations and through time. In particular, modelers have analyzed the potential impact of inconsistency in historical FICO scores.  **Future monitoring:** Same as immediately above (we monitor model performance to promptly identify deterioration on existing products). AXP maintains a contingency plan if one credit bureau is unable to furnish needed data. |
| 1. **Reject inference:**   AXP assumes the behavior of applicants who are marginally approved is substantively similar to those that are declined. | **Category:** Methodology.  **Applicability:** All (both logistic regression and GBM).  **Explanation/basis:** This assumption enables AXP to model the performance of declined applicants and introduce them into the ADSS data set, mitigating selection bias.  **Explanation/basis:** If the behavior of declined applicants differs from their imputed behavior—for example, if they have lower defaults than the marginally approved population—the resulting data set may over- or under-predict credit risk.  **Development testing**: AXP calibrates the reject inference model to establish a linkage between performance on external accounts and actual (observed) default.  **Future monitoring:** AXP monitors model performance for the marginally approved population. |

Note: This table summarizes key assumptions underlying ADSS modeling choices.

## Key limitations

Below are key limitations relating to data and the model logic. They are summarized so that AXP management, model developers, and other consumers of ADSS scores can accurately interpret the scores and to prevent use of these scores outside their intended applications. Unless otherwise noted, the limitations are relevant for both logistic regression and GBM.

| **Modeling Limitation** | **How it may impact model performance** |
| --- | --- |
| 1. **Economy and market conditions:** The model output is dependent on the economy, credit conditions, consumer behavior, and product offerings prevailing in the period it was developed. | **Category:** Data  **Explanation/basis:** Input variables may be affected by shifting macroeconomic conditions. For example, delinquent accounts may produce higher defaults in a recession than in a period of growth.  **Risk:** The model may perform sub-optimally with sudden or severe changes in the macroeconomy.  **Development testing**: Modelers test the model across different time periods to observe performance under heterogeneous conditions. With modest changes in the macroeconomy, AXP expects the model to be stable.  **Future monitoring**: Modelers monitor model performance quarterly and the ranges of input variables weekly. |
| 1. **Distribution of modeling variables:** Model accuracy may deteriorate where independent variables are highly dispersed from the mean. | **Category:** Data  **Explanation/basis:** To use a maximum of diverse historical data.  **Risk:** Model accuracy may deteriorate where independent variables are highly dispersed from the mean.  **Development testing**: Modelers check for the stability and similarity of the variable distribution across time before using in the model.  **Future monitoring**: Modelers employ a weekly tracking system known as Continuous Data Integrity Tracking (CDIT) to monitor both internal and external variables to promptly identify irregularities. |
| 1. **Business rules:** Credit management policies at the time of modeling data may differ from the rules when it enters production | **Category:** Data  **Explanation/basis:** The company’s risk management rules are based on model results from a fixed time period. After the ADSS model is complete, AXP’s risk strategy teams use its results from a test period as the foundation for business rules. If the external environment changes quickly, these rules may become outdated.  **Risk:** Therisk controls may be insufficient to control exposure.  **Development testing:** Once the model is built, the risk strategy team performs a simulation on a recent data set to align its rules with changes in the model outputs.  **Future monitoring:** Both the modeling and strategy teams perform regular tracking to ensure satisfactory model performance. |
| 1. **Vendor data:** AXP has limited visibility into how externally-sourced metrics are calculated by vendors. | **Category:** Data  **Explanation/basis**: Modelers relyon external data to obtain as complete a picture as possible of cardholder risk.  **Risk:** Externally-sourced metrics could be erroneous or incomplete. For instance, if banks do not report credit limits to the bureau, the bureau could use the highest past balance in lieu of the absent limit, inflating credit usage. This could bias model performance.  **Development testing**: Modelers check for the stability and similarity of the variable distribution across time before using in the model.  **Future monitoring**: Modelers employ a weekly tracking system known as Continuous Data Integrity Tracking (CDIT) to monitor both internal and external variables to promptly identify irregularities. |
| 1. **Variable lag:** Some independent variables reflect the financial circumstances of the cardholder with a delay (lag) of one to two months. | **Category:** Data  **Explanation/basis**: Modelers use diverse variables to maximize the angles from which modelers view cardholder risk. Some are lagged due to the operational delay of receiving vendor data and uploading it into internal AXP systems.  **Risk:** Biased model performance due to calibration on delayed data.  **Development testing**: Modelers build the model to incorporate the same lag in input data.  **Future efforts**: Modelers work continually to shorten the window between observing cardholder performance and making the data available for use in the model. |
| 1. **GBM variable selection:** The decision-tree based approach could use variables that do not have smooth trends, leading to counter-intuitive predictions. | **Category:** Methodology  **Explanation/basis:** GBM models may use a larger pool of relevant variables than historical techniques.  **Risk:** The decision-tree based approach could use variables that do not have smooth trends, leading to counter-intuitive predictions. Further,an excessively large number of variables can introduce operational risk.  **Development testing**: Modelers judgmentally assess the variable pool and only include variables that add value and demonstrate intuitive trends.  **Future monitoring:** As above, modelers monitor model performance to promptly identify deterioration (see section 13). |
| 1. **GBM sensitivity to model parameters:** GBM model parameters must be judgmentally selected. | **Category:** Methodology  **Explanation/basis:** Specifying model parameters (number of tree; tree depth; learning rate; number observations per node) is a requirement in GBM modeling.  **Risk:** Inappropriately selecting parameters could impair performance, produce over-fitting, or needlessly prolong computational time.  **Development testing:** Controls are exploratory analyses to optimize parameters and using model-specific criteria and traditional metrics of effectiveness (Gini and accuracy) to select optimal parameters.  **Future monitoring:** As above, modelers monitor model performance to promptly identify deterioration (see section 13). |
| 1. **Lack of segmentation in GBM model:** A single model is built on the entire portfolio; there is no pre-established segmentation. | **Category:** Methodology  **Explanation/basis:** GBM, as a decision-tree based approach, does not require pre-segmented data.  **Risk:** Without segments, model performance may weaken at the sub-population level.  **Development testing**: Modelers evaluate the model’s performance both overall and by key population groups.  **Future monitoring:** As above, modelers monitor model performance to promptly identify deterioration (see section 13). |
| 1. **Use of character variables in GBM:** GBM treats numeric and character variables differently. | **Category:** Methodology  **Explanation/basis:** To use maximum amount of diverse variables.  **Risk:** GBM may use character variables sub-optimally in certain cases (such as when the variable has a large number of categories).  **Development testing:** As above, modelers judgmentally assess the variable pool and only include variables that add value and demonstrate intuitive trends.  **Future monitoring:** As above, modelers monitor model performance to promptly identify deterioration (see section 13). |
| 1. **Segments with low number of defaulters (logistic):** Model segments with a low number of defaulters can lead to unstable logistic regression models. | **Category:** Methodology  **Explanation/basis:** Segmenting differentiates risk, allows for changes in portfolio composition, and captures unique relationships between subpopulations and independent variables. However, it can divide populations into sub-groups with very few defaulters, impairing statistical performance.  **Risk:** Segments with very few defaulters impair risk discrimination, which may lead to unstable models.  **Development testing:** A control is that modelers combine similar segments with very few defaulters (<1000) and create a common model for this population.  **Future monitoring:** As above, modelers monitor model performance to promptly identify deterioration (see section 13). |

Note: This table summarizes key limitations to be considered when using ADSS scores.

# Governance, policies, and controls

## Guiding standards

ADSS is created in an operational context defined by *American Express Management Policy 50: Enterprise-wide Risk Management* (January 1, 2016) and *American Express Management Policy 55: Model Governance and Validation* (July 15, 2015).

Modeling techniques and practices also embody feedback from AXP’s regulators. This feedback includes expectations communicated broadly to the industry—such as *Supervision and Regulation Letter 11-7, Supervisory Guidance on Model Risk Management* (April 4, 2011). It also includes the cumulative body of supervisory findings communicated through examinations and supervisory letters. For example, AXP changed the ADSS dependent variable for U.S. lending to 90 days past billing in 12 months in response to a 2014 finding from the U.S. Office of the Comptroller of the Currency (OCC).

The content of this document complies with the *Model-specific Supporting Documentation Standards* published by AXP’s Enterprise Model Validation Group (June 15, 2015).

## Governance framework

Modelers recognize that if ADSS outputs are incorrect or incorrectly used, AXP could face increased defaults or other adverse impacts such as brand damage or a regulatory violation. To prevent these errors or misuse, modelers follow the governance framework summarized in Table 10.1.

Table 10.1: Governance framework

|  |  |  |
| --- | --- | --- |
| **Role** | **Definition** | **Responsible party** |
| Modeling direction | Allocation of resources across modeling efforts and final decision of methodology by portfolio/market. | Decision Science’s Senior Vice President in consultation with the vice president (band 45+) who manages the specific modeling effort. |
| Model building | For a specific modeling effort, selecting appropriate data; testing variable combinations; tuning parameters; and ensuring initial model satisfies committee standards. | Decision Sciences vice president (band 45+) who manages the modeling effort. Supported by a director (band 40) who leads day-to-day operations. |
| Process controls | Ensuring controls over data processing (e.g., quality assurance checks) and integrity of statistical decision making (e.g., consistency with past models). |
| Model testing | Validating the model on out-of-sample and out-of-time datasets and assessing performance against benchmarks. |
| Internal validation | Challenge by validators external to model development. | The company’s Enterprise Model Validation Group (“EMVG”) validates Decision Science models to verify they are soundly constructed and performing as expected, consistent with their objectives and business uses. |
| Oversight | Review and challenge by executives from a variety of business lines who possess regulatory knowledge and depth of industry experience. | Responsible entities include the governing modeling committee (e.g., Modeling Strategy Committee),[[34]](#footnote-34) the Line-of-Business Compliance Officer; and the Fair Lending Program Office. |

[

## Independent validation

### Validating the Model

Once the new or redeveloped model is complete, it is validated by the company’s independent Enterprise Model Validation Group, or EMVG. EMVG’s role is to confirm that the model uses reliable data, employs sound logic, demonstrates acceptable performance, and is used in business applications as intended.[[35]](#footnote-35) See *American Express Management Policy 55: Model Governance and Validation Policy* for a description of EMVG’s validation protocol.

Validations occur not only before a new model is put into use, but also on an ongoing basis, at the discretion of EMVG, to ensure the model continues to perform as intended.[[36]](#footnote-36)

EMVG validations frequently result in findings, which fall into three groups (Table 10.2).

Table 10.2: Levels of EMVG findings

|  |  |
| --- | --- |
| **Importance level** | **Implication for Model Use** |
| High importance | The new model (or model change) cannot be implemented before EMVG’s findings are addressed. For existing models in production, a temporary fix must be implemented within 60 days of the validation report. A permanent solution must be implemented within 12 months of the validation report. |
| Medium importance | The model (or model change) can be implemented in production, but EMVG findings must be addressed within an agreed-upon timeline that cannot exceed 12 months from the validation report. |
| Observation | The model (or model change) can be implemented in production, but EMVG’s observations should be considered for future model development. |

Note: This table presents the three levels of EMVG findings and their implications for model use.

## Oversight

### Compliance and Fair Lending

In parallel with validation, AXP’s Compliance team and Fair Lending Program Office review the model for legal and regulatory concerns. This step ensures that credit decisions comply with the market’s regulations and do not discriminate on prohibited characteristics.[[37]](#footnote-37)

To receive Fair Lending and Compliance approval, the model must meet all requirements and regulations outlined in Regulation B and the Equal Credit Opportunity Act, including, but not limited to:

* Ensuring that any credit scoring system that evaluates an applicant’s creditworthiness is empirically derived and demonstrably and statistically sound.
* Ensuring that credit scoring decisions follow accepted statistical conventions and be made with the intent of preventing credit loss.
* Ensuring that any use of age is consistent with permissions outlined in Regulation B:
  + A credit risk model may be segmented into no more than two discrete segments based on the age of the applicants or the applicants being evaluated.
  + One segment must cover only a narrow range, with its high end not exceeding 35 years of age; the other segment must include all applicants or applicants above that age.
  + Each age-split segment must include attributes or variables that are predictive for the age group scored by that segment.
  + Any age-split model must meet the Regulation B standard for an “empirically derived [and] demonstrably and statistically sound credit scoring system,” including the requirement that the model be periodically revalidated statistically and adjusted to maintain predictive ability. Additionally, all age-split segments must be implemented simultaneously, not in a staged manner, as required by Regulation B.

Per FLPO guidance, U.S. models may not consider the following characteristics:[[38]](#footnote-38)

Table 10.3: Cardholder characteristics that may not be used to make credit decisions

|  |  |
| --- | --- |
| Race | Familial status |
| Age[[39]](#footnote-39) | Sexual orientation |
| Color | Receipt of public assistance |
| Religion | Receipt of income from part-time employment, retirement savings, pension, or annuity |
| National origin | Whether applicants have a telephone account in their name |
| Marital status | Age or location of dwelling |
| Gender | Good faith exercise of rights under the Federal Consumer Credit Protection Act |
| Likelihood of bearing or rearing children | Disability status |

Note: This table presents applicant characteristics that are prohibited bases for credit decisions.

To be implemented, the model must receive Compliance and Fair Lending approval at the vice president level.

### Modeling Strategy Committee

After EMVG, Compliance, and Fair Lending review the model, the model is presented to AXP’s Modeling Strategy Committee, or MSC, for review.[[40]](#footnote-40) The MSC meets roughly monthly to review new or redeveloped models, where it:

* Evaluates the model logic;
* Reviews the reliability of input data;
* Verifies compliance with consumer lending laws and other regulations;
* Ensures the model is used as intended, or, if adapted to other purposes, is appropriately tested for these purposes; and
* Reviews tracking reports and takes action if performance deteriorates below the company’s acceptance thresholds (see Table 13.1 below)

MSC membership is as follows.

**Voting members**

* Committee Chair: a band 50+ Risk and Information Management (RIM) vice president for a major business unit
* At least four vice presidents of RIM Decision Science other than the chair’s business unit
* Up to four vice president representatives at the chair’s discretion
* Vice president for Compliance
* Vice president for Fair Lending

**Non-voting members**

* Representatives from the Advance Risk Capabilities (ARC) team, Internal Audit Group (IAG), Finance, General Counsel, Global Banking Group, Global Risk Oversight, and Operational Risk, as needed

After all its concerns are addressed, the MSC officially votes to approve the model for implementation.[[41]](#footnote-41)

### Internal audit

AXP’s Internal Audit Group (IAG) provides an independent review of the company’s internal control environment, which includes a periodic review of whether model developers, validators, and oversight bodies comply with company policies and meet documentation standards.

# Implementing the Model

This section describes how the model is put into production. This occurs after all governance steps described in section 10 are complete.

## Implementation platform

As noted earlier, ADSS is implemented on the company’s the Global Decision Engine (GDE). Whenever AXP receives a new card application, GDE calculates the ADSS for the application and forwards it to the related downstream processes.

## Steps for implementing the model in the production environment

### Implementation

In this step, the ADSS model and its associated data, data treatments, and segmentation scheme are uploaded into the company’s Automated Modeling Capability (AMC), a tool for model development and implementation.[[42]](#footnote-42) AMC generates the final model code, which is implemented in the Global Decision Engine (GDE) by the Advanced Risk Capabilities (ARC) team.

### User acceptance testing (UAT)

Before placing the model into production, modelers conduct user acceptance testing (UAT), a simulation to confirm the ADSS code performs as expected under real-world conditions. Specifically, RIM’s Advanced Risk Capabilities (ARC) team runs the ADSS model code on actual data offline in a test environment. It then sends the results to RIM modelers, who compare them with results calculated via verified steps. The modeling team approves the UAT if the two samples completely match.

Once UAT is complete, the ADSS code is locked to prevent unauthorized changes.

### Post-implementation validation (PIV)

After the model is uploaded, the modeling team retrieves ADSS scores generated by GDE and tests them against authenticated results calculated offline. This step differs from the earlier User Acceptance Testing (UAT) because the tested scores are from the live stream of data produced in the company’s actual decision engine—not a test environment. The purpose is to confirm that differences in code by product and market have been accurately applied in production.[[43]](#footnote-43)

At this point, changes to the code can be initiated only through an official Risk Change Management (RCM) request approved by a modeling vice president. Change requests may be triggered by issues identified in model monitoring, by regulatory findings, or by EMVG (described above). All model changes must follow development and implementation guidelines outlined in this manual.

# Model Use and Business Impact

Table 12.1 summarizes the business uses of ADSS. The uses are described more fully thereafter.

Table 12.1: Business uses of the ADSS model

| **Business use** |  | **Description** |
| --- | --- | --- |
| Underwriting new accounts |  | In most cases, ADSS score is used deterministically alongside other application details to decide whether to approve, deny, or cancel an application. The ADSS model may flag certain applicants for closer review by an underwriter for a judgmental decision. |
| Assigning initial credit lines |  | For applicants approved for a lending card, ADSS score informs the initial credit line. |
| Determining pricing |  | For applicants approved for a lending card, ADSS score informs APR (pricing tranche). |
| Determining early activation of lending-on-charge |  | Charge applicants are evaluated for early eligibility for the lending-on-charge feature. |
| Customer management while cardholders are low-tenure |  | For the first 12 months an applicant is with AXP, ADSS score is an input into the models used for customer risk management. |

## Underwriting new accounts

ADSS scores are the primary basis for AXP’s decision to approve or deny applications for new accounts. When AXP receives an application for a new card, its automated Global Decision Engine (discussed above in section 11.1) runs the ADSS algorithm and compares the applicant’s score to pre-established cutoffs. These cutoffs are pre-established by the New Accounts Strategy team within AXP’s Risk Management group. They are based on the expected lifetime profitability to AXP of approving various tranches of cardholders, known as the AXP’s Cardholder Value (CMV) framework.[[44]](#footnote-44)

### Automated cutoff scores

Table 12.2 provides examples of ADSS cutoff scores for U.S. consumer applications; Table 12.3

provides the same for U.S. OPEN (small businesses). Note that these cutoff scores change at management’s discretion are so are illustrative only. ADSS cutoffs vary by product type because different products vary in cardholder profitability,

Table .2: ADSS cutoff scores for accepting U.S. consumer applications (illustrative)

|  |  |  |
| --- | --- | --- |
|  |  | ADSS cutoff score |
| Proprietary lending cards | Proprietary credit cards | 8% |
| Co-branded lending cards | JetBlue credit card | 8 |
| Hilton credit card | 10 |
| Starwood credit card | 10 |
| Morgan Stanley credit card | 10 |
| Mercedes credit card | 10 |
| Delta credit card | 12 |
| Charge cards | Green/Gold Charge Cards | 18 |
| Platinum/Centurion Charge Cards | 18 |

Within OPEN, the profitability of the prospective business account differs based on its size. For this reason, AXP divides OPEN applicants into two categories: “small,” or “medium or larger” (also classified as commercial) (Table 12.3).

Table .3: ADSS cutoff scores for accepting U.S. OPEN applications (illustrative)

|  | AXP product | AXP categorization of the business | ADSS cutoff score |
| --- | --- | --- | --- |
| Lending cards  (FICO must be at least 660) | Lowe's Business Rewards credit card with joint &  several liability | Small | 5% |
| Medium or larger | 6 |
| All lending cards except above | Small | 6 |
| Medium or larger | 8 |
| Charge cards  (FICO must be at least 625)\* | PLUM charge card | Any size | 12 |
| Business Gold Rewards, Gold, or Green  charge cards | Small | 17.5 |
| Medium or larger | 20 |
| Medium or larger and Dun & Bradstreet Commercial Credit Score is at least 500 and FICO is at least 600 | 25 |
| Platinum or Executive  Business charge cards | Small | 21.5 |
| Medium or larger | 22.5 |

\* When evaluating applicants for a Business Gold Rewards, Gold, or Green charge card, AXP accepts FICO as low as 600 for medium-or large-size businesses with a Commercial Credit Score of at least 500.

Table 12.4 presents the rules for assigning OPEN prospects to size categories. Note that these change at management’s discretion and should be read as illustrative.

Table .4: Rules for assigning businesses to size categories (illustrative)

|  |  |  |  |
| --- | --- | --- | --- |
| Segment | Revenue | Employee count | Other criteria |
| Large business | ≥ $1MM | ≥ 10 | * + - * 1. Business owner is a doctor or is part of the board of directors |
| Medium business | ≥ $200K | ≥ 2 | * + - * 1. Existing merchant relationship with AXP         2. USPS-confirmed business address         3. Business type is incorporated or partnership (used only if revenue and employee count are missing) |
| Small business | <$200K | <2 | * + - * 1. Presence of commercial bureau records         2. AXP has high confidence of an active storefront         3. AXP has high confidence of an authorized officer         4. Business type is incorporated or partnership |
| Unknown business | <$200K | <2 | * + - * 1. Non-match against commercial bureau         2. No match with conditions above |

### Manual underwriting

For several reasons, an application may be routed to AXP’s New Accounts Manual Underwriting team for a manual approve/decline decision. This could occur where an applicant does not meet risk thresholds but is potentially high value or high revenue. Alternatively, it could occur if the applicant was pre-approved during prospect targeting but is now declined due to an absence of records at the credit bureaus. Other reasons for manual underwriting are applications for Centurion products or requests for cards from an employee of an AXP co-brand partner. In all of these cases, ADSS is one of several factors generating a flag that routes the application for manual underwriting.

## Assigning initial credit lines

As noted earlier, ADSS is an input into the credit line models that assign lines to each new AXP cardholder. The line finally assigned depends on AXP’s estimate of the account’s debt capacity and lifetime profitability to the company (the CMV estimate described in footnote 50 and 18).

After the line model generates a recommended line for a new account, AXP checks the account against line caps for certain risk tranches (Table 12.5).

Table .5: Example of business rules for capping credit lines

|  |  |
| --- | --- |
| Criterion | Line size |
| ADSS ≥ 5% | Minimum line |
| 4% ≤ ADSS <5% | Cap line at $10K |
| 2% ≤ ADSS <4% | Cap line at $25K |
| ADSS < 1% and applicant is applying for a Delta card and owns a credit card with an external US-based bank | Cap line at $35K |

## Determining pricing (APR)

ADSS influences the risk-adjusted price (i.e., APR) assigned to each new account. Table 12.6 and Table 12.7 provide detailed illustrations.

Table .6: Example of card pricing by ADSS score (U.S. consumer cards)

|  |  |  |  |
| --- | --- | --- | --- |
| Product | Price Tier | ADSS model score | Price point |
| Blue | 1 | ≤ 1.3% | 15.24% |
| 2 | ≤ 3.8 | 18.24 |
| 3 | > 3.8 | 20.24 |
| Legacy Blue Cash Preferred or Blue Sky | 1 | ≤ 0.9 | 17.24 |
| 2 | ≤ 2.6 | 20.24 |
| 3 | > 2.6 | 22.24 |
| Blue Cash EveryDay | 1 | ≤ 0.2 | 12.99 |
| 3 | ≤ 0.3 | 14.99 |
| 5 | ≤ 1.3 | 16.99 |
| 8 | ≤ 2.7 | 19.99 |
| 10 | > 2.7 | 21.99 |
| Delta | 1 | ≤ 0.2 | 15.24 |
| 2 | ≤ 1.0 | 17.24 |
| 3 | > 1.0 | 19.24 |
| Hilton / Starwood | 1 | ≤ 0.2 | 15.24 |
| 2 | ≤ 1.0 | 17.24 |
| 3 | > 1.0 | 19.24 |
| AmEx EveryDay | 1 | ≤ 0.5 | 12.99 |
| 3 | ≤ 0.9 | 14.99 |
| 5 | ≤ 2.8 | 16.99 |
| 8 | ≤ 5.4 | 19.99 |
| 10 | > 5.4 | 21.99 |

Table .7: Example of card pricing by ADSS score (U.S. OPEN cards)

| Product | Price offered | ADSS risk cutoff |
| --- | --- | --- |
| OPEN Simply Cash | Low: prime rate +8.99%  Medium: prime rate +13.99%  High: prime rate +15.99 | ≤0.7%  ≤1.7  >1.7 |
| OPEN Blue | Low: prime rate +7.99  Medium: prime rate +11.99  High: prime rate +15.99 | ≤0.9  ≤2.0  >2.0 |
| OPEN Business Gold Rewards | Low: prime rate +9.99  Medium: prime rate +12.99  High: prime rate +15.99 | ≤0.8  ≤2.4  >2.4 |
| OPEN Lowe’s | Low: prime rate +9.99  Medium: prime rate +11.99  High: prime rate +14.99 | ≤0.7  ≤1.4  >1.4 |
| OPEN Delta | Low: prime rate +11.99  Medium: prime rate +13.99  High: prime rate +15.99 | ≤0.9  ≤3.5  >3.5 |
| OPEN Starwood | Low: prime rate +11.99  Medium: prime rate +13.99  High: prime rate +15.99 | ≤1.1  ≤3.6  >3.6 |

## Activating the Lending on Charge (LOC) feature

All applicants who open a charge card are evaluated to determine whether AXP should activate the lending-on-charge feature from the outset. ADSS is the primary criterion for this early eligibility. To be eligible, a new cardholder must satisfy multiple criteria, as follows.

* ADSS < 3%
* FICO ≥ 660
* Products are Green, Gold, and Platinum

## Managing ongoing cardholder risk

In the first year after opening an account, the cardholder’s ADSS score is an input into AXP’s Customer Decision Support System (CDSS) and Total Structural Risk (TSR) models. CDSS and TSR are the mainstay credit risk models used for the company’s everyday risk management decisions (authorizing line increases, approving point-of-sale transactions, cross-selling additional cards, and so forth).

## Business impact assessment

### Score distribution on recent data

Once the new ADSS model is complete, modelers observe how deploying it would shift the distribution of scores among applicants. This provides a broad view of the direction the scores will move after deployment; it also helps identify the score ranges where AXP can expect the most re-shuffling. Two illustrations are below.

Table .8: How deploying a new model affects ADSS deciles (U.S. consumer, 2014 Q3-4)

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Charge | | |  | Proprietary lending | | |  | Co-branded lending | | |
|  |  | | |  |  | | |  |  | | |
| Decile | Old score | New score | Change |  | Old score | New score | Change |  | Old score | New score | Change |
|  |  |  |  |  |  |  |  |  |  |  |  |
| 1 (highest risk) | 90.5% | 96.8% | + 6.3% |  | 99.8% | 87.9% | - 11.9% |  | 35.3% | 48.0% | + 12.7% |
| 2 | 77.1 | 73.0 | - 4.1 |  | 89.2 | 57.1 | - 32.1 |  | 17.0 | 17.8 | + 0.8 |
| 3 | 67.0 | 59.3 | - 7.7 |  | 68.8 | 41.1 | - 27.7 |  | 10.2 | 8.5 | - 1.7 |
| 4 | 56.8 | 46.9 | - 9.9 |  | 50.7 | 29.8 | - 20.9 |  | 6.3 | 4.4 | - 1.9 |
| 5 | 46.2 | 36.9 | - 9.3 |  | 34.9 | 20.7 | - 14.2 |  | 3.8 | 2.3 | - 1.5 |
| 6 | 36.0 | 28.1 | - 7.9 |  | 21.1 | 13.8 | - 7.3 |  | 1.9 | 1.3 | - 0.6 |
| 7 | 26.0 | 19.7 | - 6.3 |  | 12.7 | 8.8 | - 3.9 |  | 1.0 | 0.6 | - 0.4 |
| 8 | 16.3 | 12.5 | - 3.8 |  | 7.5 | 4.8 | - 2.7 |  | 0.5 | 0.3 | - 0.2 |
| 9 | 7.1 | 6.1 | - 1.0 |  | 3.7 | 2.1 | - 1.6 |  | 0.2 | 0.1 | - 0.1 |
| 10 (lowest risk) | 1.7 | 1.5 | - 0.2 |  | 1.1 | 0.5 | - 0.6 |  | 0.1 | 0.1 | 0.0 |
|  |  |  |  |  |  |  |  |  |  |  |  |
| Mean | 42.5 | 38.1 |  |  | 39.4 | 26.7 |  |  | 7.6 | 8.3 |  |

AXP obtains an alternative view by evaluating the proportion of applicants who will now fall into each key score band (Table 12.9).

Table .9: How deploying a new model affects ADSS score distribution (U.S. consumer, 2014 Q3-4)

|  |  |  |  |
| --- | --- | --- | --- |
|  | Score Band | % Population  (current ADSS score) | % Population  ( new ADSS score) |
|  |  |  |  |
| Charge | < 6 % | 14.0% | 15.0% |
| 6 - 12 | 7.0 | 9.0 |
| 12 - 18 | 6.0 | 8.0 |
| 18 - 25 | 7.0 | 9.0 |
| > 25 | 66.0 | 58.0 |
| Mean Score | 42.5 | 38.1 |
|  |  |  |  |
| Proprietary lending | < 1 | 5.0 | 9.0 |
| 1-3 | 8.0 | 10.0 |
| 3 -5 | 6.0 | 6.0 |
| 5 -8 | 7.0 | 8.0 |
| 8 -12 | 8.0 | 9.0 |
| >12 | 66.0 | 58.0 |
| Mean Score | 39.4 | 26.7 |
|  |  |  |  |
| Co-branded lending | < 1 | 36.0 | 42.0 |
| 1 - 4 | 21.0 | 22.0 |
| 4 - 8 | 14.0 | 11.0 |
| 8 - 10 | 5.0 | 3.0 |
| 10 - 12 | 4.0 | 3.0 |
| 12 -15 | 4.0 | 3.0 |
| >15 | 17.0 | 17.0 |
|  | Mean Score | 7.6 | 8.3 |

### Swap-set analysis

When the model is put into production, some applicants who were previously above a critical threshold (cutoff value) will now fall below it. Conversely, some applicants who were previously below this value will now rise above it. After building the model, modelers perform a “swap set” to replace the first population with the second. This highlights the net number of applicants who will be subject to different treatment should the model take effect. Contrasting the “swap in” from the “swap out” set helps AXP understand the change in portfolio mix—for example, a shift towards higher-income, lower-utilization cardholders (below).

Table .10: Example of swap-set analysis (ADSS model for U.S. consumer lending, Sep. 2015)

|  |  |  | In-In  (approved by both new and old model) | Swap in  (rejected by old model, approved by new one) |  |  | Out-Out (rejected by both new and old model) | Swap out (approved by old model, rejected by new one) |  |  | Difference swap out » swap-in |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| No. of applicants |  |  | 1.32 M | 59.2 k |  |  | 1.17 MM | 69.6 k |  |  |  |
| % population |  |  | 50.0% | 2.0% |  |  | 45.0% | 3.0% |  |  |  |
| Avg. old ADSS score » Avg. new ADSS score |  |  | 1.5 » 1.0 | 10.3 » 3.9 |  |  | 21.1 » 23.4 | 5.2 » 10.0 |  |  |  |
| Defaulters |  |  | 0.9% | 4.6% |  |  | 21.5% | 7.6% |  |  | - 3.0% |
| Average FICO |  |  | 743 | 677 |  |  | 582 | 671 |  |  | - 6 |
| Average income |  |  | $127.4 k | $112.1 k |  |  | $103.2 | $82.6 k |  |  | +$29.5k |
| Inquiry index |  |  | 3.9 | 12.3 |  |  | 15.2 | 12.6 |  |  | - 0.3 |
| Bankcard credit usage |  |  | 35.0% | 58.0% |  |  | 78.0% | 66.0% |  |  | - 8.0% |
| % expatriates |  |  | 10.0% | 17.0% |  |  | 6.0% | 11.0% |  |  | + 6.0% |
| Bureau tenure (years) |  |  | 14.8 | 10.5 |  |  | 9.6 | 8.7 |  |  | + 1.8 |
| External payments |  |  | $19.0k | $8.9 k |  |  | $4.0 k | $7.0 k |  |  | + $1.9 k |

Knowing how the new model affects the portfolio of new acquisitions, modelers next estimate the net impact on the company’s profitability (pre-tax income, or PTI). They do so by observe how the changed distribution will affect billed business, account receivables, write off dollars, and more. Table 12.11 and Table 12.12 present a sample analysis.

Table .11: Simulated metrics to obtain pre-tax income (PTI) of new ADSS model (illustrative only)

|  |  |  |  |
| --- | --- | --- | --- |
| Period | 0-6 months after change | 7-9 months | 10-12 months |
|  |  |  |  |
| ∆ in accounts | (1,152) | (1,086) | (1,047) |
| Credit attrition | (20) | (19) | (23) |
| Non-credit attrition | (46) | (20) | (22) |
| Ending no. of accts | (1,086) | (1,047) | (1,002) |
| Avg. no. of accts | (1,119) | (1,066) | (1,025) |
|  |  |  |  |
| Spending | (3,860,501) | (1,229,860) | (1,076,612) |
| Average daily balance | (688,379) | (862,815) | (847,098) |
| Average A/R | (1,331,796) | (1,272,769) | (1,205,969) |
| Write-off events | (5) | (10) | (15) |

Table .12: Profitability impact of metrics simulated in Table 12.11 (illustrative only)

|  |  |  |  |
| --- | --- | --- | --- |
|  | 0-6 months after change | 7-9 months | 10-12 months |
|  |  |  |  |
| Spending variable income | (60,349) | (19,226) | (16,830) |
| Finance charge revenue | (120) | (565) | (760) |
| Fee revenue | 558 | 266 | 255 |
| Write-offs | (11,457) | (23,521) | (32,469) |
| Reserve expense | (138,440) | (132,304) | (125,360) |
| Provision expense | (149,897) | (17,385) | (25,525) |
| Cost of funds | (15,116) | (7,223) | (6,844) |
| Incentive cost | (69,120) |  |  |
| Contribution Margin | 174,222 | 5,083 | 15,034 |

# Tracking Model Performance

Modelers continually compare ADSS estimates against real-world defaults to evaluate the model’s ongoing performance and identify opportunities for fine-tuning. Table 13.1 summarizes these monitoring activities.

Table 13.1: Activities for monitoring the ongoing performance of ADSS

|  |  |  |
| --- | --- | --- |
| **Activity** | **Frequency** | **Details** |
| Tracking Gini (discrimination), accuracy, and population stability | **US:** quarterly  **International:** quarterly for lead markets;[[45]](#footnote-45) biannually for the rest  For portfolios of individual card products (consumer, OPEN, and international), modelers track performance within six months of the model’s implementation | Modelers monitor the model’s performance by portfolio and segment.[[46]](#footnote-46) Modelers measure accuracy via the method presented in Section **Error! Reference source not found.**. (In short, they compare ADSS scores to default trends from the last period with six months of future data, imputing 12-month default rates from this trend.)  Modelers measure discriminatory power by comparing the current-quarter Gini with the Gini of the previous quarter and early validation data. The comparison is based on actual default rates over a 6-month window.  Modelers calculate the population stability index through the method presented in the appendix. |
| Monitoring input variables (checking that the characteristics of current applicants resemble the development data set) | Monthly | A system called Continuous Data Integrity Tracking (CDIT) provides automated weekly updates on applicant behavioral variables. CDIT sends alerts when the variables vary beyond prescribed ranges. These alerts are a trigger to investigate the discrepancy and identify corrective steps. |
| Verifying that ADSS scores are accurately calculated | Monthly | Each month, AXP’s risk data team randomly samples 1 percent of ADSS scores and compares them with scores calculated offline using verified logic. |
| Reviewing specific cases for insights that may lead to refinements | Monthly | Each month, AXP’s risk team reviews instances of misprediction—for example, where the ADSS score indicates modelers should approve a transaction, but the applicant defaults shortly after. In these cases, modelers scrutinize the applicant’s risk characteristics with the goal of identifying a new variable or technique that may provide earlier warning of default.  Modelers may also perform case reviews when new products are introduced, with the goal of determining how well ADSS predicts default for the new product. |

Note: this table illustrates AXP’s standard practices for tracking the performance of the ADSS model and the stability of its inputs.

Using an ADSS example from the consumer model, the consumer modeling team tracks the ADSS model by calculating the Gini index, accuracy index, and population stability quarterly.

Table 13.2: Illustration of model tracking report

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Segment | Population | 12-months default | Predicted | Accuracy index | Lift from previous quarter | Lift from early validation | Accuracy metric |
| *Low tenure,  thin file* | 12,438 | 49.7% | 57.9% | 83.6% | Satisfactory | Satisfactory | Satisfactory |
| *Low tenure,  thick file* | 5,400 | 6.95% | 4.1% | 45% | Satisfactory | Satisfactory | Needs  Improvement |
| *Med. Tenure,  thick file* | 14,052 | 10.2% | 12.5% | 84.4% | Satisfactory | Satisfactory | Satisfactory |

Note: this table provides an illustration of an ADSS model tracking report.

In this example, the accuracy index for the *Low tenure, thick file* segment falls into the “Needs improvement category” as defined by Modeling Strategy Committee. Thus the segment will be considered for accuracy adjustments to align the segment score with observed default rates. Table 13.3 presents the model rating criteria.

Table 13.3: Rating criteria for model tracking

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Portfolio level** |  |  | **Segment level** |  |
| Criteria | Rating |  | Criteria | Rating |
| Accuracy index less than 50% | Model fail and escalation to Institutional Credit Risk Committee (ICRC)/relevant committees (F) |  | Accuracy index less than 50% and number of defaulters ≥ threshold[[47]](#footnote-47) | Model fail (f) |
| Accuracy index between 50% to 70% | Needs Improvement (I) |  | Accuracy index less than 50% and number of defaulters < threshold | Needs improvement (1) |
| Accuracy index between 70% to 85% | Requires Attention (A) |  | Accuracy index between 50% to 80% | Requires attention (a) |
| Accuracy index > 85% | Satisfactory (S) |  | Accuracy index > 80% | Satisfactory (s) |

Note: this table provides standard AXP protocols for assessing and acting on deteriorating model performance.

### Post-model adjustments

In post-model adjustments, modelers modify base ADSS scores to match recent default trends observed in tracking data. Over time, the market context may change—the combined effect of shifts in consumer behavior, credit availability, product offers, and the economy. The purpose of post-model adjustments is to bring ADSS scores into line with today’s conditions.

Imputing future default rates for tracking data

As noted above, post-model adjustments fine-tune ADSS scores to match the default rates modelers observe in tracking data. But tracking data, which typically begins six months back, is too recent to furnish 12 months of future data. This means modelers must impute default rates *as if* this six-month-old vintage had a full 12 months for its performance to unfold over time.

Modelers do so by taking a vintage of applicants from 12 months ago and calculating its mean default rate at a point six and 12 months later.[[48]](#footnote-48) [[49]](#footnote-49) Let AXP denote the default rate six months after the start point as *Short-Term DRhistorical* and the rate 12 months after as *Long-Term DRhistorical*. Then the ratio [*Long-Term DRhistorical /Short-Term DRhistorical*] acts as a translation factor relating the earlier and later periods. This translation factor is then applied to the tracking data that is too recent for AXP to foresee full performance.

Fig. 13.1: Method for calculating the translation factor

**Translation factor data**

**Tracking data**

**Today**

**TF**

**=**

**T+12**

**t**

**-**

**6**

**t**

**-**

**18**

**t**

**Projected default rate**

Note: this figure illustrates AXP’s method for imputing long-term (18-month) default rates from six months of data.

The table below illustrates. In it, modelers take a vintage of applicants from the low tenure, thick file segment from 12 months before time (t0).[[50]](#footnote-50) Modelers capture the default rates six months (t-12) (77/ 4318 = 1.78 percent) and 12 months (t0) (204/4044 = 5.07 percent) after this start point. modelers then take the ratio between these default rates (5.07/1.78 = 2.85). Multiplying this ratio by one plus the difference in weighted average ADSS scores [2.85\*(1 + 0.16)] yields the translation factor (**3.31**).

**Table 13.4: Illustration of how to obtain the translation factor (from a vintage 12 months ago)**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Default 6 months after start of vintage (time t-12 from the perspective of today)** | | | | **Default 12 months after start of vintage (time t0 from the perspective of today)** | | | |  | | |
| **A** | **B** | **C** | **D** | **G** | **H** | **I** | **J** | **L** | **M** | **N** |
| **# Applicants** | **# Defaulters** | **Default Rate** | **ADSS** | **# Applicants** | **# Defaulters** | **Default Rate** | **ADSS** | **Ratio of default rates** | **Δ in weighted average ADSS** | **Final factor = (1+M)\* L** |
| 4318 | 77 | 1.78% | 4.48% | 4044 | 204 | 5.07% | 3.76% | 2.85 | 0.16 | 3.31 |

Note: this table continues the illustration of extrapolating 18-month default rates from six months of data.

Switching to the point of view of today, if the low tenure, thick file segment has a default rate of 1.95 percent after only six months of observation, modelers can impute that it would have a (1.95 percent x **3.31**) 6.45 percent default if it matured to a full 12 months. Modelers use this imputed default rate of 6.45 percent to assess the accuracy of today’s ADSS scores.

**Applying post-model adjustments**

As illustrated in Table 13.3, ADSS scores must typically be adjusted if (1) this comparison produces an accuracy index of less than 50 percent and (2) the segment under review has more than 5,000 defaulters. For segments with fewer defaulters or an accuracy index greater than 50, the choice to adjust is at the discretion of functional modeling vice president.

The adjustment approach outlined above can be computed at two different levels:

**Segment-level factors**

In this case, using the below equation one unique factor is created for each model segment. This factor is then applied on the raw unadjusted score to get the final accuracy adjusted score.

**Granular (Segment and Score Band) Level Factors**

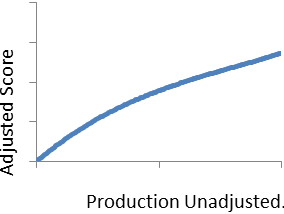
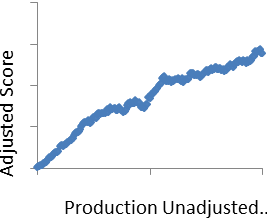
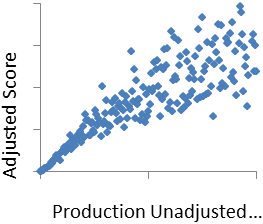
In this scenario, multiple distinct adjustment factors are computed within a model segment – one each for each score band level within that segment. If any score band doesn’t contain sufficient number of observations and/or defaulters, then KNN (k-Nearest Neighborhood) KNN (k-Nearest Neighborhood) technique is used to identify nearest neighborhood of that band (by coalescing adjacent score bands) so that the minimum sample size requirements are satisfied.

The table above also illustrates how a few granular score bands need to be combined as they don’t satisfy the minimum sample size requirements.

Once the score-band/neighborhood is finalized, the adjusted score is calculated on the neighborhood-level summarized data using the same equation as above.

To maintain the rank-ordering order across the borders of the score bands, modelers smooth neighboring adjustment factors using a polynomial function of 4th order. This ensures there is no sharp jump in the model adjustment function. The table below illustrates.

**Figure 13.2: Smoothing of adjustment factor**



**1. Compute granular adjustments**

**2. Apply KNN (clustering) to make adjustments robust**

**3. Apply smoothening function to maintain rank order**

Note: this figure illustrates the use of a k-Nearest Neighborhood smoothing function in the application of accuracy adjustments.

**Deciding between granular and segment-level adjustments**

Both segment and granular level adjusted scores are calculated for each applicant record in the

short-term data. The final choice of whether to proceed with an adjustment and if so, to select

segment-level versus granular-level adjusted score, is driven by the performance of these scores

vis-à-vis the projected 12 month default rates in the short-term data.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Tracking Segment | Default Rate | Raw Score | Accuracy | Adjustment Decision | New Score | New Accuracy |
| Charge | Delinquent | 10.4% | 12.3% | 75.0% | Yes | 10.4% | 0.86 |
| Non-delinquent, thin file | 7.5 | 8.4 | 84.0 | No | 8.4 | NA |
| Non-delinquent, thick file, file, low-bureau tenure | 6.8 | 8.1 | 80.0 | No | 8.1 | NA |
| Non-delinquent, thick file, file, mid-bureau tenure | 6.2 | 6.1 | 85.0 | No | 6.1 | NA |
| Non-delinquent, thick file, file, high-bureau tenure | 4.7 | 4.4 | 85.0 | No | 4.4 | NA |
|  |  |  |  |  |  |  |
| Lending | Delinquent | 3.3 | 3.6 | 84.0 | No | 3.6 | NA |
| Non-delinquent, thin | 4.2 | 2.6 | 61.0 | Yes | 4.2 | 0.85 |
| Non-delinquent, thick file, file, low-bureau tenure | 2.6 | 2.3 | 88.0 | No | 2.3 | NA |
| Non-delinquent, thick file, file, mid-bureau tenure | 1.4 | 1.4 | 86.0 | No | 1.4 | NA |
| Non-delinquent, thick file, file, high-bureau tenure | 0.7 | 0.9 | 70.0 | Yes | 0.7 | 0.93 |

# Key Terms

Table 14.1: Definitions of common terms

|  |  |
| --- | --- |
| **Term** | **Definition** |
| ADSS | The Acquisition Decision Support System, an internal model used by AXP to evaluate credit applications for first-time cardholders. ADSS measures the probability that a new account will enter default in either the next 12 months (for charge) or 24 months (for lending). Lower scores generally represent increased creditworthiness. Also called Q-score. |
| AECB | American Express Centurion Bank, one of AXP’s two U.S. bank operating subsidiaries. AECB issues AXP’s proprietary credit cards and certain consumer charge cards. In addition, AECB offers loans through its lending-on-charge program. |
| A/R | Accounts receivable, which generally represents the total outstanding amount due from cardholders at a given time for charges made on AXP cards. |
| Authorization | The process of granting merchants permission to accept an AXP charge for payment. The process begins when the cardholder attempts payment. The payment request is transmitted to AXP, which calculates the transaction risk and, if the transaction is approved, reserves the sales amount on the cardholder’s account. See CAS. |
| Balance transfer (BT) | The reassignment of part or all of a cardholder balance from another institution into an AXP account, often as part of an initiative to attract new cardholders or increase A/R growth. |
| CARE | The Customer Assistance and Relief Environment, an AXP program intended for cardholders experiencing financial hardship. CARE lowers the required minimum due amount for a limited time and enables cardholders to remain in their current delinquency stage if they fulfill the program’s terms. |
| CAS | The company’s Credit Authorization System, which authorizes transactions for cardholders globally. ADSS scores are inputs into DAC (below), which operates on CAS. |
| Case set up | A flag on an account indicating an AXP customer service representative should contact the cardholder to request a payment. Case set ups are generally initiated because the account is overdue or high-risk (even if current). |
| Credit Bust Out (CBO) | Where cardholders acquire cards with no plan to repay their balance. In these instances, cardholder generally acquire cards from various issuers, and, once their credit history is established, quickly ramp up their expenditure (“bust out”), often using bounced checks to maximize spending. |
| Credit limit | The pre-established amount of money a cardholder may charge to an account at a given time. |
| CPS | Consumer Products and Services, the AXP business unit encompassing card products for individual U.S. cardholders (rather than businesses). |
| Co-brand | See SAC. |
| DAC | Dynamic Authorizations Capability (DAC), an information processing system on CAS. For each incoming transaction, DAC makes an authorization decision by comparing TSR, ADSS, and FICO scores to pre-established business rules that weigh the balance of profitability and risk.Referral Reduction Rules – provide the ability to override the original decline decision by DAC depending on user defined rules |
| Delinquent | An AXP account is typically considered delinquent when the owed balance is 60 or more days past the billing date (30 or more days past the due date). |
| EMVG | Enterprise Model Validation Group (EMVG), the business unit responsible for validating the statistical models AXP uses to make business decisions. |
| FICO | A quantitative measure of an individual’s risk of default, based on consumer files at a credit bureau. A higher score generally represents increased creditworthiness. |
| FSB | American Express Federal Savings Bank, one of AXP’s two U.S. bank operating subsidiaries. FSB issues OPEN charge and credit cards as well as U.S. consumer co-branded credit cards. In addition, FSB has outstanding lines of credit in association with certain OPEN and consumer charge cards and offers loans through its lending-on-charge program. |
| Global limit (GL) | A communicated spending limit that AXP imposes on a small proportion of higher-risk charge accounts to control exposure. AXP manages the limits based on cardholder risk and economic conditions. Cardholders’ transactions may be declined if their balance hits the global limit. As such, global limits operate as a temporary de facto credit limit. |
| Lending cards | Credit cards, which allow cardholders to make purchases and either pay the balance in full each month or revolve the balance month-over-month and keep making charges up to the credit limit (so long as they pay the required monthly minimum due). |
| LOC | Lending-on-charge, a feature on some charge cards that enables cardholders to revolve certain balances (that is, carry these balances over from one billing cycle to the next). |
| MAS | Modeling Automation Suite, a proprietary tool used widely in AXP’s risk management practice. The tool recommends the most predictive initial logistic regression model by identifying an optimal subset of independent variables. |
| MR | Membership Rewards points, which cardholders earn through eligible spending on AXP cards. MR points are redeemable for a wide array of rewards, including travel, retail merchandise, and dining and entertainment. |
| MSC | The Modeling Strategy Committee, a body of RIM executives that convenes roughly monthly. For the models within its purview, MSC evaluates the model logic and input data; tracks model performance; directs corrective action where performance deteriorates; and ensures models comply with relevant laws and regulations. |
| OPEN | The AXP business unit encompassing card products for U.S. small businesses. |
| Portfolios | Groupings of accounts by business. An example portfolio is consumer proprietary lending cards. A portfolio may encompass multiple segments. |
| Proprietary lending cards | Credit cards that solely carry the AXP brand, in contrast to strategic alliance and co-brand (SAC) cards. AXP bears the full cost of marketing, operations, member rewards, and credit risk associated with proprietary cards. |
| Re-age | Reclassifying an account’s delinquency status (for example, reclassifying a delinquent account as current). Cardholder accounts are typically re-aged when they enter a hardship program (see definition for CARE, above). |
| Responsible lending actions (RLA) | Actions taken by AXP management to control exposure by repricing or restricting credit. Examples of RLA actions are reducing credit lines, imposing global limits, canceling cards, and suspending the LOC feature on charge cards. |
| Revolvers | Cardholders who typically roll over unpaid balances from one billing cycle to the next. |
| Segments | Categories of accounts with common risk characteristics (for example, FICO score less than 700). Segments are typically subsets of portfolios. |
| Strategic alliance co-brand (SAC) (also known as co-brand) | Cards that are issued by AXP under co-brand marketing agreements with U.S. companies. Cardholders earn rewards based on spending through the partners’ loyalty programs, for example, frequent flyer miles, hotel loyalty points, or cash back. Generally, AXP’s partner is responsible for providing rewards and AXP retains the credit risk. |
| Transactors | Cardholders who typically pay their full balance at the end of each billing cycle. |
| TSR | Total Structural Risk, the model used by AXP to assess the probability that AXP cardholders will enter default on any of their relationships both within and outside AXP in the next 18 months. TSR differs from ADSS in that it depends more heavily on external data. |
| United States | U.S. states plus territories, including Puerto Rico, U.S. Virgin Islands, American Samoa, Guam, Northern Mariana Islands, Marshall Islands, Palau, and Micronesia. |
| Q-score | See ADSS. |
| Write-offs | Accounts receivable (A/R) that is recorded as a loss, either due to prolonged delinquency or a cardholder’s bankruptcy, death, or settlement agreement. Synonymous with charge-offs. |

Note: this table defines key terms related to credit risk management at AXP.

# Appendix: bibliography

**External references**

Abdou, H.A. and J. Pointon (2011) “Credit Scoring, Statistical Techniques and Evaluation Criteria: A Review of the Literature,” *Intelligent Systems in Accounting, Finance & Management*, Vol. 18, No. 2-3, pp. 59-88.

Chakravorti, S. and T. To (2007) “A theory of credit cards,” *International Journal of industrial organization*, Vol. 25, No. 3, pp.583-595.

[Crone](https://scholar.google.co.uk/citations?user=sVW29BIAAAAJ&hl=en&oi=sra), S.F. and S. Finlay (2012) “[Instance sampling in credit scoring: An empirical study of sample size and balancing](http://www.sciencedirect.com/science/article/pii/S0169207011001403),” *International Journal of Forecasting*, [Vol. 28, No. 1](http://www.sciencedirect.com/science/journal/01692070/28/1), pp. 224–238.

Crook, J.N., D.B. Edelman and L.C. Thomas (2007) “Recent developments in consumer credit risk assessment,” *European Journal of Operational Research*, Vol. 183, No. 3, pp. 201447-1465.

Finlay, S. (2009) “Are modelers modeling the right thing? The impact of incorrect problem specification in credit scoring,” [*Expert Systems with Applications*](http://www.sciencedirect.com/science/journal/09574174), [Vol. 6, No. 5](http://www.sciencedirect.com/science/journal/09574174/36/5), pp. 9065–9071.

Friedman, Jerome H. (1999) “Greedy Function Approximation: A Gradient Boosting Machine,” *Annals of Statistics*. Vol. 29.

Friedman, Jerome; Hastie, T. and Tibshirani, R. (1998), “Additive Logistic Regression: a Statistical View of Boosting.” *Annals of Statistics*. Vol. 28.

Hamerle, A., R. Rauhmeier and D. Rösch (2003) “Uses and Misuses of Measures for Credit Rating Accuracy,” available at: <http://ssrn.com/abstract=2354877>.

Hastie, T., Tibshirani, R. and Friedman (2009). The Elements of Statistical Learning; Data Mining, Inference and Prediction. Springer Verlag.

Keenan, S.C., J.R. Sobehart and D.T. Hamilton (1999) “Predicting default rates: a forecasting model for Moody's issuer-based default rates,” available at: <http://ssrn.com/abstract=1020303>

[Keramati](https://scholar.google.co.uk/citations?user=nTxkJqoAAAAJ&hl=en&oi=sra), A. and [N. Yousefi](https://scholar.google.co.uk/citations?user=DsciMcMAAAAJ&hl=en&oi=sra) (2011) “A Proposed Classification of Data Mining Techniques in Credit Scoring,” *Proceedings of the 2011 International Conference on Industrial Engineering and Operations Management*, Kuala Lumpur, Malaysia, January 22 – 24.

Lessmann, S., B. Baesens, H. Seow and L.C. Thomas (2015) “Benchmarking state-of-the-art classification algorithms for credit scoring - A ten-year update,” [*European Journal of Operational Research*](http://www.sciencedirect.com/science/journal/03772217), [Vol. 247, No. 1](http://www.sciencedirect.com/science/journal/03772217/247/1), pp. 20124–136.

Liu, F., Z. Hua and A. Lim (2015) “Identifying future defaulters: A hierarchical Bayesian method,” [*European Journal of Operational Research*](http://www.sciencedirect.com/science/journal/03772217), [Vol.241, No. 1](http://www.sciencedirect.com/science/journal/03772217/241/1), pp. 202–211.

Marqués, A.I., V. García and J.S. Sánchez (2013) “A literature review on the application of evolutionary computing to credit scoring,” *Journal of the Operational Research Society*, Vol. 64, No. 9, pp. 201384–1399.

Martens, D., B. Baesens, T. Van Gestel and J. Vanthienen (2007) “Comprehensible credit scoring models using rule extraction from support vector machines,” [*European Journal of Operational Research*](http://www.sciencedirect.com/science/journal/03772217), [Vol. 183, No.3](http://www.sciencedirect.com/science/journal/03772217/183/3), pp. 201466–1476.

Morrison, J.S., (2010) “Marrying Credit Scoring and Time-Series Data,” *RMA Journal*.

Sobehart, J.R., S.C. Keenan and R. Stein (2000) “Benchmarking quantitative default risk models: a validation methodology,” *Moody’s Investors Service, Global Credit Research, Rating Methodology*, March.

Stein, R.M., (2007) “Benchmarking default prediction models: Pitfalls and remedies in model validation,” *Journal of Risk Model Validation*, Vol. 1, No. 1, pp. 77-113.

Thomas, L.C. (2000) “A survey of credit and behavioural scoring: forecasting financial risk of lending to consumers,” *International Journal of Forecasting*, Vol. 16, No. 2, pp. 20149–172.

Thomas, L.C., D.B. Edelman and J.N. Crook (2002) “Credit scoring and its applications,” Society for Industrial and Applied Mathematics, Philadelphia.

Tong, E.N.C., C. Mues and L.C. Thomas (2012) “Mixture cure models in credit scoring: If and when borrowers default,” [*European Journal of Operational Research*](http://www.sciencedirect.com/science/journal/03772217), Vol. 218, No. 1, pp. 20132-139.

**Internal references**

“American Express Management Policy 50: Enterprise-wide Risk Management,” 01-01-2016.

“American Express Management Policy 55: Model Governance and Validation,” 15-07-2015.

“Model documentation Standards,” 15-06-2015, Available at: *06\_10\_2015\_Model Documentation Standards (Final).pdf.*

“MAS Modeling Document,” Available at: *Enhanced MAS Modeling Document.docx.*

# Appendix: multivariate logistic regression

## Specification

The logistic regression model considers independent response variables, , where each response variable is dependent on predictor variables taking values , with . Each follows a Bernoulli distribution with probability of success , where

.

The odds ratio of the response variable is defined to be . If for example, the odds ratio is , or to . If , the odds ratio is , or to . Our assumption is that the natural logarithm of the odds ratio depends linearly on the predictor variables, ie.

(A12.1)

The value is usually denoted . This model can be re-expressed in vector form as

(A12.2)

where . Taking the exponent on both sides and rearranging, one obtains

(A12.3)

## Maximum-likelihood estimation

For each let be the observed value of the variable , and let . The values of the parameters may be estimated using the method of maximum likelihood. The likelihood function is defined to be the value of the joint probability distribution at . Since the are independent, this can be expressed as

(A12.4)

Finding the values of the parameters that maximize the likelihood function requires the use of numerical methods.

## Diagnostics

### Wald statistics

The Wald test is used to compare the maximum likelihood estimate for a distribution parameter with a proposed value. Suppose that is the maximum-likelihood estimate for a parameter and that is the proposed value of the parameter, then the Wald statistic is defined as

(A12.5)

The statistic is compared to a chi-squared distribution to determine whether to reject the hypothesis that .

### Variance inflation factors

The variance inflation factor measures the inflation of the variance of a regression coefficient due to correlation with other regression coefficients. Suppose there are predictor variables, , with regression coefficients . Then the variance inflation factor of is given by the formula

(A12.6)

where is the coefficient of determination of regressed on the remaining predictor variables. If then the multicollinearity is high.

### Hosmer-Lemeshow test

The Hosmer-Lemeshow test may be applied to a collection of variables taking one of two outcomes. To apply the Hosmer-Lemeshow test, the observations of these variables are first ordered by their fitted probabilities. They are then divided into groups of sizes . The group sizes should be approximately equal.

Let denote the outcome of the th observation in the th group, and let denote the fitted probability. The Hosmer-Lemeshow statistic is defined as

(A12.7)

The asymptotic distribution of this statistic is approximately chi-squared.

# Appendix: model performance metrics

## Measures of discrimination and accuracy

To evaluate model performance on a data set containing both predicted and actual metrics, the data set is separated into rank-groups based on the predicted metrics. These rank-groups are ordered.

For each let be the fraction of the population falling within the first rank-groups. Suppose modelers have some subset of the population for which a specified event occurs, then for each let be the fraction of this subset of the population that falls within the first rank-groups. The Lorenz curve for this event is constructed by plotting the points and interpolating linearly between adjacent points.

Such Lorenz curves could be constructed for any event including predicted events, actual events, negative of actual events, or random events. In such cases modelers will use the notations , , and .

Figure 13-1: Lorenz curve example



The area enclosed between the actual event line and the random event line will be denoted , and the area between the predicted event line and the random event line will be denoted . The area of the half of the graph above and to the left of the random event line is denoted ; clearly .

Let denote the actual outcome of the th observation, and let denote the predicted outcome. Let be the total number of observations. In the context of Lorenz curves such as those shown above, the following model performance metrics may be constructed and are widely used throughout this document:

* Gini coefficient (): This is defined as the ratio of to . This simplifies to the formula

(A13.1)

* Maximum possible Gini coefficient (): This is defined as the ratio of to . This simplifies to the formula

(A13.2)

* Modified Gini coefficient (): This is defined as the ratio of to . This can be reformulated as

(A13.3)

* Spearman’s rank correlation (): Suppose each observation is ranked based on predicted and actual outcomes, and that is the difference between the two ranks for the th observation. Spearman’s rank correlation is given by the formula

(A13.4)

* Accuracy index (): Let denote the actual outcome of the th observation in the th rank-group, and let denote the predicted outcome. For each let be the size of the th rank-group. The accuracy index is given by the formula

(A13.5)

* Modified accuracy index (): This is given by the formula

(A13.6)

where is the total number of defaults.

* Kolmogorov-SmirNov. (): The maximal distance between the actual event line and the non-event line. It is given by the formula

(A13.7)

* Root-mean-square error (): A measure of the difference between the actual outcomes and the predicted outcomes. It is given by the formula

(A13.8)

## Population stability

Segmentation stability is measured by tracking quarter-over-quarter changes in percentage of overall population falling within each segment, and then aggregating changes over all segments using the following metric, constructed to be non-negative:

(A14.1)

# Appendix: Using MAS to impute or transform data

Data transformations are performed using an AXP proprietary data analysis application called the Modeling Automation Suite (MAS). MAS suggests the best transformation for an independent variable; transformations considered include (1) capping and flooring values, (2) change of variable to linear, natural logarithm or square root of the original variable; and (3) missing value imputation.

To compute the preferred transformation for an independent variable (), the following methodology is applied.

1. Three datasets are created corresponding to the linear (untransformed), log and square root of the variable. Perform the following steps for each of the three datasets.
2. Observations are split into buckets based on the values of . Each contains an equal number of elements. The number of buckets is determined by the user. Missing observations are treated as a separate bucket. The average value of the variable is calculated for each bucket, , for .
3. The application computes a minimum permitted difference () between the minimum and maximum bucket specific averages, . For any subset of buckets for which the difference in average values for the minimum and maximum values of exceeds, the following steps are followed:
   1. Calculate the fraction of the population, denoted , contained in the buckets under consideration denoted.
   2. For each bin, , calculate the and the log odds (). Calculate the difference, , between the log odds for the two bins for which is highest and lowest
   3. Perform an ordinary linear regression of on . Let denote the R-square statistic of this regression. The log odds of the dependent variable from the missing bucket, denoted is inserted into the inverted regression equation to obtain an imputed value of the independent variable for the missing bucket.
   4. Calculate the MAS statistics defined as:
4. If the combination of bin range and functional form that gives the maximum MAS statistic contains first or last or both buckets of the variable, further binning is performed in that the first two buckets or the last two buckets or both the first two and last two buckets are split and these additional bins are appended to the existing bins. Steps 1 to 5 are then repeated. Otherwise, the combination of transformation and set of bins that yields the maximum MAS statistic is selected.

The selected transformation provides the functional form and the imputed value for missing values. The cap and floor values are the corresponding to the minimum and maximum bins of the suggested transformation.

[

# Appendix: Cardholder Value (CMV) calculation

To calculate an applicant’s Cardholder Value (CMV), AXP estimates cash flows for each year after the account’s approval. Cash flow in a given year is a function of net income adjusted for changes in risk-based economic capital requirements, as follows.

Once cash flow is determined for each year, CMV is calculated as:

# Appendix: Overview of card products and portfolios

AXP’s principal products are charge and credit payment cards. Credit cards, internally called lending cards, allow cardholders to make purchases and either pay the balance in full each month or revolve it month-over-month and keep making charges up to the credit limit.

Charge cards generally carry no preset spending limit and are intended as a method of payment rather than a way to finance purchases. They are designed to be paid in full each month, but may include a lending-on-charge (LOC) feature, which enables cardholders to revolve certain balances. Cardholders enrolled in lending-on-charge may choose to automatically “sweep” all charges of a specific type or over a designated amount (for example, $100 or $500) into their revolving balance. The intent of this feature is to help when traveling or to extend repayment for a larger purchase.[[51]](#footnote-53)

To remain current, lending cardholders must pay their minimum due payment on time; charge cardholders must pay their pay-in-full monthly balance and make minimum due payments on the lending-on-charge portions of their accounts.

AXP offers both proprietary and co-branded cards, which are also called Strategic Alliance and Co-Brand (SAC) cards. An example of a proprietary card is Blue Cash Everyday®; an example of a co-branded card is Gold Delta SkyMiles® Credit Card from American Express.

# Appendix: AXP application fields by market

**Table 3.2.1:** Application Data Requirements

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  |  |  |
|  | | | **ARG** | | **AUS** | | **AUT** | | **CA** | | **FIN** | | **FRA** | | **GER** | | **HK** | | **IDC** | | **IND** | | **ITA** | | **JP** | | **MX** | | **ND** | | **NZ** | | **SGP** | | **SPA** | | **SWE** | | **TWN** | **THA** | **UK** |
| Basic | | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  |  |  |
| Name | | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | \*\* | \*\* |
| Home Address | | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | \*\* | \*\* |
| SSN / National ID | | | \*\* | | \* | | \*\* | | \* | | \*\* | | \* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \* | | \*\* | | \*\* | | \* | | \*\* | | \*\* | | \*\* | | \*\* | \*\* | \* |
| Email ID | | | \*\* | | \* | | \*\* | | \* | | \*\* | | \* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \* | | \*\* | | \*\* | | \* | | \*\* | | \*\* | | \*\* | | \*\* | \*\* | \* |
| Date of Birth | | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | \*\* | \*\* |
| Phone (any one) | | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | \*\* | \*\* |
| Own or Rent home | | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  |  |  |
| Employer Name | | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  |  |  |
| Annual Income | | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | \*\* | \*\* |
| Asset Type# (Checking/Savings/Retirement/Other) | | | \* | | \* | | \*\* | | \* | | \*\* | | \*\* | | \*\* | | \* | | \*\* | | \* | | \*\* | | \* | | \* | | \*\* | | \* | | \* | | \*\* | | \*\* | | \* | \* | \*\* |
| Asset Amount# | | | \* | | \* | | \*\* | | \* | | \*\* | | \*\* | | \*\* | | \* | | \*\* | | \* | | \*\* | | \* | | \* | | \*\* | | \* | | \* | | \*\* | | \*\* | | \* | \* | \*\* |
| email (Online) | | | \*\* | | \*\* | | \*\* | | \* | | \* | | \*\* | | \*\* | | \*\* | | \*\* | | \* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \* | | \*\* | \*\* | \*\* |
| Choose your card (JetBlue) | | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  |  |  |
| Partner Program ID | | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  |  |  |
| Signature | | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  |  |  |
|  |  |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |
| Name | | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | \*\* | \*\* |
| Date of Birth | | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | \*\* | \*\* |
| SSN/ National ID | | | \*\* | | \* | | \*\* | | \* | | \*\* | | \* | | \*\* | | \*\* | | \*\* | | \* | | \*\* | | \* | | \*\* | | \*\* | | \* | | \*\* | | \*\* | | \*\* | | \*\* | \*\* | \* |
| Basic card | | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  |  |  |
| Name | | |  | | \*\* | | \*\* | | \*\* | |  | | \*\* | | \*\* | | \*\* | |  | |  | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | |  | | \*\* | |  | | \*\* |  | \*\* |
| Home Address | | |  | | \*\* | | \*\* | | \*\* | |  | | \*\* | | \*\* | | \*\* | |  | |  | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | |  | | \*\* | |  | | \*\* |  | \*\* |
| SSN / National ID | | |  | | \* | | \*\* | | \* | |  | | \* | | \*\* | | \*\* | |  | |  | | \*\* | | \* | | \*\* | | \*\* | | \* | |  | | \*\* | |  | | \*\* |  | \* |
| Date of Birth | | |  | | \*\* | | \*\* | | \*\* | |  | | \*\* | | \*\* | | \*\* | |  | |  | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | |  | | \*\* | |  | | \*\* |  | \*\* |
| Phone (any one) | | |  | | \*\* | | \*\* | | \*\* | |  | | \*\* | | \*\* | | \*\* | |  | |  | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | |  | | \*\* | |  | | \*\* |  | \*\* |
| Total Annual Income | | |  | | \*\* | | \*\* | | \*\* | |  | | \*\* | | \*\* | | \*\* | |  | |  | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | |  | | \*\* | |  | | \*\* |  | \*\* |
| Asset Type (Checking/Savings/Retirement/Other) | | |  | | \* | | \*\* | | \* | |  | | \*\* | | \*\* | | \* | |  | |  | | \*\* | | \* | | \* | | \*\* | | \* | |  | | \*\* | |  | | \* |  | \* |
| Asset Amount | | |  | | \* | | \*\* | | \* | |  | | \*\* | | \*\* | | \* | |  | |  | | \*\* | | \* | | \* | | \*\* | | \* | |  | | \*\* | |  | | \* |  | \* |
| email (Online) | | |  | | \*\* | | \*\* | | \* | |  | | \*\* | | \*\* | | \*\* | |  | |  | | \*\* | | \*\* | | \*\* | | \*\* | |  | |  | | \*\* | |  | | \*\* |  | \*\* |
| Choose your card (JetBlue) | | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  |  |  |
| Partner Program ID | | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  |  |  |
| Company Name | | |  | | \*\* | | \*\* | | \*\* | |  | | \*\* | | \*\* | | \*\* | |  | |  | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | |  | | \*\* | |  | | \*\* |  | \*\* |
| Company Structure | | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  |  |  |
| Tax ID Number | | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  |  |  |
| Company Address | | |  | | \*\* | | \*\* | | \*\* | |  | | \*\* | | \*\* | | \*\* | |  | |  | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | |  | | \*\* | |  | | \*\* |  | \*\* |
| Company phone | | |  | | \*\* | | \*\* | | \*\* | |  | | \*\* | | \*\* | | \*\* | |  | |  | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | |  | | \*\* | |  | | \*\* |  | \*\* |
| Industry Type | | |  | | \*\* | | \*\* | | \*\* | |  | | \* | | \*\* | | \*\* | |  | |  | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | |  | | \*\* | |  | | \*\* |  | \*\* |
| Years in Operation | | |  | | \*\* | | \*\* | | \*\* | |  | | \* | | \*\* | | \*\* | |  | |  | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | |  | | \*\* | |  | | \*\* |  | \*\* |
| Company Name on the Card | | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  |  |  |
| Number of Employees | | |  | | \*\* | | \*\* | | \*\* | |  | | \* | | \*\* | | \*\* | |  | |  | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | |  | | \* | |  | | \*\* |  | \*\* |
| Title of Authorized Officer | | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  |  |  |
| Business Revenue | | |  | | \*\* | | \*\* | | \* | |  | | \* | | \*\* | | \*\* | |  | |  | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | |  | | \* | |  | | \*\* |  | \*\* |
| Estimated Monthly Spend | | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  |  |  |
| Business Registration Certificate | | |  | | \*\* | | \* | | \* | |  | | \*\* | | \* | | \*\* | |  | |  | | \*\* | | \*\* | | \* | | \*\* | | \*\* | |  | | \* | |  | | \*\* |  | \*\* |
| Annual Net Income | | |  | | \*\* | | \*\* | | \* | |  | | \* | | \*\* | | \*\* | |  | |  | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | |  | | \* | |  | | \*\* |  | \* |
| *Supps* | | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  |  |  |
| Name | | |  | |  | | \*\* | | \*\* | |  | | \*\* | | \*\* | | \*\* | |  | |  | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | |  | | \*\* | |  | | \*\* |  | \*\* |
| SSN/ National ID | | |  | |  | | \*\* | | \*\* | |  | | \*\* | | \*\* | | \*\* | |  | |  | | \*\* | | \*\* | | \*\* | | \*\* | | \*\* | |  | | \*\* | |  | | \*\* |  | \*\* |
| Date of Birth | | |  | |  | | \*\* | | \* | |  | | \* | | \*\* | | \* | |  | |  | | \*\* | | \* | | \*\* | | \* | | \*\* | |  | | \* | |  | | \*\* |  | \*\* |

\*\*: Mandatory Field; \*Non-Mandatory Field; In International markets Bank information is sought

ADSS variables and their definitions

|  |  |
| --- | --- |
| *dall010* | Number of trades 30 days delinquent, but never more than 30 days delinquent, reported within the last 6 months |
| *dall020* | Number of open trades with a current 30 day delinquency |
| *dall022* | Number of open trades at 100% or more credit usage |
| *dall024* | Number of months the oldest trade has been open |
| *dall025* | Number of months the newest trade has been open |
| *dall028* | Number of open trades with a balance > 100 |
| *dall031* | Number of trades to new to rate, or being disputed |
| *dall033* | Number of trades opened with the last 24 months |
| *dall038* | Maximum available credit on an open trade with activity within the last 12 months |
| *dall041* | Number of months since the most recent 90+ day delinquency on a trade |
| *dall043* | Number of months since the most recent 30, 60, or 90+ day delinquency on a trade |
| *dall044* | Average percentage of credit usage on open trade opened within the last 24 months |
| *dall048* | Avg util% trd opn opnd≤12 |
| *dall051* | Percentage of the maximum balance on open trades with activity within the last 12 months and the sum of the balances on trades with activity within the last 12 months |
| *dall053* | Number of months since most recent trade with rating of 60+ days delinquent |
| *dall087* | Ratio of number of open trades within the last 24 months to the number of open trades |
| *dall123* | Number of trades considered for major derogatory event processing that are presently 90+ days delinquent |
| *dall170* | Number of trades considered for major derogatory event processing where the worst rating was 00 (current) within the last 24 months |
| *dall340* | Number of accepted trades that are not too new to rate, reported within the last 24 months |
| *dall344* | Number of accepted trades that have a credit usage percentage greater than 75% reported within the last 24 months |
| *dall347* | Sum of available credit on accepted trades reported within the last 24 months |
| *dall353* | Number of accepted trades opened within the last 12 months reported within the last 24 months |
| *dall376* | Number of open trades with non-missing balance and non-missing credit opened within the last 6 months |
| *dall404* | Bureau tenure |
| *dall405* | Charge segment |
| *dall412* | Golden eye composite indicator |
| *dall437* | Positive index |
| *dall438* | External delinquency index |
| *dall449* | Number of authorized user trades |
| *dappv017* | Checking/savings code (no=1;chk=2;bth=3;sav=4;unk=.) |
| *dappv052* | Internet channel |
| *dappv063* | Application supps requested |
| *dappv078* | Application type (1=to,2=cs,3=dm,4=it,5=ot, .=unk) |
| *dappv107* | Blue indicator |
| *dappv110* | Costco indicator |
| *dappv111* | Plum indicator |
| *dappv112* | Lowe’s indicator |
| *dappv214* | Checking / saving |
| *dappv216* | Bank account information = both checking or savings |
| *dappv217* | Sac indicator |
| *dappv228* | Annual income |
| *dappv231* | Pa indicator |
| *dappv244* | Ssn issuance in the last 5 years indicator |
| *dccrscls* | Ccs class in 1,2 |
| *dccrsper* | Ccs score |
| *dfico001* | Fico score from credit bureau (valid range 200-900) |
| *dinqv005* | Number of inquiries with valid member number reported within the last 1 month |
| *dinqv006* | Number of inquiries |
| *dinqv007* | Number of inquiries reported within the last 6 months |
| *dinqv031* | Number of inquiries reported within the last 12 months |
| *dinqv034* | Number of months since the newest inquiry was reported |
| *dinqv035* | Number of inquiries (that are not amex) reported within the last 6 months |
| *dinqv099* | New inquiry index |
| *dloos022* | Number of rvbc trades opened within the last 24 months that have ever been 30+ days delinquent |
| *dloos030* | Months since oldest open rvbc trade with balance greater than 50 was opened |
| *dloos035* | Sum of the number of months open rvbc trades have been opened or reported |
| *dloos038* | Average number of months opened on open rvbc trades |
| *dloos040* | Number of rvbc trades opened within the last 3 months |
| *dloos049* | Number of rvbc trades opened at least 72 months ago |
| *dloos064* | Percentage of open rvbc trades opened within the last 6 months |
| *dloos106* | Average percentage of credit usage for open rvbc trades with a balance greater than 50 and a non-missing credit |
| *dloos120* | Student loan indicator |
| *dloos155* | Number of valid trades with balance > 0 |
| *dmrtg025* | Number of months the newest trade has been open |
| *dmrtg034* | Number of trades with activity with the last 24 months |
| *dmrtg400* | 0 |
| *dmrtg401* | Mortgage pay down |
| *dmsc001* | Amex supp relationship |
| *dnbccs7* | Dnb ccs7 score |
| *dplst022* | Number of open trades at 100% or more credit usage |
| *dplst036* | Maximum percentage of credit usage on an open trade with activity within the last 12 months |
| *dplst038* | Maximum available credit on an open trade with activity within the last 12 months |
| *dplst048* | Average percentage of credit usage on open trade opened with the last 12 months |
| *dplst049* | Average percentage of credit usage on open trade opened with the last 6 months |
| *dplst070* | Number of trades considered for major derogatory event processing |
| *dplst079* | Sum of available credit on non-collection trades that are considered for major derogatory event processing |
| *dplst085* | Avergae credit usage of trades considered for major derogatory event processing reported within the last 6 months |
| *dplst344* | Number of accepted trades that have a credit usage percentage greater than 75% reported within the last 24 months |
| *dplst351* | Number of months since the oldest accepted trades was opened reported within the last 24 months |
| *dplst372* | Maximum available credit for valid trades that are non-collection |
| *dplst378* | Average credit usage of credit for trades with non-missing balance and non-missing credit opened within the last 12 months |
| *dplst380* | Sum of balances on trades with balance greater than or equal to 0 and credit greater than 0 |
| *dplst382* | Ratio of sum of balances over max credit on trades with non-missing balance and non-missing credit |
| *dprtl028* | Number of open trades with a balance > 100 |
| *dprtl036* | Maximum percentage of credit usage on an open trade with activity within the last 12 months |
| *dprtl044* | Average percentage of credit usage on open trade opened within the last 24 months |
| *dprtl400* | Number of retail cards |
| *drvbc022* | Number of open trades at 100% or more credit usage |
| *drvbc024* | Number of months the oldest trade has been open |
| *drvbc031* | Number of trades to new to rate, or being disputed |
| *drvbc038* | Maximum available credit on an open trade with activity within the last 12 months |
| *drvbc039* | Minimum available credit on an open trade with activity within the last 12 months |
| *drvbc051* | Percentage of the maximum balance on open trades with activity within the last 12 months and the sum of the balances on trades with activity within the last 12 months |
| *drvbc060* | Percentage of the number of open trades with up-to-date payments and the number of open trades |
| *drvbc080* | Average credit usage of trades considered for major derogatory event processing reported within the last 24 months |
| *drvbc148* | Number of trades considered for major derogatory event processing where the worst rating was 00 (current) within the last 12 months |
| *drvbc340* | Number of accepted trades that are not too new to rate, reported within the last 24 months |
| *drvbc350* | Number of months since the most recent accepted trades was opened reported within the last 24 months |
| *drvbc351* | Number of months since the oldest accepted trades was opened reported within the last 24 months |
| *drvbc373* | Average credit usage of maximum available credit for valid trades that are non-collection |
| *drvbc376* | Number of open trades with non-missing balance and non-missing credit opened within the last 6 months |
| *dsowv001* | Amount of balance transfers in last quarter (sw-amt-bl-xfr-1) |
| *dsowv002* | Preferred revolving card credit limit (sow-loc-max-rev-bal) |
| *dsowv003* | Preferred transacting card credit limit (sow-loc-max-xact-bal) |
| *dsowv004* | Preferred revolving card balance (sw-mxrev-bl-1) |
| *dsowv005* | Preferred transacting card balance (sw-mxtrn-bl-1) |
| *dsowv006* | Sow-num-rvl-crd |
| *dsowv007* | Sow-num-xfer-crd |
| *dsowv008* | Preferred revolving card pay down percent (sow-pay-pct-max-rev-bal) |
| *dsowv009* | Preferred revolving card spend (sw-spnd-mxrev-bl-1) |
| *dsowv010* | Preferred transacting card spend (sw-spnd-mxtrn-bl-1) |
| *dsowv011* | Total annual external spend (sow-tot-spend) |
| *dsowv012* | Total transacting balance (sw-tot-trn-bl-1) |
| *dsowv013* | Average monthly external spend in q4 (sw-av-spnd4-1) |
| *dsowv014* | Total revolving balance (sw-tot-rvl-bl-1) |
| *dsowv015* | Number of balance transfers in last quarter (sw-num-bl-xfr-1) |
| *dsowv016* | Average pay down percent on revolving cards (sw-av-rev-pyd-1) |
| *dsowv017* | Average monthly external payment in last quarter (sw-av-pay-3m-1) |
| *dsowv018* | Number of trades with line decreases in last year (sw-line-decrease-1) |
| *dsowv019* | Number of trades with line increases in last year (sw-line-increase-1) |
| *dsowv020* | Number of balance parker trades (sw-num-bt-parker-1) |
| *dsowv021* | Probability of high airline spend (sw-prob-air-1) |
| *dsowv022* | Total annual external payment (sw-tot-payment-1) |
| *dsowv023* | Total annual spend on retail trades (sw-tot-spnd-rtl-1) |
| *dsowv024* | Size-of-wallet total revolving balance on retail trades (sw-totrvl-bl-rtl-1) |
| *dsowv025* | Total annual spend on loc trades (sw-tot-spnd-loc-1) |
| *dsowv026* | Total revolving balance on loc trades (sw-totrvl-bl-loc-1) |
| *dsowv027* | Size-of-wallet strategy variable 3 (sw-sow-03-1)- credit line on preferred spend card |
| *dsowv028* | Size-of-wallet strategy variable 4 (sw-sow-04-1); total spend from joint accts. |
| *dsowv029* | Size-of-wallet strategy variable 5 (sw-sow-05-1 |
| *dsowv030* | Average monthly external spend in q1 (sw-av-spnd1-1) |
| *dsowv031* | Size-of-wallet strategy variable 2 (sw-sow-02-1); spend on preferred spend card |
| *dsowv032* | Size-of-wallet strategy variable 9 (sw-sow-09-1)- number of switching preferred cards in the last 4 quarters |
| *dsowv033* | Size-of-wallet strategy variable 10 (sw-sow-10-1) |
| *dsowv034* | Size-of-wallet average pay down percent on revolving retail trades |
| *dsowv035* | Size-of-wallet - no of balance transfers last year |
| *dsowv039* | Size-of-wallet number of balance transfers in last year |
| *dsowv040* | Size-of-wallet number of trades with accelerated revolving bal |
| *dsowv043* | Total revolving balance from the most recent quarter divided by that from the same quarter 1 year ago (sow strategy variable 8) |
| *dsowv044* | Bt's amount in one year |
| *dsumv023* | Percentage of major derogatory actions (dsumv001/dsumv022\*100) |
| *dsumv041* | Debt to income ratio for plst (transformation) |
| *dwoff070* | Number of non-trade writeoffs ever |
| *dwoff083* | Sum of collection amount for collections (if from trade record, it must also be a candidate for major derogatory event processing) |
| *dwoff101* | Number of months since most recent collection (if from trade record, it must also be a candidate for major derogatory event processing) |
| *dwoff105* | Number of collections with a balance greater than 200 (if from trade record, it must also be a candidate for major derogatory event processing) |
| *loc\_max\_rev\_bal* | Maximum line of credit revolving balance amount |
| *loc\_max\_spend* | Credit line on preferred spend card |
| *loc\_max\_trans\_bal* | Maximum line of credit transacting balance amount |
| *max\_revolve\_bal* | Maximum revolving balance amount |
| *num\_bal\_transfer* | Balance transfer count |
| *num\_bal\_transfer\_12m* | Last year balance transfer number |
| *num\_bt\_parker* | Balance parker trade number |
| *num\_line\_decrease* | Line decrease 12 months trade number |
| *num\_line\_increase* | Line increase 12 months trade number |
| *num\_pref\_card* | Number of switching preferred cards in the last 4 quarters |
| *num\_revolve\_accel* | Number of cards with accelerated revolve |
| *num\_revolve\_card* | Revolving card count |
| *num\_spend\_accel* | Number of cards with accelerated spend |
| *num\_trans\_card* | Transacting card count |
| *pay\_pct\_max\_rev\_bal* | Maximum revolving balance pay down percent |
| *sbefaa* | Faa score |
| *scedidx* | Commercial edi |
| *scmbhit* | Comm hit charge |
| *scmposet* | Commercial positive |
| *scmrsket* | Commercial negative |
| *scsr* | Csr scoe |
| *slmxtnbl* | Highest line of credit on a transacting card |
| *smartrevenue* | Smart revenue |
| *snfinact* | Non financial accounts |
| *snrevcrd* | Count of revolving cards (number of commercial card trades on which as per our estimates this person pays less than 50% |
| *sntrncrd* | Transacting card count |
| *snuminq6* | No of inq in 6m |
| *spend\_max\_rev\_bal* | Maximum spend revolving balance amount |
| *spend\_max\_spend* | Spend on preferred spend card |
| *spend\_max\_trans\_bal* | Maximum spend transacting balance amount |
| *srevoemp* | Large biz ind |
| *ssmarttn* | Smart tenure |
| *ssmrtutil* | Smart credit usage |
| *sthinfil* | Thin file ind |
| *stotbal* | Total balance on revolving accounts ≥ $30,000 |
| *strdu3mo* | # Of times more than 90% utilized |
| *strevbal* | Rev balance |
| *tot\_payment* | Total annual external payment amount |
| *tot\_revolve\_bal* | Total revolving balance amount |
| *tot\_spend* | Total annual spend amount |
| *tot\_trans\_bal* | Total transacting balance amount |
| *amt\_bal\_transfer* | Balance transfer amount |
| *amt\_bal\_transfer\_12m* | Last year balance transfer amount |
| *ave\_rev\_pay down* | Average revolving pay down |
| *ave\_spend\_1* | Average spend quarter1 |
| *ave\_spend\_2* | Average spend quarter2 |
| *ave\_spend\_3* | Average spend quarter3 |
| *ave\_spend\_4* | Average spend quarter4 |
| *Cchit* | Commercial hit lending |

1. Analysis of underwriting pathways, U.S. consumer portfolio, 2015 data. [↑](#footnote-ref-1)
2. For applicants denied an AXP account, the model estimates default on external trade lines 12 months after the point of evaluation. [↑](#footnote-ref-2)
3. A second internal model, the Consumer Decision Support System, or CDSS, predicts the likelihood of default by current cardholders on AXP debt. A third internal model, Total Structural Risk, or TSR, predicts the likelihood of default by current cardholders on AXP or external debt. [↑](#footnote-ref-3)
4. Within the U.S. consumer portfolio, ADSS is also used to pre-screen prospective customers for targeted marketing. [↑](#footnote-ref-4)
5. The extension of additional cards to existing AXP cardholders is referred to as cross-selling; because cross-selling occurs on accounts already under management, AXP bases its credit evaluation on the CDSS and TSR models discussed in footnote 3. [↑](#footnote-ref-5)
6. For other internal models, the “decline” decision is not rejecting an application but rejecting a point-of-sale charge. [↑](#footnote-ref-6)
7. The 11 models will be (1) U.S. Consumer Products and Services (CPS); (2) U.S. small business (OPEN); (3) Canada; (4) Australia; (5) Mexico; (6) UK; (7) Japan; (8) India; (9) Argentina; (10) a cluster market for continental Europe (Italy; Germany; France; Netherlands; Spain; Austria; Finland; and Sweden) and (11) a cluster model for Asia Pacific (Hong Kong; Singapore; Thailand; Taiwan; and New Zealand). [↑](#footnote-ref-7)
8. In mid-2016, as this document was written, three models (Canada, Australia and India) used GBM; the remainder used logistic regression. By the end of 2016, AXP expects that all models except Argentina will use GBM. [↑](#footnote-ref-8)
9. For the U.S. consumer portfolio, a “front-end” ADSS model is also used to pre-screen applicants as part of prospect targeting. The “front-end” model is identical to the more widely used “back-end” model, except that it does not have those applicant characteristics made available at the time of application (e.g., income and the card product sought). AXP’s risk group uses the front-end model in consultation with the marketing group to target prospects for card solicitations. During this step, the marketing group also identifies the product to offer and the channel through which it is offered, e.g., email or direct mail. [↑](#footnote-ref-9)
10. An additional category, “pre-approved,” is used only for cross-selling to existing cardholders. Pre-approved is similar to pre-screened, but with a more committed offer of credit. Pre-approved applicants are denied only if they fail to meet ability-to-pay criteria or undergo a major financial change (for example, bankruptcy) between the solicitation and AXP’s receipt of the application. [↑](#footnote-ref-10)
11. The ADSS model used for credit underwriting is internally referred to as the “back-end” ADSS model, to distinguish it from the “front-end” model used to pre-screen applicants in the U.S. consumer portfolio. Unless otherwise noted, the ADSS model described in this document refers to the back-end version. The back-end and front-end versions are identical, except that the back-end has the additional independent variables that become available at the point of application. For example, the back-end ADSS model uses income from the application as well as the marketing channel through which the applicant applied, two variables that were unavailable at the time of pre-screening. [↑](#footnote-ref-11)
12. A third outcome is that the application may be cancelled if AXP does not have sufficient information to process it, even after multiple attempts to collect this information. [↑](#footnote-ref-12)
13. EMVG assigns these categories based on the model’s business use; its potential impact on the company’s financial position and brand; and regulatory expectations. The categories determine the escalation thresholds for alerting leadership that performance is deteriorating. Further details on these models are available in AXP’s internal catalog of risk models, accessible at <https://arc.idn.aexp.com/MDOT/Default.aspx>. [↑](#footnote-ref-13)
14. The Asia Pacific cluster model includes Hong Kong; Singapore; Thailand; Taiwan; and New Zealand. [↑](#footnote-ref-14)
15. The continental Europe cluster model includes Italy; Germany; France; Netherlands; Spain; Austria; Finland; and Sweden. [↑](#footnote-ref-15)
16. An account is also classified as a default if it is cancelled within 12 months of the application because it entered debt collections. This could occur if a cardholder missed his or her first payment and so received a call from AXP debt collections requesting payment. If AXP cannot reach the cardholder after multiple attempts, it will place the account in collections even before it is 90 days past billing. [↑](#footnote-ref-16)
17. As noted earlier, an applicant’s ADSS score may change between the point of pre-screening and the point of underwriting because of financial changes between the two events. Although the model remains substantively the same for both evaluations, the applicant’s actual data is refreshed. [↑](#footnote-ref-17)
18. AXP’s Modeling Strategy Committee requires that each model segment include at least 1,000 defaulters. [↑](#footnote-ref-18)
19. For example, defaults rates often fall after the federal government issues tax refunds. [↑](#footnote-ref-19)
20. Vantage is a credit score created by the three major U.S. consumer credit bureaus (Equifax, Experian, and TransUnion). AXP uses it over FICO, the competitor score, because of the nature of AXP’s vendor agreements. [↑](#footnote-ref-20)
21. AXP offers the Morgan Stanley Credit Card from American Express,® the Schwab Investor Card™ from American Express, and the American Express Platinum Card® for Schwab. [↑](#footnote-ref-21)
22. Note that ADSS does not use any application data that forms a prohibited basis under the Fair Housing Act (FHA) and Equal Credit Opportunity Act (ECOA) (for more on fair lending, see section 10.4.1). [↑](#footnote-ref-22)
23. IDN is an internal relational database with data on account billings, payments, fees, delinquency status, and more. It is typically populated on a cycle-cut basis, meaning account characteristics are recorded each month when the cardholder’s billing cycle concludes, or “cuts.” Thus if an account’s cycle date is October 7, the balance is recorded as of this date and subsequent account activity falls into the November cycle. [↑](#footnote-ref-23)
24. Throughout this document, “development data” refers to the data used to build the model, in contrast to the data used to test or adjust it. [↑](#footnote-ref-24)
25. The use of these four data sets—the (1) development data set, (2) out-of-sample data set, (3) out-of-time data set, and (4) early validation data set—is required by AXP’s Modeling Strategy Committee, or MSC. [↑](#footnote-ref-25)
26. In this process, modelers give precedence to categorical variables over continuous ones, since categorical variables promote stable segments. [↑](#footnote-ref-26)
27. A p-value represents the probability that a result is produced by chance. A p-value threshold of 0.05 means there is a 5 percent chance that the observed result is caused by chance rather than an underlying real-world relationship. The choice of a p-value threshold indicates what weight of evidence modelers will consider conclusive. [↑](#footnote-ref-27)
28. Minimum or maximum truncations are unnecessary for GBM, where outliers will not bias coefficients. [↑](#footnote-ref-28)
29. Significance is measured by the Wald Chi-square statistic, which compares the Wald statistic to a Chi-square distribution. For details of this metric, see Section 16.3: Diagnostics. [↑](#footnote-ref-29)
30. This assumes the default rate is less than one percent. If the default rate is two percent, for example, the perfect model will capture all defaulters in the riskiest two percentiles, and so forth. [↑](#footnote-ref-30)
31. Hamerle, Rauhmeier and Rösch (2003) and Hand (2009) discuss the non-comparability of Gini scores for different portfolios. [↑](#footnote-ref-31)
32. Modelers use the final ADSS score after all post-model adjustments. [↑](#footnote-ref-32)
33. For simple regression, the marginal sensitivity of the dependent variable to a change in an independent variable is measured by the latter’s coefficient. When building a logistic regression model, however, modelers transform the dependent variable and alter independent variables through transformations or capping and flooring. These complications make interpreting the coefficient less straightforward. [↑](#footnote-ref-33)
34. All Decision Science models other than CCAR and Basel fall under the purview of AXP’s Modeling Strategy Committee, or MSC. CCAR and Basel models fall under the CCAR/Basel/Economic Capital Modeling Committee. [↑](#footnote-ref-34)
35. The extent of the validation is proportional to model importance: models classified as critical, significant, or moderate impact require full independent validations. Low-impact models are validated by EMVG or by an independent peer reviewer. [↑](#footnote-ref-35)
36. The frequency of re-validations depends on the use of the model and its importance to AXP. A validation could occur earlier than planned due to material changes to the model or its data. [↑](#footnote-ref-36)
37. At the beginning of this process, the modeling team completes Fair Lending’s Risk Intake Form to record the variables used in the model. Variables are subject to review if they are available to models, regardless of whether they are actually used. This includes both independent and segmentation variables. [↑](#footnote-ref-37)
38. FLPO review is not required for variables that are newly excluded or were approved in previous models. [↑](#footnote-ref-38)
39. Age may be considered when certain conditions are met. Modelers seeking to use age-related variables should discuss these conditions with the FLPO. [↑](#footnote-ref-39)
40. MSC is a sub-committee of AXP’s Individual Credit Risk Committee (ICRC). In this context, “individual” refers to models relating to consumer or small business customers. Separate committees manage models for institutional borrowers. [↑](#footnote-ref-40)
41. At this point, the MSC, along with EMVG and the vice president managing the modeling effort, officially approve a Risk Change Management (RCM) request. The history of RCMs serves as the archive of past changes to the model. They include descriptions of model changes; the justification for these changes; results of implementation tests; and approvals. [↑](#footnote-ref-41)
42. Data treatments include statistical transformations of independent variables as well as floors (minimum values), caps (maximum values), and procedures for imputing missing or zero values. [↑](#footnote-ref-42)
43. Specifically, RIM’s Advanced Risk Capabilities (ARC) team sends the modeling team a Post Implementation Validation (PIV) file with the results. Testing is complete when the modeling team signs and dates this file. [↑](#footnote-ref-43)
44. CMV estimates the five-year discounted cash flow associated with an account. It considers the acocunt’s risk attributes (expected attrition rate, expected default rate, expended spending, expected unpaid balance, and so forth); product economics (weighted APR, variable margin of spending, and rewards cost); and background economic considerations (cost of funds, recover rates, and so forth). [↑](#footnote-ref-44)
45. Lead markets are Australia, Canada, Japan, Mexico, and the UK. [↑](#footnote-ref-45)
46. Segments may be combined for model tracking if they contain fewer than 1000 defaulters. [↑](#footnote-ref-46)
47. The threshold is 5,000 defaulters for consumer and OPEN models, and 2,000 for non-U.S. models. [↑](#footnote-ref-47)
48. More specifically, modelers typically select the three most recent vintages that allow 18 months of future data. [↑](#footnote-ref-48)
49. For larger segments, modelers may perform this analysis at the more granular level of score band. [↑](#footnote-ref-49)
50. The cash usage segment pools cardholders across products (charge, lending, and linked). [↑](#footnote-ref-50)
51. Modelers typically cap the maximum revolving LOC balance to $35,000 to $50,000 and communicate this limit through terms and conditions. This limit is not a committed line of credit and can be suspended at our discretion. [↑](#footnote-ref-53)