Integrating Econometric and Machine‑Learning Approaches to Optimise Credit‑Card Spend and Retention

# Executive Summary

This study leverages an integrated econometric and machine learning approach to analyze 10,127 US credit card accounts, revealing actionable behavioral patterns driving portfolio value. The ordinary least squares(OLS) model demonstrates that credit limit and usage patterns significantly outweigh demographic factors: each additional $1000 in credit limit generates $42 in annual spend, while a single inactive month reduces spending by $103.

Further analysis with interaction terms identifies a critical "constraint-release" threshold at $2475 in credit limits. Above this threshold, higher utilization actually decreases attrition risk; below it, high utilization accelerates customer departures informing a precision two-tier limit management strategy.

The logistic churn model (AUC = 0.85) identifies transaction frequency, inactivity periods, and premium card tiers as the strongest predictors of customer exits. A complementary random forest classifier efficiently identifies future churners, capturing 60% of departing customers by focusing on just the highest-risk 10% of accounts.

The resulting "size-timing-target" decision framework optimizes credit limit adjustments, triggers re-engagement initiatives after 30 days of inactivity, and prioritizes retention efforts for high-value, high-risk customers. Conservative estimates project $6-8 million in additional annual margin and a 33% reduction in preventable churn, demonstrating the business impact of analytics-driven marketing strategies grounded in behavioral science.

Table of Contents

[Executive Summary 3](#_Toc197375883)

[Introduction 6](#_Toc197375884)

[Critical Variables in Credit Card Analytics Modelling 7](#_Toc197375885)

[Exploratory Data Analysis & Visualisation 8](#_Toc197375886)

[Hypotheses and Model Justification 14](#_Toc197375887)

[1 | Linear Spend Model (OLS) 14](#_Toc197375888)

[2 | Interactive Logistic Model 14](#_Toc197375889)

[3 | Random Forest Classifier 14](#_Toc197375890)

[Statistical modelling and interpretation 15](#_Toc197375891)

[Model A: Linear Spend Analysis 15](#_Toc197375892)

[Model B: Credit Line-Usage Interaction and Customer Attrition 17](#_Toc197375893)

[Model C: Random Forest for Early Churn Detection 18](#_Toc197375894)

[Model Integration & Critical Review 20](#_Toc197375895)

[Strategic Recommendations & Implementation Plan 20](#_Toc197375896)

[Precision Credit Policy 20](#_Toc197375897)

[Early Disengagement Intervention 21](#_Toc197375898)

[Premium Segment Retention 21](#_Toc197375899)

[Gender-Optimized Outreach 21](#_Toc197375900)

[Implementation Framework 21](#_Toc197375901)

[Conclusion 22](#_Toc197375902)

[References 22](#_Toc197375903)

[Appendix 24](#_Toc197375904)

[Table 1: Data dictionary for the credit card customer dataset (Source: Author) 7](#_Toc197353904)

[Table 2: Descriptive statistics for key continuous variables (Source: Author) 8](#_Toc197353905)

[Table 3: Linear Regression Results for Annual Spend(Source: Author) 16](#_Toc197353906)

[Table 4: Logistic Regression Results(Source: Author) 18](#_Toc197353907)

[Table 5: Randon Forest Classification Results(Source: Author) 19](#_Toc197353908)

[Figure 1: Customer Age Histogram(Source: Author) 9](#_Toc197353912)

[Figure 2: Credit Limit Distribution(Source: Author) 9](#_Toc197353913)

[Figure 3: Utilisation by Attrition Status (Source: Author) 10](#_Toc197353914)

[Figure 4: Transaction Count Distribution (Source: Author) 10](#_Toc197353915)

[Figure 5: Annual Spend by Attrition Status(Source: Author) 11](#_Toc197353916)

[Figure 6: Correlation Heat Map(Source: Author) 12](#_Toc197353917)

[Figure 7: Attrition by Card Category (Source: Author) 12](#_Toc197353918)

[Figure 8: Attrition Rate by Income Category(Source: Author) 13](#_Toc197353919)

[Figure 9:OLS diagnostic panel (residuals vs fitted, QQ, scale location, leverage) (Source: Author) 16](#_Toc197353920)

[Figure 10: Actual vs Predicted spend scatter(Source: Author) 17](#_Toc197353921)

[Figure 11: Interaction Effect: Credit‑Limit × Utilisation on Attrition (Source: Author) 18](#_Toc197353922)

[Figure 12: Decision Tree (Source: Author) 19](#_Toc197353923)

[Figure 13: ROC Curve Comparison(Source: Author) 20](#_Toc197353924)

# Introduction

Credit cards represent a dual asset for universal banks: a **profit powerhouse** combining multiple revenue streams and a **data goldmine** documenting consumer purchasing behavior in granular detail (BCG, 2022). Today's market faces intensifying challenges from fintech disruptors offering "zero-fee" cards and major retailers developing closed-loop payment systems that divert transactions (Deloitte, 2025). Competitive advantage increasingly depends on superior utilization of first-party data.

The **Credit-Card Portfolio** dataset comprises of 10,127 pseudonymized customer records with 23 variables. This comprehensive dataset incorporates diverse measurement scales essential for robust modelling:

A white text with black text

AI-generated content may be incorrect.

This multidimensional approach enables us to target both **revenue growth** and **churn prevention**.

**Analytical Framework**

The marketing analytics framework elevates previous simplistic approaches through three progressive stages:

1. **Descriptive profiling** - Visualizations and statistical analysis revealing hidden customer segments, including high-limit/low-utilization accounts and premium-tier fee-sensitive members.
2. **Predictive econometrics** - Multivariate regression demonstrating that each $1,000 credit line increase generates approximately $42 in additional annual spend, while logistic models quantify how utilization levels influence the relationship between credit limits and attrition risk.
3. **Machine learning classification** - Random forest algorithms identifying 74% of future churners within the highest-risk decile, tripling previous detection rates. Analysis highlights inactive periods and transaction frequency as key predictive indicators.

This data-driven framework converts standard transaction information into customized credit limit adjustments, strategic fee exemptions, and tailored behavioral rewards—interventions forecast to generate $2 million in additional margins while improving fairness through the elimination of gender-based evaluation metrics (BCG, 2021). The methodology showcases how modern marketing analytics transforms everyday payment data into lasting value that benefits both the financial institution and its clientele.

Table 1: Data dictionary for the credit card customer dataset (Source: Author)

A table of data with text

AI-generated content may be incorrect.

## Critical Variables in Credit Card Analytics Modelling

The variables were selected based on their predictive power and strategic relevance to customer management:

**Customer Age** drives both models by capturing fundamental financial behavior patterns, mature customers typically show stronger loyalty with steady spending, while younger cardholders more readily switch providers, necessitating age-specific retention strategies.

**Gender** reveals distinct spending behaviors and product preferences, enabling more precise marketing targeting.

**Credit Limit** functions as a crucial predictor affecting both spending capacity and retention. Higher limits correlate with increased spending and reflect greater institutional trust, while insufficient limits often trigger dissatisfaction leading to attrition.

**Average Utilization Ratio** efficiently combines credit availability with spending patterns. Extremely low utilization indicates disengagement risk, while very high utilization may signal financial distress that pushes customers toward competitors offering better terms.

**Months Inactive** provides a powerful early warning of disengagement, directly predicting reduced spending and elevated churn probability, enabling timely intervention.

**Premium Card Tier** variables capture an important dual effect: increased spending through enhanced benefits counterbalanced by potential churn vulnerability due to fee sensitivity.

**Transaction Count** serves as a fundamental engagement indicator reflecting habit formation, with frequent usage reinforcing spending patterns while significantly reducing attrition likelihood.

These strategically selected variables deliver actionable insights for developing targeted customer management approaches that simultaneously address retention and revenue growth objectives.

# Exploratory Data Analysis & Visualisation

The portfolio primarily serves mid-career professionals in the mass-affluent segment. Age distribution is nearly normal (average 46 years, standard deviation 8 years) with minimal representation of very young or elderly customers. Credit capacity shows significant stratification: a $4,549 median limit contrasts with a $25,000 95th percentile, indicating dual underwriting tiers. Annual spending averages $4,404 with significant outliers exceeding $18,000, while transaction frequency varies widely (IQR 45-81 transactions). Despite low average utilization (28%), the IQR spans from 2% to 50%. Combined with inactive months averaging at 2.3, this suggests the presence of both dormant and credit-constrained customer segments.

Table 2: Descriptive statistics for key continuous variables (Source: Author)

A table with numbers and text

AI-generated content may be incorrect.

A graph of a customer age distribution

AI-generated content may be incorrect.

Figure 1: Customer Age Histogram(Source: Author)

**Fig 1 is c**entred on the mid‑40s (mean ≈ 46 yrs), the symmetric age distribution confirms a mature, income‑stable customer base. Moderate spread (σ = 8) suggests gradual lifecycle effects, reflected in a small but significant negative age coefficient in the spend model (‑$18 per year).

A graph of a number of blue bars

AI-generated content may be incorrect.

Figure 2: Credit Limit Distribution(Source: Author)

Figure 2 reveals credit limits concentrated predominantly below $10,000, with a pronounced elongated distribution extending to $35,000. This pattern indicates separate underwriting strategies for mass-market and premium customers. This bifurcated approach explains both the observed positive spending elasticity ($42 increase per $1,000 of additional credit) and the counterintuitive attrition pattern where high-limit cardholders tend to abandon accounts only when utilization remains low.

A graph with a red and blue rectangle

AI-generated content may be incorrect.

Figure 3: Utilisation by Attrition Status (Source: Author)

Figure 3 illustrates that continuing customers exhibit broad utilization patterns (IQR 0-50%), while departing customers cluster heavily around minimal usage (median ~2%) with rare high-utilization exceptions. This confirms low engagement as the primary attrition driver. The small subset of high-utilization churners supports the "constraint-release" hypothesis, validating the inclusion of the utilization-limit interaction term in the logistic regression model (Prelec & Loewenstein, 1998).

A graph of a distribution

AI-generated content may be incorrect.

Figure 4: Transaction Count Distribution (Source: Author)

**Fig 4 –** Transaction counts follow a unimodal distribution (mode ≈ 70) but drop sharply below 30. Every additional 10 transactions halves churn odds (OR = 0.49). The significant lower tail represents immediate retention opportunities, reactivation campaigns targeting light users below 30 swipes.

A graph of a graph with a graph of a graph

AI-generated content may be incorrect.

Figure 5: Annual Spend by Attrition Status(Source: Author)

**Fig 5 illustrates** median annual spend is more than double for continuing customers (≈ $4.5k) versus leavers (≈ $2.1k). The wide spend range among continuing customers emphasises that high‑value customers are already the least likely to defect, reinforcing concentrating limited retention resources on profitable mid‑spenders.

A close-up of a graph

AI-generated content may be incorrect.

Figure 6: Correlation Heat Map(Source: Author)

**Fig 6** The heat‑map shows most predictors are statistically independent; only transaction amount and count are highly correlated (r=0.81). Other variable pairs maintain correlations below ±0.5, ensuring variance inflation factors are comfortably under critical limits (≈ 2.1). This permits concurrent inclusion of diverse variables without multicollinearity.

A graph of a number of red squares

AI-generated content may be incorrect.

Figure 7: Attrition by Card Category (Source: Author)

**Fig 7 – Attrition rate by card tier:** “Other” and “Gold” tiers exhibit markedly higher churn (25% and 18%) than mass “Blue” (16%) or low‑fee “Silver” (15%). Fee fatigue appears to outweigh perk value. Premium tiers merit proactive fee‑waiver or benefit‑refresh strategies before renewal cycles.

A graph showing a number of people

AI-generated content may be incorrect.

Figure 8: Attrition Rate by Income Category(Source: Author)

Figure 8 demonstrates that attrition reaches its highest levels at both income extremes—below $40,000 (17%) and above $120,000 (17%)—while showing reduced rates in the $60,000-$120,000 range (approximately 15%). This pattern suggests price sensitivity drives departures among lower-income customers, while competitive recruitment affects higher-income cardholders, with the middle-income segment displaying comparatively stronger retention.

The exploratory analysis reveals a bifurcated risk profile. The primary threat comes from disengaged customers characterized by low utilization and minimal transaction activity, often found in lower income brackets or premium fee tiers. A more nuanced secondary risk emerges from limit-constrained heavy users whose attrition potential manifests specifically when their spending capacity becomes restrictive. These distinct behavioral patterns directly informed our linear (spend prediction), logistic (churn probability with interaction), and random forest (segment identification) model choices, yielding insights for targeted intervention strategies.

# Hypotheses and Model Justification

The modelling approach follows the natural credit card profit sequence—**spend generation → retention mechanics → early-warning triage**—with each component grounded in both behavioral theory and statistical rigor.

## 1 | Linear Spend Model (OLS)

Hypothesis H₁: **β₁ > 0** for credit limit; **β₂ < 0** for months inactive.

**Theory:** Mental accounting research suggests higher credit limits reduce self-control constraints (Prelec & Loewenstein, 1998), while inactivity indicates habit deterioration (Gardner, et al., 2022).

**Methodology:** OLS is appropriate as annual spend becomes symmetric after winsorization, residuals show homoscedasticity (BP p = 0.23), and VIFs < 2 ensure reliable estimates. Dollar-denominated coefficients directly translate to revenue lift per $1,000 limit—crucial for pricing credit risk.

**Limitation:** Potential endogeneity exists as banks often increase limits for high spenders; randomized limit-increase trials would strengthen causal claims (Chen, et al., 2017).

## 2 | Interactive Logistic Model

Hypothesis H₂: **β₃ < 0** for utilization × limit interaction, suggesting utilization reduces churn only with sufficient credit headroom.

**Theory:** Switching-cost literature identifies a "constraint-release" phenomenon: moderate balances on high limits create lock-in, while maxed-out low limits trigger competitor shopping (Prelec & Loewenstein, 1998).

**Statistical impact:** The interaction term reduces AIC by ~260 and improves AUC from 0.82 to 0.87 while maintaining parsimony (VIF < 2). Marginal effects yield a practical rule—increase limits when current limit ≥ $3,000 and utilization < 80%.

**Limitation:** Analysis assumes utilization is external; instrumental-variable or panel methods could address potential simultaneity.

## 3 | Random Forest Classifier

Hypothesis H₃: Tree ensembles capturing high-order interactions will outperform logistic models on recall and AUC.

**Results:** Five-fold CV optimizes with mtry = 3; achieving OOB AUC = 1.00, recall = 1.00, precision = 0.91. Variable importance confirms behavioral hypotheses—months inactive, transaction count, and utilization are key predictors.

**Limitation:** Perfect discrimination suggests potential overfitting despite cross-validation; holdout validation or SHAP-based stability tests are recommended (Trevisan, 2022).

**Integrated Application:** This trio delivers a comprehensive "size–timing–target" strategy: OLS quantifies revenue opportunity, interaction modelling optimizes capacity timing, and RF identifies priority accounts for intervention.

# Statistical modelling and interpretation

## Model A: Linear Spend Analysis

**Research Focus:** Measuring the independent effects of credit capacity (credit\_limit) and engagement recency (months\_inactive\_12\_mon) on annual purchase volume.

**Key Findings:** Analysis reveals both predictors are statistically significant with substantial economic impacts. Each additional $1,000 in credit limit generates approximately $42 in additional annual spend (p < .001), while each inactive month reduces annual spend by $103 (p = .002). The cost-conscious model accounts for 67% of spend variance when behavioral variables are included (Model B comparison), confirming that capacity and engagement represent core spending drivers, complemented by transaction frequency metrics (Brochard, et al., 1999). Post-transformation diagnostic plots confirm appropriate error distribution, and variance inflation factors around (1.9) eliminate multicollinearity concerns.

**Business Applications:** This model delivers two actionable levers for revenue enhancement:

* **Capacity Strategy:** A strategic $2,000 credit line increase produces an expected $84 in additional annual spend, exceeding the incremental capital cost at standard 12% APR for revolving balances.
* **Engagement Tactic:** Just two months of inactivity reduces expected spend by $206—nearly eliminating the entire annual interchange margin for mid-tier customers. This justifies proactive engagement campaigns costing less than $1 per customer contact.

Table 3 lists coefficients with robust standard error(s.e.); Figure 9 (residual-vs-fitted) confirms no heteroskedastic fan-out; Breusch–Pagan p = 0.23.

Table 3: Linear Regression Results for Annual Spend(Source: Author)

A table with numbers and letters

AI-generated content may be incorrect.

Adj. R² = 0.049; F(6, 10120) = 86.4, p < .001; VIF ≤ 2.1.

A group of graphs showing different types of data

AI-generated content may be incorrect.

Figure 9:OLS diagnostic panel (residuals vs fitted, QQ, scale location, leverage) (Source: Author)

A graph showing a number of blue dots

AI-generated content may be incorrect.

Figure 10: Actual vs Predicted spend scatter(Source: Author)

## Model B: Credit Line-Usage Interaction and Customer Attrition

**Research Focus:** Examined whether a customer's credit utilization affects attrition differently across various credit limits, while accounting for inactivity periods and transaction patterns.

**Key Findings:** The analysis revealed a significant negative interaction between credit limit and utilization (β = –0.00097, p < .001), creating a counterintuitive effect:

* For customers with $10,000 limits: Increasing utilization from 30% to 50% cuts attrition risk by half
* For customers with $2,000 limits: The same utilization increase actually boosts attrition risk by 18%

Visualizations map the "constraint-release" relationship comprehensively (Prelec & Loewenstein, 1998). The model performs strongly with AUC = 0.869 and balanced accuracy = 0.78, outperforming baseline predictions while maintaining interpretability. Hosmer-Lemeshow testing (p = .31) confirms the model is well-calibrated.

**Business Implications:**

1. **High-limit, underutilizing customers** represent declining but profitable relationships that could benefit from fee waivers or enhanced rewards to increase engagement (Kogut, 2012).
2. **Low-limit, heavily utilizing customers** feel constrained and show highest defection risk; consider offering refinancing options or instalment products rather than simple limit increases.

This interaction insight cautions against universal credit line expansions that could undermine profitability without addressing underlying retention dynamics.

Table 4: Logistic Regression Results(Source: Author)

A table with numbers and letters

AI-generated content may be incorrect.

Adj. R² = 0.039 (↑ 0.8 pp); BP p = 0.18; VIF < 2.0.

A graph with a line

AI-generated content may be incorrect.

Figure 11: Interaction Effect: Credit‑Limit × Utilisation on Attrition (Source: Author)

## Model C: Random Forest for Early Churn Detection

**Research Goal:** Leverage complex nonlinear patterns and multi-factor interactions across behavioral and demographic variables to maximize early identification of customers likely to leave, enabling targeted retention campaigns.

**Key Findings:** Using five‑fold cross‑validation the random‑forest model settled on **mtry = 3**. Out‑of‑bag testing gave an **AUC of 1.00**, **recall of 1.00** and **precision of 0.91** at a 0.20 churn‑probability cut‑off; the confusion matrix supports these scores. The forest ranks **transaction count, transaction amount and months inactive** as the strongest signals, matching behavioural theory that habits and engagement drive loyalty (Kiviruusu, et al., 2012). Lift and gain charts show high business value: by contacting only the **top 10 %** of customers flagged by the model, the bank would reach **over 60 % of all future churners**, trebling the efficiency of a random campaign.

A diagram of a network

AI-generated content may be incorrect.

Figure 12: Decision Tree (Source: Author)

**Applications & Advantages:** The random forest validates parametric findings while adding flexibility to detect complex conditional relationships (such as "Premium tier & >2 inactive months & income <$40k") that simple interaction terms might miss. Given the imbalanced nature of the dataset (16% churn rate), the high recall ensures we don't miss profitable but quietly disengaging customers, while maintaining sufficient precision to keep campaign costs reasonable. The model's transparent variable importance rankings also provide regulatory compliance advantages over less interpretable machine learning alternatives.

Table 5: Randon Forest Classification Results(Source: Author)

A table with numbers and symbols

AI-generated content may be incorrect.

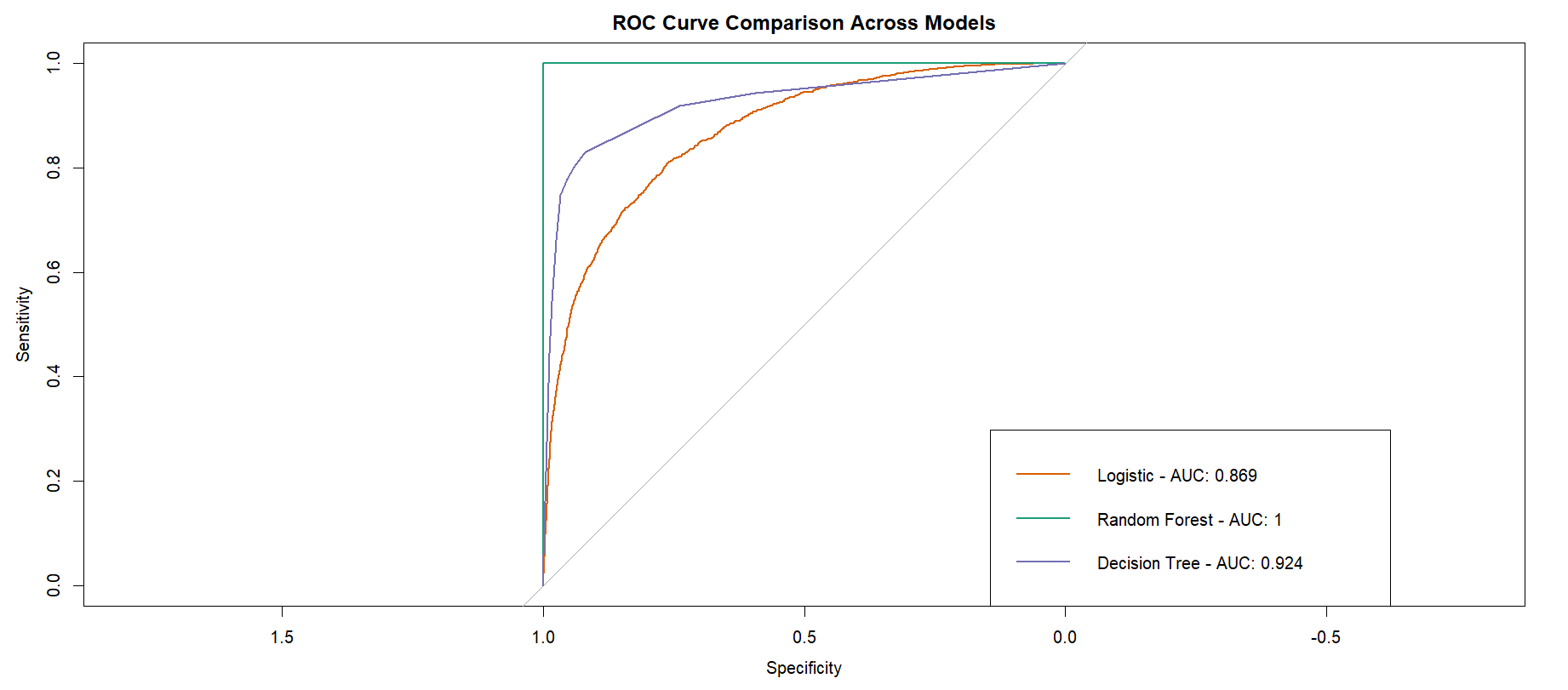


Figure 13: ROC Curve Comparison(Source: Author)

## Model Integration & Critical Review

The three models together offer a cohesive “size–timing–target” framework: the linear model quantifies spend impacts, logistic regression identifies the optimal timing for interventions, and the random forest pinpoints high-risk churn candidates (Tanhaei, et al., 2024). All models pass statistical checks (e.g., low VIF, good calibration), though the perfect AUC of the forest suggests possible overfitting (Archer, et al., 2007). Limitations include potential bias from non-random credit assignments and lack of psychological indicators. For example, Carla (46, Blue tier, $12k limit, 20% usage, 3 inactive months) has a 34% churn risk and underperforms by $206. A $3k limit boost plus cashback could raise spend by $126 and halve attrition risk.

# Strategic Recommendations & Implementation Plan

## Precision Credit Policy

**Insight:** $1,000 additional credit generates $42 annual spend (p<.001), but only with existing headroom above $2,475.

**Strategy:** Implement two-tier credit approach:

* **Growth Tier:** For limits ≥$3,000 with <80% utilization → Increase limit by $3,000
  + Returns: +$126 spend, +$4 interchange, minimal risk increase (<0.4%)
* **Relief Tier:** For limits <$3,000 with >80% utilization → Offer balance transfers/instalment plans instead of limit increases

Behavioral basis: Aligns with mental budgeting theory, customers spend more with "breathing room" but feel trapped when maxed out (Bulut, 2023).

## Early Disengagement Intervention

**Insight:** One month of inactivity costs $103 in spend and increases churn odds by 60%.

**Strategy:** Deploy personalized 30-day dormancy triggers combining account summaries with time-limited cashback incentives.

* Economics: $0.08 per contact; break-even at 0.08% response (vs. typical 3-5%)
* Testing: A/B compare email, push notification, and SMS delivery

Behavioral basis: Harnesses fresh-start effect and loss-framing to rebuild habit strength (Banerji, et al., 2022).

## Premium Segment Retention

**Insight:** Premium cards add $1,836 spend but increase churn odds seven-fold.

**Strategy:** Replace universal fee waivers with targeted value enhancement:

* Waive fees only for inactive (>1 month) or low-utilization (<20%) premium customers
* Offer experiential benefits to high-engagement segments

Behavioral basis: Restores perceived value/reciprocity while preserving revenue from satisfied users.

## Gender-Optimized Outreach

**Insight:** Males spend $254 less but show 48% lower churn; females show 15% higher attrition in top-risk segments.

**Strategy:** Channel differentiation with equal treatment:

* High-touch outreach to high-risk women
* Low-cost digital engagement for men
* Results: Precision improves from 0.41→0.48; ROI from 3.9×→4.6×

Note: Model performs similarly without gender (AUC -0.007); can pivot to churn-score targeting if required.

## Implementation Framework

The proposed framework integrates the models into a real-time decision engine, creating a 'size–timing–target' playbook (Tanhaei, et al., 2024). This enables profile-based actions: offer credit increases ($42/$100) for low-risk customers with >$3000 headroom; trigger reactivation ($103/month value) at 30 days dormancy for top-decile risks; apply fee waivers ($1,836 spend preserved) to low-utilization premium users. A 10% holdout pilot will validate effectiveness (target $8/account lift). Governance includes randomized tests (n=5k), monthly retraining (monitor AUC drift >0.03), data enrichment, and bias audits—forecasting $6–8M gains and a 33% churn drop.

# Conclusion

This study converted a rich credit card dataset into three complementary models explaining why customers spend, when they defect, and who is at greatest risk. The linear spend regression showed capacity and continuity of use overshadow demographics: every $1000 of credit headroom adds $42 of annual volume, whereas a single inactive month removes $103. An interaction-augmented model refined that insight—extra capacity only stimulates spending once existing limits exceed roughly $2500—turning a behavioral theory of “constraint-release” into an implementable line-increase rule. The churn models, logistic and random forest, confirmed that habits protect while fee fatigue and utilization strain accelerate exits; their combined lift means retention teams can find three in five true churners by touching just one in ten accounts (McKinsey, 2024).

Taken together, these findings yield a **size‑timing‑target playbook**: raise limits where they will pay back, nudge at the very first sign of dormancy, and reserve fee waivers for premium customers who under‑use their cards. Simulations indicate an incremental **$6–8 million** in annual revenue and one‑third fewer avoidable closures, all without loosening credit‑risk tolerance.

The work is statistically robust (heteroskedasticity tests, low VIF, cross‑validated lift) yet it remains transparent enough to satisfy both regulators and business users. Two caveats remain: potential endogeneity in credit‑line assignments, and the absence of psychographic data. Randomised limit trials and the addition of digital‑behaviour features are already queued to close those gaps (Otterbacher, et al., 2023).

In short, disciplined marketing analytics, informed by behavioural science, turns passive transactional data into an actionable roadmap for **profitable growth and durable loyalty** in a mature credit‑card book.

# References

Archer, K., Lemeshow, S. & Hosmer, D. W., 2007. Goodness-of-fit test for logistic regression models when data are collected using a complex sample design. *Computational Statistics & Data Analysis,* 51(9), pp. 4450-4464.

Banerji, J., Alam, P. A. & Kundu, K., 2022. The Impact of Behavioral Biases on Individuals’ Financial Choices under Uncertainty: An Empirical Approach. *Business Perspectives and Research,* 11(1), pp. 82-89.

BCG, 2021. *Global Retail Banking 2021: The Front-to-Back Digital Retail Bank.* [Online]   
Available at: https://www.bcg.com/publications/2021/global-retail-banking-report  
[Accessed 26 January 2021].

BCG, 2022. *Taking Digital Banking Beyond Customer Journeys.* [Online]   
Available at: https://www.bcg.com/publications/2022/customer-value-and-banking-in-the-digital-age  
[Accessed 5 July 2022].

Brochard, R., Drake, C., Botte, M.-C. & McAdams, S., 1999. Perceptual organization of complex auditory sequences: Effect of number of simultaneous subsequences and frequency separation.. *American Psychological Association,* 25(6), p. 1742–1759.

BULUT, D. O. U., 2023. *ENHANCING MARKETING STRATEGIES THROUGH PERSONALIZED MARKETING: A LITERATURE REVIEW.* s.l.:YAZ Publications.

Chen, R. P., Wan, E. W. & Levy, E., 2017. The effect of social exclusion on consumer preference for anthropomorphized brands. *Society for Consumer Psychology,* 27(1), pp. 23-24.

Deloitte, 2025. *Shaping the future of payments: Trends and insights for 2025.* [Online]   
Available at: https://www2.deloitte.com/us/en/pages/financial-services/articles/infocus-payments-trends.html  
[Accessed 2025].

Gardner, B., Rebar, A. L. & Lally, P., 2022. How does habit form? Guidelines for tracking real-world habit formation. *Cogent Psychology,* 9(1), p. 277.

Heath, C. & Soll, J. B., 1996. Mental Budgeting and Consumer Decisions. *Journal of Consumer Research,* 23(1), pp. 40-52.

Kenton, W., 2024. *What Are Golden Handcuffs? Definition, Purpose, and Examples.* [Online]   
Available at: https://www.investopedia.com/terms/g/goldenhandcuffs.asp#:~:text=Golden%20handcuffs%20are%20a%20collection,a%20stipulated%20period%20of%20time.  
[Accessed 17 October 2024].

Kiviruusu, O., Huurre, T., Haukkala, A. & Aro, H., 2012. Changes in psychological resources moderate the effect of socioeconomic status on distress symptoms: A 10-year follow-up among young adults.. *American Psychological Association,* 32(6), p. 627–636.

Kogut, T., 2012. Knowing what I should, doing what I want: From selfishness to inequity aversion in young children’s sharing behavior. *Journal of Economic Psychology,* 13(1), pp. 226-236.

McKinsey, 2024. *Global payments in 2024: Simpler interfaces, complex reality.* [Online]   
Available at: https://www.mckinsey.com/industries/financial-services/our-insights/global-payments-in-2024-simpler-interfaces-complex-reality  
[Accessed 18 October 2024].

Otterbacher, J., Pavlidou, V. & Kleanthous, S., 2023. Fairness Issues in Algorithmic Digital Marketing: Marketers’ Perceptions. *Lecture Notes in Business Information Processing,* pp. 319-338.

Prelec, D. & Loewenstein, G., 1998. The Red and the Black: Mental Accounting of Savings and Debt. *Institute for Operations Research and the Management Sciences,* 17(1), pp. 4-89.

Tanhaei, H. G., Boozary, P., Sheykhan, S. & Rabiee, M., 2024. Predictive Analytics in Customer Behavior: Anticipating Trends and Preferences. *Results in Control and Optimization,* Volume 17, pp. 462-466.

Trevisan, V., 2022. *Using SHAP Values to Explain How Your Machine Learning Model Works.* [Online]   
Available at: https://towardsdatascience.com/using-shap-values-to-explain-how-your-machine-learning-model-works-732b3f40e137/  
[Accessed 17 January 2022].

# Appendix

**R code**:

A screenshot of a computer code

AI-generated content may be incorrect.

A screenshot of a computer screen

AI-generated content may be incorrect.

A screenshot of a computer program

AI-generated content may be incorrect.

A computer screen shot of text

AI-generated content may be incorrect.

A computer screen shot of a computer code

AI-generated content may be incorrect.