Buy-Now-Pay-Later Risk Analysis

Tony Ngari, Haiwei Du, He Jiang, Turner McLaurin

December 9, 2025

Contents

§0. Introduction

§0.1. Background

Buy-Now-Pay-Later (BNPL) has rapidly emerged as a dominant force in the U.S. consumer credit market, processing over $100 billion in gross merchandise volume in 2024 alone [She21]. BNPL has capitalized on current macroeconomic pressures, including high inflation and stagnant real wages, that have strained household budgets and fueled a demand for flexible, alternative financing. However, as the industry scales to serve 91 million users in the US, and over 380 million users worldwide, understanding the underlying risks is critical [Res24].

§0.2. Service Mechanics and Risk

Buy Now, Pay Later platforms operate by offering consumers instant, short-term, unsecured installment credit directly at checkout [Con23]. When a shopper selects BNPL, the provider pays the merchant immediately (usually the full purchase amount, minus a fee), then collects repayment from the consumer over a series of scheduled installments. Because approvals occur within milliseconds and require minimal upfront data, the experience is designed to be nearly frictionless, expanding consumer liquidity at the exact moment of purchase.

For merchants, the value proposition is commercial rather than financial, since BNPL tends to increase conversion, reduce cart abandonment, and elevate Average Order Value, often by 20% - 40% [Res24]. This "stimulus effect" on demand is why merchants willingly subsidize the cost of what appears to consumers as free credit. For brands competing in saturated digital markets, BNPL functions not just as a payment option but as a growth lever.

BNPL providers make money primarily through the Merchant Discount Rate (MDR), typically 3% - 8% per transaction; substantially higher than standard credit-card interchange and intended to cover financing, risk, and the merchant-side commercial uplift [Con23]. Additional revenue sources include consumer interest on longer-duration payment plans, late fees designed mainly as behavioral incentives, and occasionally merchant marketing or referral partnerships.

However, the business is exposed to several loss vectors. The most significant is credit losses: because BNPL accepts minimal application data, some users will default outright or accumulate delinquent balances [Fit22]. Providers also face funding costs, debt or equity capital used to advance merchant payments, as well as payment-processing fees when collecting each installment through card networks. In periods of macro stress or tightening capital markets, these costs can compress margins rapidly.

Risk management is uniquely challenging for BNPL because the model depends on real-time decisions with sparse information. Traditional lenders rely on rich credit histories, stable income data, and days or weeks of underwriting time. BNPL must assess risk in under a second, often for first-time users with thin credit files. As a result, providers deploy adaptive, behavioral scoring models based on repayment patterns, transaction frequency, and purchase magnitude [Ris25]. These systems effectively "underwrite forward," using a consumer's initial low-value transactions to calibrate their credit limits and eligibility for higher-value purchases. The objective is to minimize losses not by preventing all risk, but by sequencing exposure intelligently, scaling credit only as the user demonstrates reliability. This tension between frictionless growth and disciplined risk controls defines the core operational challenge of the BNPL model.

§0.3. Objective

Because BNPL lenders issue unsecured credit with limited visibility into a user's broader indebtedness, while competing on instant, frictionless approvals, the sector may be structurally overexposed to hidden default risk. Determining the extent of this exposure, and evaluating how resilient BNPL business models truly are, is critical given the growing ubiquity of these services. This concern directly motivates the three core objectives of the project.

1. Analyze and model BNPL consumer behavior to understand spending patterns, repayment habits, and default tendencies.

2. Develop quantitative risk models to measure and visualize the financial exposure and vulnerability of BNPL providers.

3. Simulate real-world market scenarios to test business model resilience and evaluate the long-term viability of BNPL firms under varying economic conditions.

§1. Methodology

§1.1. Overview

To analyze credit risk, three distinct predictive models, *Logistic Regression*, *Random Forest*, and *Gradient Boosting*, were employed to estimate the probability of loan default. These models assessed risk as a function of key covariates, including loan-to-income ratio, borrower income, loan amount, loan term, nominal interest rate, income verification status, and selected borrower characteristics. Model fit and performance were evaluated using likelihood-based criteria.

The simulation employed a 24-month Monte Carlo financial model to assess BNPL viability across 36 combinations of economic conditions and user-base scales. By simulating revenue, costs, credit losses, and cash flow under each scenario, the model evaluates whether scale improves unit economics and how macro environments shape business sustainability, using parameters grounded in real BNPL disclosures and recent economic data.

§1.2. Risk Analysis Model

1.2.1. Data Preparation

The data used in this study are anonymized consumer "pay later" loans from a digital lender, obtained from the *Paylater Loan Data* on Kaggle [Ibi25]. The raw dataset contains approximately 160,000 loan records and 28 variables, including borrower income and basic characteristics, loan amount, nominal interest rate, contractual term (months), limited prior borrowing information, and performance indicators. The main outcome is a binary indicator of whether a loan ever entered default.

Data preprocessing and feature construction, analysis were performed in Python. Variables were renamed to descriptive labels (e.g., borrower income, loan amount, interest rate, loan term in months, income verification status, settlement timing in days) and restricted to those relevant to the research questions. Core quantitative fields (loan amount, interest rate, loan term, borrower income, an existing affordability ratio based on loan amount and income, prior maximum loan amount, default indicator, and settlement timing) were converted to numeric form, and loans missing any of these critical fields were excluded. Binary indicators (such as income verification, whether the first payment was missed, and default status) were standardized to 0/1.

1.2.2. Feature Selection

Feature selection was carried out in a few clear steps in the Jupyter notebook. First, the raw Kaggle dataset was reduced to variables that are directly related to credit risk and affordability. This meant keeping core loan terms (amount, term, interest rate), borrower income and its verification status, simple demographics, basic borrowing history, and a small set of repayment-burden measures. Columns that were not linked to these ideas, or that did not support the research questions, were removed at this stage.

Next, the notebook used a simple "signal" check for numeric variables. For each candidate numeric feature, it computed a Kolmogorov–Smirnov (KS) statistic comparing the distribution of that feature for defaulted and non-defaulted loans. Variables with a small KS value or very few non-missing observations were dropped. Variables with larger KS values were kept as "strong numeric drivers" and examined more closely. For these drivers, the notebook produced a correlation heatmap and default-rate-by-decile plots (with approximate confidence intervals). These plots showed that variables such as settlement days, payment ratio, loan-to-income ratio, loan amount, term in months, and maximum prior amount all have clear and stable relationships with default risk.

Based on this analysis, the final feature set used in the models includes these strong numeric drivers, a small number of engineered features (such as payment-to-income ratio, utilization relative to a borrower-specific maximum, and simple interactions between rate, term, and amount), and a few key categorical indicators (for example, loan purpose and basic borrower characteristics). Variables that showed weak signal in the KS screening, very little variation, or high redundancy with stronger predictors were excluded from the modeling pipeline and used only for descriptive purposes. The stage-specific feature sets for the registration, early-stage, and established-user models in the simulation are structured subsets or extensions of this same core feature pool, aligned with the information that a BNPL lender would realistically observe at each point in the customer life-cycle.

1.2.3. Model Training

All models are fitted on the cleaned and engineered dataset using a common train–test split and the same preprocessing steps. Numerical variables are standardized using z-scores so that each has a mean of 0 and a variance of 1. Categorical variables such as gender, loan purpose, and residence status are converted to one-hot encoded indicators, with unseen categories handled safely. The resulting feature matrix X combines all numeric and encoded categorical variables, and the target vector y is the 0/1 default indicator. The data are split into a training set (80%) and a test set (20%), stratified by default status to preserve the overall default rate. These steps are implemented using a ColumnTransformer within scikit-learn Pipeline objects, ensuring preprocessing is applied consistently during both training and evaluation.

On this prepared data, we estimate three main classifiers. First, a logistic regression classifier with L2 regularization and `class weight` = *balanced* is used to give extra weight to default cases and address class imbalance. This is a more flexible version of the simpler logistic regression estimated in R, since it can handle a larger number of predictors and helps stabilize coefficients when predictors are correlated. Second, a random forest classifier is fitted with many trees ($n_{\text{estim}} = 500$), a minimum leaf size of 2, and `class weight` = *balanced subsample*. This model captures non-linear relationships and interactions between variables without needing to specify interaction terms by hand. Third, a histogram-based gradient boosting classifier is estimated. This model builds boosted decision trees on binned features, which is computationally efficient and often performs well on tabular data. All three models use the same preprocessing pipeline and the same train-test split to make their performance comparable.

Before fitting these final models, we conduct signal-driven exploratory data analysis to screen variables. For each numeric variable, I compute a Kolmogorov–Smirnov (KS) statistic to measure how different its distribution is between defaulted and non-defaulted loans, and we focus on those with the largest KS values as "strong numeric drivers." We then examine a correlation heatmap among these strong drivers and the default indicator, and for each such variable (for example, `settle days`, `payment ratio`, `loan income ratio`, `term months`, `amount`, and `max amount taken`, we plot default rates by decile with approximate confidence intervals. These plots show that variables related to repayment behavior and timing (`settle days`, `payment ratio`, `loan income ratio`) and to loan size and maturity (`amount`, `term months`, `max amount taken`) provide meaningful separation between defaulting and non-defaulting loans.

Finally, to interpret the models in economic terms, we run simple scenario-based simulations using the fitted classifiers. We construct synthetic borrower profiles that vary in affordability (loan-to-income and payment-to-income ratios), term length, and interest rate, while holding other characteristics at typical values such as median income and age, a common loan purpose, and verified income status. For each profile, the models produce predicted default probabilities. Comparing these predictions across low, medium, and high affordability buckets shows how changes in affordability and loan design translate into changes in predicted default risk. The detailed implementation of these simulations is provided in the accompanying notebook.

§1.3. Simulation

Our analysis uses a forward-looking, month-by-month financial simulation to evaluate the viability and resilience of BNPL business models under varying economic and scale conditions. The methodology focuses on replicating how a real-life BNPL balance sheet and P&L evolve over time under varying conditions. For each scenario, we initialize a starting user base, capital level, and loan book, then iterate through a 24-month horizon where Gross-Merchandise Volume (GMV) generation, default losses, recoveries, capital costs, operating expenses, and user acquisition dynamics update the simulated financial position. This procedure allows us to observe how cash runway, profitability, and solvency respond to shifts in macroeconomic conditions or scale. All specific parameter values (e.g., economic scenarios, cost assumptions, user behavior estimates) and supporting tables are provided in the appendix.

§2. Results

§2.1. Risk Analysis Model Evaluation

Variance inflation factors were computed for all predictors in the final default model. All VIF values were well below commonly used thresholds with a maximum VIF of 1.048 (all under 3), indicating that multi-collinearity was not a concern in this analysis.

Table 1: ROC-AUC, PR-AUC, and Brier Scores

Model	ROC-AUC	PR-AUC	Brier Score
logistic	0.896	0.847	0.1117
random_forest	0.898	0.846	0.0980
hist_gboost	0.907	0.857	0.0901

The histogram gradient boosting model is the best performer on all three metrics. Its ROC-AUC above 0.90 means that, in randomly chosen pairs of loans (one that defaults and one that does not), the model assigns a higher predicted default probability to the defaulted loan in about 90% of cases. The Precision–Recall AUC is also high, which is important because default is less frequent than non-default. The lower Brier score indicates that predicted probabilities are, on average, closer to the true 0/1 default outcomes.

We also examine calibration plots for each model, which compare average predicted probabilities to observed default rates in probability bins. All models are reasonably well calibrated in the middle range of predicted probabilities. The gradient boosting model, in particular, shows good average calibration, consistent with its lower Brier score.

Finally, we compute lift tables, which group loans into score bins and report the default rate and the cumulative share of all defaults captured. For each model, the bins with higher model scores have higher observed default rates than bins with lower scores, indicating meaningful separation. Although the exact lift pattern differs across models, the gradient boosting classifier tends to concentrate defaults more strongly in higher-score bins, which is consistent with its higher ROC-AUC and Precision–Recall AUC.

Overall, these results show that default risk is not purely random. When rich affordability measures, term and rate information, and simple interactions are included, default probabilities can be predicted with high out-of-sample accuracy. The weaker performance of the original simple logistic regression appears to come mainly from the limited feature set, rather than from an inherent lack of signal in the data.

§2.2. Simulation

The simulation produces an unambiguous outcome: every BNPL configuration tested becomes insolvent within the 24-month horizon. Although survival times vary across scale levels and economic environments, the overall pattern is remarkably consistent. All modelled firms, from small 10,000-user entrants to 50-million-user market leaders, exhaust their capital well before reaching break-even. The results suggest that BNPL failure is not a function of poor execution or adverse conditions but reflects underlying structural economics inherent to the model.

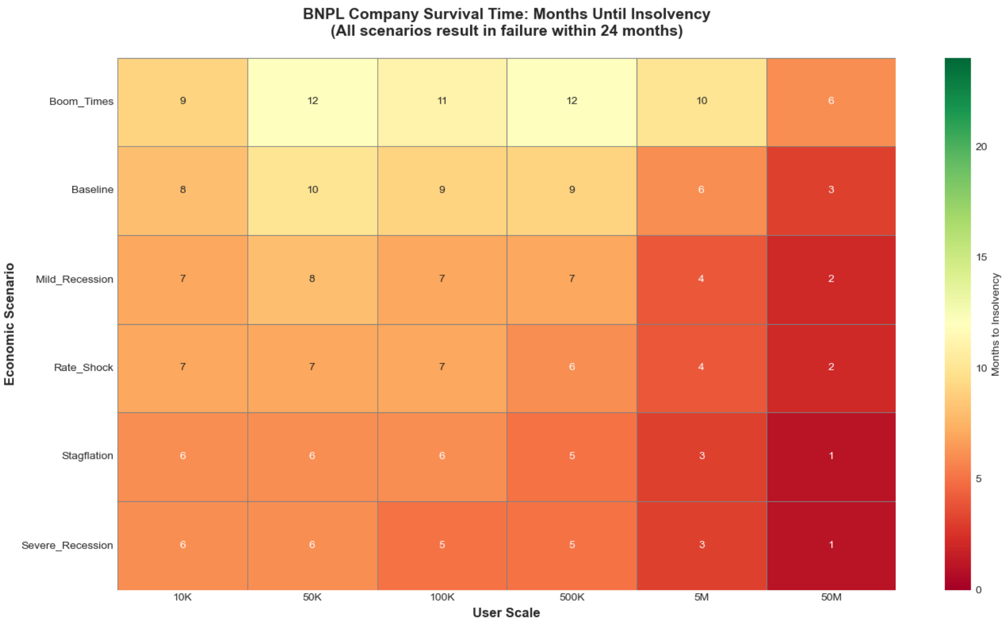

Figure 1: BNPL Survival Time

The insolvency heatmap in Figure 1 summarizes performance across all 36 combinations of user scale and macroeconomic assumptions. Even a cursory glance shows the absence of any configuration that survives a full 24 months. Firms operating in boom conditions, marked by high GMV growth, low default rates and favorable funding costs, last marginally longer, typically around 10 to 12 months. By contrast, firms operating in recessionary or stagflationary environments become insolvent in as little as 5 to 7 months.

What is most striking is how little scale changes the trajectory. Moving from 10000 users to 50 million users, a 5000-fold increase in customer base and a forty-fold increase in starting capital, extends survival by only a few additional months. At no point does increased scale introduce meaningful operating leverage. Costs scale in lockstep with GMV, meaning that larger user bases simply produce larger absolute losses rather than improved margins. The heatmap shows this clearly: insolvency is inevitable across all scenarios, with variation limited to the speed rather than the direction of decline.

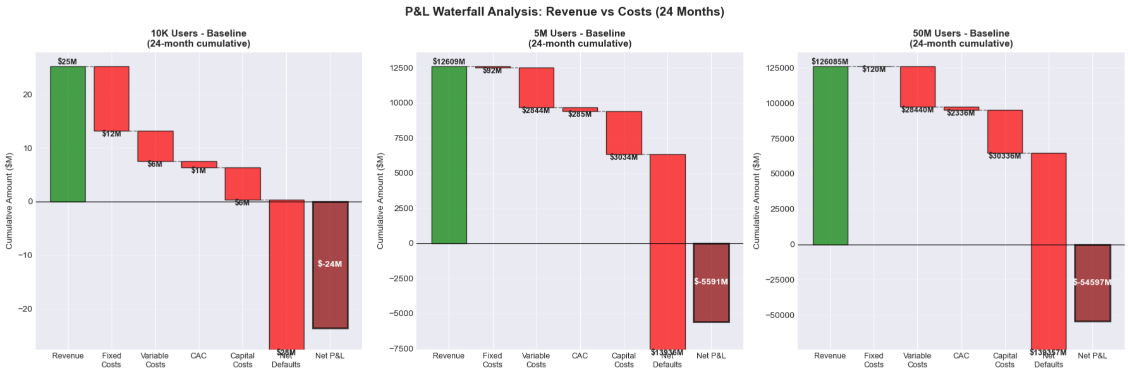

Figure 2: P&L Waterfall Analysis

To understand why scale fails to improve financial outcomes, it is helpful to examine the cumulative P&L results. The waterfall charts in Figure 2 illustrate revenue and cost accumulation over the 24-month simulation period for firms with 10000, 5 million and 50 million users, all under baseline economic conditions. Revenues grow substantially as scale increases, but losses expand even more quickly. At each scale level, the company finishes the period with deeply negative net income: roughly -$24 million at 10000 users, -$5.6 billion at 5 million users and –$54.6 billion at 50 million users.

The composition of these losses demonstrates why scale offers little relief. The dominant expense categories are credit defaults and the interest costs associated with financing the loan book. Both of these scale directly with GMV, and therefore grow proportionally as the user base expands. Unlike traditional lending operations, where underwriting improvements, diversification or customer maturation can reduce loss rates over time, BNPL loans remain short-duration, low-information and high-risk throughout the simulation. Collection recoveries are weak and capital costs remain substantial, meaning that each incremental unit of GMV introduces more cost than margin. Fixed operating costs, although material, quickly become insignificant relative to the scale of credit losses and interest expenses at higher user levels.

What the waterfall visualizations ultimately reveal is that the BNPL model does not generate improving unit economics at any scale. Larger firms simply incur proportionally larger losses. In this respect, the model resembles a high-velocity version of subprime unsecured lending, in which rapid turnover and thin margins leave no buffer against even moderate default rates.

§2.3. Insights

One of the most important contextual findings from this work is that predicting default in BNPL is not the fundamental challenge. Our credit models consistently achieve ROC–AUC scores above 0.90, indicating that lenders can distinguish high-risk and low-risk users with considerable accuracy. Given our ability to do so, it is a near guarantee that companies in the space are also able to, and likely to superior extent.

The financial results generated by our simulations mirror the real-world performance of BNPL leaders. Klarna, for example, reported only around $20 million of profit on roughly $2.8 billion of revenue after years of scale expansion, effectively a rounding-error margin for a platform processing tens of billions of dollars annually [oApp25]. Affirm's public filings tell a similar story: rapid GMV growth and strong user adoption, yet chronic difficulty converting that scale into sustainable profitability [Sal25]. These outcomes are consistent with what the model shows. BNPL operators can grow large, reach tens of millions of users and billions of dollars of GMV, and still fail to produce margins that resemble those of mature lenders.

The implication is clear: these companies are not operating in pursuit of lending profitability, but in spite of its absence. When we model the business from first principles, the BNPL unit economics are fundamentally incapable of producing durable profits under reasonable assumptions. Default losses, funding costs and variable expenses compound faster than revenues scale, even when user growth is explosive and credit-risk models perform well. The largest firms appear to have reached the same ceiling the simulation reveals: enormous top-line volume paired with negligible or negative profit, a boundary that cannot be crossed simply by adding more customers or more capital.

This raises a broader strategic question about why rational firms would continue to operate a model that performs so poorly as a standalone business. Part of the answer lies in recognizing that BNPL's value proposition is not limited to its direct financial returns. For many operators, BNPL functions less as a lending product and more as an acquisition engine, a data-collection mechanism or a merchant-enablement tool. It drives engagement, stimulates incremental consumer spending and strengthens strategic positioning within broader ecosystems. In this sense, the "losing" economics of BNPL may be tolerable, or even strategically attractive, when the product functions as an entry point to higher-margin activities such as advertising, data monetization, or cross-selling into more profitable financial products like credit cards. BNPL firms may also be operating according to a familiar private-equity and venture-capital logic: sustain losses long enough to build a user base, create strategic relevance, and develop infrastructure that a larger player will eventually acquire. This dynamic is already visible in the market, most notably in Block Inc.'s $29 billion acquisition of Afterpay in 2021 [Blo22].

§3. Conclusion

§3.1. Additional Considerations

The modeling presented in this report demonstrates that BNPL lending, when evaluated as a standalone credit product, does not achieve sustainable profitability under reasonable economic or behavioral assumptions. Although real-world operators may benefit from merchant subsidies, ecosystem synergies, or data-driven cross-selling advantages, these factors sit outside the unit economics of the loans themselves and do not materially alter the core result: BNPL firms consistently struggle to generate durable margins, and the largest players provide empirical confirmation of this pattern. An important point is that credit risk is not the source of failure. Our probability-of-default models, built on a broad set of behavioral and affordability features, achieve high predictive accuracy. The issue is that even accurately predicted defaults remain too large, and too expensive to finance, for revenues to offset. The financial unsustainability arises from structural characteristics of BNPL credit, not from limitations in underwriting.

§3.2. Closing Remarks

At the heart of this project is a machine learning pipeline designed to estimate consumer default risk using three model families: logistic regression, random forests, and histogram-based gradient boosting. These models draw on features reflecting loan terms, affordability signals, demographic factors, and repayment behavior. Their strong performance underscores that BNPL lenders' credit challenges are not due to inadequate prediction. Taken together, the findings show that BNPL can succeed as a strategic product, driving user acquisition, increasing merchant conversion, or supporting broader platform economics, but not as a profitable lending business in isolation. The model's conclusion is therefore straightforward: BNPL persists because of the value it creates around the loan, not within it.

Career Opportunities in Data Center Financing

Non-Technical Project

§4. Background

This report analyzes a specific emerging job opportunity within the Data Center sector, focusing on the strategic shift by Goldman Sachs to formalize an "AI Infrastructure" financing team . As the demand for artificial intelligence capabilities surges, the physical infrastructure required to support it, specifically data centers, power grids, and processors, requires massive capital investment [KKRnd]. Goldman Sachs has responded by creating a dedicated team within its Global Banking and Markets division to capture this financing market [Tip25]. This role represents a hybrid intersection of traditional project finance, structured credit, and specialized technical knowledge regarding AI hardware and energy assets.

§5. Market Context and Strategic Motivation

The primary driver for this new role is the explosive growth in financing demand for AI-related infrastructure. Deals in this sector are currently valued in the billions of dollars, creating a significant opportunity that investment banks are eager to share. The financing landscape covers three critical pillars:

1. Data Centers: the physical facilities housing the servers.
2. Power facilities: the energy grids and renewable sources required to run energy-intensive AI workloads.
3. Processors: the financing of the actual compute hardware (GPUs / TPUs).

Goldman Sachs has formed this dedicated team to achieve dual commercial objectives. the first is *direct revenue*, whereby they are increasing loan business revenue by originating debt for infrastructure projects. The second is *asset management synergy*, where they are creating tools that can be offered to asset management and wealth management clients, effectively packaging these infrastructure loans as investment products. This move builds upon an already successful foundation; Goldman Sachs' FICC (Fixed Income, Currency, and Commodities) financing business has reported quarterly revenues exceeding $1 billion, indicating a robust environment for expanding into this specific vertical.

§6. Business Scope and Coverage

While the catalyst for this team is the AI wave and the associated data center construction boom, the operational scope is broader, and they have a global reach mandate to cover both developed and emerging markets [Bai25]. Reports currently indicate that the team won't limit itself to AI assets, and will manage traditional infrastructure construction and upgrade projects, including essential civil infrastructure such as toll roads and airports. This creates a diversified portfolio where high-growth, high-risk AI assets are balanced against stable, traditional infrastructure projects.

§7. Job Role Analysis: Required Skill Set

The role requires a sophisticated blend of financial modeling, legal structuring, and deep technical understanding of the assets being financed.

§7.1. Core Financial Competencies

1. Project Finance Modeling: Candidates must possess the ability to build complex Discounted Cash Flow (DCF) models tailored for long-term cash flows. This is specifically applied to the unique revenue models of data centers and power facilities.

2. Debt Structuring: Expertise is required in designing the capital stack. This involves structuring loans into different tranches, such as senior debt (lower risk/return) versus mezzanine debt (higher risk/return), and understanding strategies for retaining assets on the balance sheet versus syndicating them.

3. Securitization: A key function of the role involves packaging debt obligations to sell to third-party investors, particularly insurance companies. This requires knowledge of how to structure these packages to meet the risk profiles of the broader securitization market.

§7.2. Sector-Specific Technical Knowledge

7.2.1. AI Asset Valuation

Candidates must understand the life-cycle of the hardware. Unlike a toll road, AI processors and GPU equipment depreciate rapidly. Understanding this depreciation curve is vital for accurate financing.

1. The Economic versus Accounting Lifecycle Paradox: Candidates must distinguish between the accounting useful life often standardized at 5 to 6 years by hyperscalers to smooth earnings and the economic useful life of cutting-edge AI chips often only 1.5 to 3 years due to rapid obsolescence.

2. Workload Specific Valuation: Valuation models must differentiate between Training hardware which faces the fastest obsolescence curve losing roughly 50 to 70 percent value within 18 months and Inference hardware which can remain economically viable for 3 to 4 years by running less demanding optimized models.

3. Residual Value Risk: Financing structures must account for the earnings cliff meaning the risk that the collateral GPUs becomes a liability rather than an asset if secondary market prices crash before the loan term matures, specifically monitoring performance per watt metrics that render older chips uneconomical to power.

7.2.2. Infrastructure insights

A deep understanding of the physical assets is mandatory. This includes data center operations, power grid dynamics, and specifically PPA Power Purchase Agreements and renewable energy financing, which are critical for sustainable data center operations [CFA25].

1. Grid Interconnection and Timing Mismatches: A critical financing risk is the timeline mismatch where AI data centers can be built in 18 to 36 months, but the necessary utility substations and high voltage transmission upgrades often take 3 to 5 years or more, creating stranded asset risks where completed facilities sit idle waiting for power.

2. PPA Structuring Physical versus Virtual: Expertise is required in structuring Power Purchase Agreements to match data center load profiles. This includes Physical PPAs involving direct delivery of electrons via the grid often requiring proximity versus Virtual PPAs involving financial swaps for renewable attributes. Candidates must navigate the tension between developers needing 15 to 20 year guaranteed revenue terms and tech tenants preferring shorter 5 to 10 year terms to maintain flexibility.

3. Energy Density and Cooling Economics: Financing models must account for the shift in CapEx density. AI facilities require significantly higher power density such as 40 to 100 kilowatts per rack versus traditional 5 to 10 kilowatts, necessitating expensive liquid cooling infrastructure that fundamentally changes the construction cost basis and operating expense models compared to legacy cloud data centers.

7.2.3. Market Analysis

The ability to analyze macro trends, specifically the surge in borrowing demand from both governments and corporations related to AI development and defense spending will be crucial.

1. Defense and Operational AI Spending: Analysis should focus on the shift from R&D to Operational AI in defense. With the DoD requesting over 13 billion dollars for AI and autonomy for fiscal year 2026, financing opportunities are moving toward contractors deploying mission critical autonomous systems rather than just experimental research, requiring distinct risk assessment frameworks for government backed receivables.

2. Corporate Shadow Debt and Leases: Candidates must identify borrowing demand disguised as financing leases. Major corporations are increasingly using capital leases to keep massive GPU expenditures off immediate cash flow statements, creating a hidden layer of leverage that requires sophisticated credit analysis to uncover.

3. The CapEx Gap: Borrowing demand is driven by the widening gap between operating cash flows and the massive capital expenditures required for AI projected to reach multi trillion dollar levels. This forces even cash rich tech giants to seek external debt and private credit partners to fund the infrastructure of modern society without diluting shareholder returns.

§8. Candidate Profile and Qualifications

§8.1. Educational Background

1. Academic Focus: The standard requirement is a strong background in Finance, Economics, or Quantitative fields.

2. Technical Edge: There is an increasing valuation of candidates with Engineering or Computer Science backgrounds. This technical literacy is crucial for understanding the underlying physical assets (Compute/AI) that are being financed.

§8.2. Professional Certifications

1. CFA & SIE: The Chartered Financial Analyst (CFA) designation and the Securities Industry Essentials (SIE) exam are cited as strong indicators of technical proficiency and are considered standard industry requirements for this level of role.

8.2.1. Experience

1. Capital Solutions: Experience in private credit, loan arrangement, or alternative capital markets is highly preferred.

2. Deal Execution: Candidates should have a proven track record in executing billion-dollar scale financing deals or complex infrastructure projects.

3. Cross-Functional Agility: The role sits within the global banking and markets division but requires collaboration across FICC and asset management teams.

§9. Career Trajectory and Future Opportunities

This role positions professionals at the nexus of finance and technology, opening three distinct high-value career paths:

1. Infrastructure Private Equity: A natural transition is to top-tier Private Equity firms such as Blackstone or KKR . These firms are aggressively acquiring data centers and digital infrastructure, making the skillset of valuing and financing these assets directly transferable [KKRnd].

2. Tech Strategy (Hyperscalers): Professionals may move to strategic finance roles at major technology firms (Hyperscalers) like Google, Microsoft, or Amazon. These roles involve managing the massive capital expenditure budgets required for AI infrastructure expansion.

3. Specialized Investment Funds: Opportunities exist in niche funds that focus specifically on the energy transition and digital assets, capitalizing on the intersection of power consumption and AI growth.

§10. Conclusion

The Goldman Sachs AI infrastructure team represents a strategic pivot to capture the financing demand created by the AI revolution. For job seekers, this presents a high-barrier, high-reward opportunity. It requires a candidate who is not only proficient in complex financial structuring and modeling but also possesses a non-technical but deep understanding of the hardware and energy systems that power the modern economy.

§11. Appendix

§11.1. Codebase

1. Risk Model: https://github.com/Campiomeee/BNPL-Project-Modeling/tree/main.

2. Simulation: https://github.com/ngnk/789_bnpl_simulation/tree/main

§11.2. Simluation

11.2.1. Economic Parameters

Economic conditions fundamentally shape BNPL performance through four interconnected parameters that simultaneously affect both revenues and costs. Our model incorporates six distinct economic scenarios grounded in historical precedent:

11.2.2. Default Rates

Default rates represent the monthly probability that a user fails to make a required payment. Our baseline 3.5% monthly default rate aligns with industry disclosures: Klarna reported loss rates of 3-4% in 2021-2022 investor materials, while Affirm's SEC filings showed charge-offs in the 4-6% range. These rates significantly exceed traditional credit card defaults (2-3% in stable conditions) due to BNPL's weaker underwriting standards

Table 2: Macroeconomic Scenarios and Assumptions

Scenario	Default Rate	GMV Growth	Recovery Rate	Capital Cost	Historical Period
Baseline	3.5%	+5%	40%	8%	2019 stable economy
Mild Recession	5.0%	0%	35%	10%	2001, 2020 COVID
Severe Recession	7.5%	-5%	30%	15%	2008–2009 GFC
Boom Times	2.5%	+15%	45%	6%	2021 stimulus era
Stagflation	6.0%	-2%	30%	18%	2022–2023 inflation
Rate Shock	4.0%	+3%	38%	20%	2023 rapid rate hikes

and the "phantom debt" problem-BNPL obligations frequently go unreported to credit bureaus, allowing users to accumulate unsustainable debt loads.[2]

During economic downturns, defaults increase substantially. FDIC data from the 2008-2009 recession shows credit card default rates reaching 6-10%, which we use as the basis for our severe recession scenario (7.5% monthly). The mild recession scenario (5.0%) reflects shorter, less severe downturns like the 2001 dot-com bust or the brief 2020 COVID shock before stimulus programs took effect.

11.2.3. GMV Growth Rates

Gross Merchandise Volume (GMV) growth affects top-line revenues but also scales capital requirements and default exposure proportionally. During the 2020-2021 boom period, Afterpay reported GMV growth exceeding 90% year-over-year, while Affirm grew 70%+. Our boom scenario's +15% monthly growth (+435% annualized) captures this extraordinary expansion.

Conversely, recessions compress discretionary spending. During 2008-2009, U.S. consumer discretionary spending contracted 3-8% annually. Our severe recession scenario models -5% monthly GMV growth, while baseline assumes sustainable 5% monthly growth typical of maturing fintech products.

11.2.4. Recovery Rates

Recovery rates quantify how much BNPL companies recoup from defaulted loans through collections. Traditional credit card recoveries range 20-40% of charged-off balances, which we use as our baseline (40%). BNPL's smaller average loan amounts ($150 in our model vs. $1000+ for credit cards) make collections less economically viable, potentially yielding lower recoveries.

Economic conditions significantly impact collections effectiveness. During recessions, borrowers have fewer resources and collection agencies face higher volumes, reducing recovery rates by 20-30%. Our severe recession scenario assumes 30% recovery, while boom conditions allow for 45% recovery as borrowers regain financial stability.

11.2.5. Capital Costs

BNPL companies fund their loan books primarily through warehouse credit lines, which carry costs tied to the Federal Reserve rate plus credit spreads. In our baseline scenario (8% annual), we assume a 3% Fed funds rate plus a 5% spread, typical for fintech lenders in stable conditions During crises, credit markets tighten dramatically. The 2008 recession saw spreads widen to 10-12% over base rates as lenders repriced risk. Our severe recession scenario (15% capital cost) reflects these conditions. The rate shock scenario (20%) models the 2023 situation: 5.5% Fed funds rate plus widened spreads for distressed fintech companies facing valuation pressure and increased regulatory scrutiny.

11.2.6. Business Model Parameters

Beyond macroeconomic conditions, BNPL companies face structural business parameters that shape their unit economics and determine whether they can achieve profitability at any scale.

11.2.7. Revenue Streams

BNPL companies generate revenue primarily through merchant discount rates (MDR)- the fee merchants pay for BNPL services. Our model uses 3.5% of GMV, reflecting disclosed rates from major players. Affirm's SEC filings show merchant fees ranging 2.5-3.5%, while Afterpay historically charged 3-6% depending on merchant size and negotiating power. We use the conservative mid-range estimate.

$$\text{Monthly Revenue from Merchant Fees} = \text{GMV} \times \text{Users} \times 0.035 \tag{11.0.1}$$

Secondary revenue comes from recovery on defaulted loans. When users default, companies recover a portion through collections, though collection costs consume approximately 25% of recovered amounts.[12] Our model therefore credits only 75% of recovered defaults as revenue:

$$\text{Recovery Revenue} = \text{Default Amount} \times \text{Recovery Rate} \times 0.75 \tag{11.0.2}$$

§11.3. User Behavior and GMV Generation

We model user behavior through two key metrics: average purchase size ($150) and transaction frequency (2.5 purchases per user per month). These parameters yield $375 in monthly GMV per user. Industry data supports these assumptions, Afterpay reports average purchase values around $150, while Affirm's average is approximately $200. Our frequency assumption is conservative, as not all registered users transact monthly; active user rates vary significantly by season and economic conditions.

$$\text{Monthly GMV per User} = 150 \times 2.5 \text{ dollars} = 375 \text{ dollars} \tag{11.0.3}$$

11.3.1. Fixed Opearting Costs

Fixed costs scale with organizational complexity rather than GMV volume. Our tiered cost structure reflects realistic staffing and infrastructure requirements, displayed below in Table 3.

Table 3: Operational Scale and Cost Structure

User Scale	Monthly Fixed Costs	Annual Costs	Organization Profile
$<$ 100K	\$500K	\$6M	$\sim$ 50 employees, basic infra
100K – 1M	\$1.5M	\$18M	$\sim$200 employees, compliance
1M – 10M	\$3M	\$36M	$\sim$500 employees, multiple markets
$>$ 10M	\$5M	\$60M	1000+ employees, full operations

These estimates align with actual BNPL company disclosures. Affirm employed approximately 2,000 people with estimated annual operating expenses exceeding \$300M in 2023, while Klarna had roughly 7,000 employees (before 2022 layoffs) with annual operating costs exceeding \$1B. Fixed costs include engineering, compliance, risk management, general & administrative expenses, legal, and office infrastructure.

11.3.2. Variable Costs

Variable costs scale directly with transaction volume at 1.5% of GMV, comprising:

1. Payment processing (0.5-0.8%): Fees paid to payment processors like Stripe or Adyen.
2. Customer service (0.3-0.5%): Support tickets and dispute resolution.
3. Fraud prevention (0.2-0.3%): Third-party fraud detection services.
4. Credit bureau pulls (0.1-0.2%): Underwriting data costs.

11.3.3. Customer Acquisition Cost (CAC)

User acquisition costs decline with scale as brand awareness reduces marginal marketing spend:

1. Early stage (<500K users): \$50 per new user-reflecting high digital acquisition costs for unknown brands
2. Growth stage (500K - 10M users): \$30 per new user-economies of scale in marketing
3. At scale (>10M users): \$20 per new user-strong organic growth and brand recognition

11.3.4. Capital Structure and Interest Expense

BNPL companies finance their loan books primarily through debt. We model an 80% debt-to-GMV ratio, meaning 80% of outstanding loans are funded by warehouse credit lines rather than equity capital. This creates significant interest expense that scales with loan volume:

$$\text{Monthly Interest Expense} = (\text{Total Loan Book} \times 0.80) \times (\text{Annual Capital Cost}/12) \tag{11.0.4}$$

The loan book size is determined by the average loan duration (3 months in our model) multiplied by monthly GMV. This capital structure amplifies the impact of rising interest rates, a critical vulnerability demonstrated during the 2022-2023 rate shock period.

11.3.5. Scale Scenarios and Starting Capital

To address the fundamental question of whether scale solves BNPL's unit economics, we test six user scale scenarios ranging across four orders of magnitude. Each scale level receives starting capital calibrated to real-world funding patterns:

Table 4: Capital Requirements by User Scale

Users	**Capital**	**$ / User**	**Funding Stage**	**Comparable**
10,000	$5M	$500	Seed	Early startup
50,000	$15M	$300	Series A/B	Regional player
100,000	$30M	$300	Series B	National expansion
500,000	$100M	$200	Series C	Established brand
5,000,000	$500M	$100	Series D+	Pre-IPO (Afterpay)
50,000,000	$2B	$40	Late stage	Affirm/Klarna scale

These capital levels reflect actual BNPL fundraising history. Affirm raised $1.3B+ through its 2021 IPO at a $24B valuation, while Klarna's 2021 Series H round raised $800M at a $45.6B peak valuation. The declining cost-per-user reflects both economies of scale and reduced perceived risk as companies mature. Notably, even our largest scenario ($2B capital for 50M users) falls short of the cash these companies actually raised, making our analysis conservative regarding starting liquidity.

§References

[Bai25] Bain & Company, *Next phase of data center growth to be more disciplined, but risks of power constraints and construction delays remain*, Press Release, Accessed December 9, 2025, Oct. 2025. [Online]. Available: https://www.bain.com/about/media-center/press-releases/20252/next-phase-of-data-center-growth-to-be-more-disciplined-but-risks-of-power-constraints-and-construction-delays-remain-bain--co-research/.

[Blo22] Inc. Block, *Completes acquisition of afterpay*, Block Website — Press Release, Accessed December 9, 2025, Jan. 2022. [Online]. Available: https://block.xyz/inside/press-release-afterpay-acquisition.

[CFA25] CFA Institute, *Seven key considerations for data center investors*, CFA Institute Research, 2025.

[Con23] Consumer Financial Protection Bureau, "Buy now, pay later: Market trends and consumer impacts," CFPB, 2023. [Online]. Available: https://www.consumerfinance.gov/data-research/research-reports/buy-now-pay-later-market-trends-and-consumer-impacts/.

[Fit22] Fitch Ratings, "Buy now, pay later: Business model and credit risks," Fitch Ratings, 2022. [Online]. Available: https://www.fitchratings.com/research/corporate-finance/buy-now-pay-later-business-model-credit-risks-10-02-2022.

[Ibi25] G. Ibikunle, *Paylater loan data*, Kaggle Dataset, Accessed December 2025, 2025.

[KKRnd] KKR Global Institute, *Data centers: The hubs of digital infrastructure*, KKR Insights, n.d.

[oApp25] Business of Apps, *Klarna statistics*, Business of Apps Website, Accessed December 9, 2025, 2025. [Online]. Available: https://www.businessofapps.com/data/klarna-statistics/.

[Res24] Juniper Research, *Buy now, pay later users to exceed 670 million globally by 2028*, Juniper Research Press Release, Accessed December 9, 2025, Apr. 2024. [Online]. Available: https://www.juniperresearch.com/press/pressreleasesbuy-now-pay-later-users-to-exceed-670-million-globally/.

[Ris25] Inc. RiskSeal, *Digital signals for bnpl risk assessment*, RiskSeal Blog, April 15, 2025, Apr. 2025. [Online]. Available: https://riskseal.io/blog/digital-signals-for-bnpl-risk-assessment.

[Sal25] Beck Salgado, *Affirm reports 34% revenue spike in latest earnings*, Revenue Brew, Accessed December 9, 2025, Nov. 2025. [Online]. Available: https://www.revenuebrew.com/stories/2025/11/12/affirm-reports-34-revenue-spike-in-latest-earnings.

[She21] Ron Shevlin, *Buy now, pay later: The "new" payments trend generating $100 billion in sales*, Forbes, Accessed December 9, 2025, Sep. 2021. [Online]. Available: https://www.forbes.com/sites/ronshevlin/2021/09/07/buy-now-pay-later-the-new-payments-trend-generating-100-billion-in-sales/.

[Tip25] TipRanks, *Goldman sachs banks on AI boom with new financing team*, TipRanks News, 2025. [Online]. Available: https://www.tipranks.com/news/goldman-sachs-banks-on-ai-boom-with-new-financing-team.