Dynamic decision-making under customer abandonment risk

Motivation. Service firms are often faced with decisions where the options available to the firm differ in their risk-reward profiles. That is, when delivering service to customers, they often must choose between safer and riskier courses of action, with different expected rewards from the different choices. For example, a financial advisor might suggest a risky strategy to a given customer (a portfolio consisting mostly of stocks) or a safer one (a portfolio with more cash and bonds). Online platforms often face similar tradeoffs. Consider the case of a meal delivery service such as Blue Apron. Every week, they must decide whether to send a particular customer safe dishes or riskier fare.

In this paper, we consider how a service firm should make such choices when faced with the risk of customer abandonment. Customers are expensive to acquire and often leave quickly if disappointed with the service they receive, especially online. We consider how firms should dynamically choose between safer and riskier actions when faced with customers that are more likely to quit if they are unhappy with their recent few experiences. We consider both cases where the firm may or may not mix between the two actions.

Our Model. We study a continuous-time model with one firm repeatedly interacting with one customer. At any given time, the firm has to choose between two modes of service to the customer: a safe mode that generates rewards at a constant rate, and a risky mode that accumulates rewards according to a positive-drift Brownian Motion, where the drift may be larger or smaller than the reward rate in safe mode. We also allow for mixing between the two strategies in order to account for a setting like a financial advisor’s problem. We assume the firm earns rewards that are proportional to that of the customer.

We model customer abandonment using a hazard rate function of a one-dimensional state of the customer. We label the state happiness. We model customer happiness via a behavioral model where at each point in time her happiness equals an exponentially weighted
moving average of her recent rewards. Happiness captures a customer’s “perception” of the reward she can get from interacting with the firm under recency bias. The hazard rate of customer abandonment is, for simplicity, a step function that drops to zero at a happiness value we call the happiness threshold. That is, customers do not quit while their happiness is above the threshold. The happiness threshold captures a customer’s minimum “tolerance” of the reward she can get from the firm. We assume her happiness threshold is higher than the reward rate generated by the safe action so the customer will eventually quit.

We study how the firm can maximize the sum of rewards over the finite lifetime of the customer, by optimally trading off myopic rewards with the risk going forward of the customer leaving the system.

**Results.** Succinctly, the firm should use the mode with higher average reward rate, _except:_

- If the customer is currently not a flight risk but may become one soon if not served well, this makes the safe mode relatively more attractive.
- If the customer is currently a flight risk, but may no longer be one if served well, this makes the risky mode relatively more attractive.

The strength of these two effects is increasing in the volatility of the risky mode and decreasing in the difference between the (average) reward rates of the service modes.

We now describe our results in more detail. In the optimal policy, the service mode employed by the platform is a function of the current happiness of the customer. We find that the optimal policy always follows a “sandwich-type” structure, using the high-average-reward service mode when happiness is high or low, but sometimes use the low-average-reward service mode when happiness takes an intermediate value. In the case of a financial advisor, the optimal strategy mixes between the two modes when happiness is in this intermediate range.

This characterization holds regardless of whether the safe mode yields higher or lower
expected return than the risky mode; and for every possible value of the happiness threshold. The intuition at high and low happiness values is fairly straightforward: when far away from the happiness threshold, the firm should act myopically and choose the service mode that generates higher immediate rewards.

What happens near the happiness threshold is far less obvious. To understand this question, we need to break our analysis into two cases that depend on the problem parameters:

1. When the risky mode has higher expected rewards than the safe mode: the optimal policy uses the safe mode (or mix the safe mode with the risky mode if allowed) in an interval (possibly empty) just above the happiness threshold because it increases the time before the customer incurs a positive hazard rate of quitting, along with the risky mode elsewhere.

2. When the safe mode has higher expected rewards than the risky mode: the optimal policy uses the risky mode for happiness in an interval just below the happiness threshold to decrease the time the customer spends with a positive hazard rate, and the safe mode elsewhere.

One notable feature of these results is asymmetry of where the center of the “sandwich” policy lies. **Risk aversion just above the threshold:** As an intermediate strategy, the safe mode is only used above the happiness threshold to slow down the customer from entering the positive hazard rate zone. **Risk-seeking just below the threshold:** In contrast, the risky mode is only used for intermediate happiness below the happiness threshold, in an attempt to shorten the time spent by the customer in the positive hazard rate zone.