



Munich Personal RePEc Archive

Assessing the Impact of Chokepoints in a Customer Onboarding Process

Rodriguez, A.E. and Rosen, John

University of New Haven, University of New Haven

21 June 2023

Online at <https://mpra.ub.uni-muenchen.de/117997/>
MPRA Paper No. 117997, posted 19 Jul 2023 07:11 UTC

Assessing the Impact of Chokepoints in a Customer Onboarding Process

A.E. Rodriguez (Corresponding author)

Department of Economics & Business Analytics

Pompea College of Business

University of New Haven

West Haven, CT 06514, United States

Tel: 203-932-7372 Email: arodriguez@newhaven.edu

John D. Rosen

Department of Economics & Business Analytics

Pompea College of Business

University of New Haven

West Haven, CT 06514, United States

Tel: 203-434-4964 Email: jrosen@newhaven.edu

Abstract

Customer onboarding processes have become dysfunctional, especially with regards to the increasing number, complexity, and often, competing demands, of regulatory and law enforcement bodies with oversight over a firm's practices. Prospective customers are screened across any number of considerations ranging from conventional ones such as financial considerations (i. e., "Does this customer have an acceptable balance sheet?") to the more recent socio-cultural ones (i. e. "does this customer have an effective diversity program?" "Has this customer expressed a commitment to environmentally sustainable business practices?"). An impaired sales pipeline resulting from an impaired customer vetting process may diminish economic returns, reduce profitability, and erode market share. A firm intent on repairing their customer intake processes could examine whether rescinding or reducing extant customer acceptance thresholds will enhance their performance. However, many firms are beset by a peculiar outcome that complicates auditing the onboarding process. Customer portfolios are routinely culled of non-performing customers or costly-to-serve customers, leaving a selection of seemingly successful customers – a data artifact known as a one-class problem. We simulate the onboarding process to isolate the effect of changes in established acceptance thresholds on customer's likelihood of success. When only One Class ("Successful" or "Performing") customers are available, One-Class algorithms can be used for resolving the matter. This study illustrates the use of Support Vector Machines and Isolated Random Forests in reconstructing a representative sample of the customer pool. We find there exists a tradeoff between reductions in customer thresholds and the firm's commitment to ensuring customer success.

Keywords: Customer onboarding, Model-shift, Data-shift, One-Class, Isolation Forests, Support Vector Machines.

*“The ability to reflect on one’s past actions and envision
alternative scenarios is the basis of free will and social responsibility.”*

Judea Pearl

1. Introduction

Customer onboarding may have once been a simple, one-stop process.¹ But for many modern manufacturers, distributors, and other upstream suppliers, customer onboarding has increasingly become a complex, elaborate, multi-step ordeal (Chen & Lee, 2022) (Lee & Tang, 2017). Onboarding increases in complexity with increases in number, stringency, and changes in regulatory, legal, and financial oversight by federal, state, and international regulatory agencies.

Modern day scrutiny of prospective customers, once limited to financial considerations, has now ballooned to encompass a wide array of societal and political concerns spanning multiple continents and political jurisdictions. Increasingly wide-ranging “Know Your Customer” requirements compel firms to perform due diligence and examine relevant information of the prospective customer prior to onboarding. Today, onboarding due-diligence may address climate change sensibilities, green practices, labor concerns, disposition towards ethnic and racial minorities, genetically modified organisms, board composition, drug-trafficking, money laundering, terrorism counter-measures, Title VII considerations, *inter alia*. For today’s corporation, the process is extensive, complex, time consuming, challenging, and expensive—even when assisted by automated workflow processes.

Given the perceived potential for increased operational cost, risk, and sizable financial losses from accepting improperly vetted customers, multiple departments in modern corporations are effectively granted *de facto* veto power to summarily reject a customer candidate under scrutiny. The authority to terminate, or condition, a client or customer, proceeds from the firm’s recognition of and deference to, their specialized department’s individual area of expertise. For example, the Finance department can object to a candidate that fails to meet certain debt ratios, operational margins, or sales levels; the Compliance department may express concern over particular regulation-compelled security or safety protocols; and Legal may object to the political or reputational quality of the customer or, lately, perhaps an inadequate minority composition of the customers workforce.

Over time, the vetting left unchecked, at least two problems are likely to unfold. First, increased internal scrutiny results in the inevitable lengthening of the approval process thus delaying acceptance. This to the consternation of both the sales department and the customer. The lags may balloon to become of sufficient length so as to discourage aspiring customers. At that point, frustrated prospective customers may choose to withdraw from the vetting process and decide to do business with competing suppliers. Second, the lengthening of the approval process is likely to often devolve into increased internecine rivalry and stakeholder finger-pointing, blame-shifting, general ill-will and an associated decline in employee morale, (Barki & Hartwick, 1994).

¹ Onboarding is a managerial term originating in the human resources (“HR”) world. In HR it refers to the process of introducing new employees into an organization culture. Applied to a sales setting, it entails all the activities required of a company intent on supplying a new firm.

Unattended, a troubled sales pipeline may lead to lowered economic returns, reduced profitability (Halvorsrud, Kvale, & Folstad, 2016).² In fact, the primary objective of our work is to examine this assertion: specifically, how does relaxing administrative vetting thresholds impact the firm? In a glimpse of results explained in more detail later, we note here that relaxing customer acceptance thresholds may increase the firm’s Success Ratio – defined as the proportion of successful firms out of those accepted, those that passed the selection gauntlet.

Returning to our discussion, why would senior executives remain impassive in the presence of possible dysfunction – lengthened decision processes, frustrated or lost customers, etc. – in their onboarding operations? We believe this constitutes examples of *bounded awareness* whereby the required specialization and necessary expertise underscoring departmental vetoes is unmatched by senior management and thereby result in an instance of cognitive blinders that preclude any sound decision-making (Bazerman & Chugh, 2006). Bounded or limited awareness can happen when decision makers don’t gather relevant data, consider critical facts, or understand the relevance of the information they have – or, as is the case here, the relevance of the information they don’t have (Tett, 2015). We represent the firm’s dysfunction derived from the observed principal-agent problem as an implicit correlation between the key features of the company; increasing correlation entails more effective collaboration among decision-makers.

Decision-makers need to understand the nature of the prospective tradeoffs involved; tradeoffs that would reap considerable information essential for informed decision-making. What would happen to performance if they relaxed some of the thresholds underscoring one or several of the various onboarding vetting processes? An appropriate response to this question – barring an actual experiment – could be obtained via simulation of the onboarding process. Specifically, the likely scenarios that would represent the requisite situation can be modeled.

Modeling counterfactuals is always a challenge. It requires an appraisal of the population distribution of prospective customers, and it requires knowledge of the financial outcomes of those prospective customers selected and those not selected. Under erstwhile selection protocols, this would not have been problematic from a modeling perspective. But it is accessible via propensity modeling or matching (Leite, 2017), correcting for misclassification, or any of a variety of semi-supervised or one-class machine learning methods (Khan & Madden, 2014) (Zhou, 2018).

An onboarding process, especially in a regulated environment, presents peculiar modeling challenges. First, the vetting process actually selects fiscally and otherwise sound customers, customers that met all onboarding hurdles. Over time, boosted by collaboration and synergies among the contractual partners, most chosen customers perform well. This represents the “self-fulfilling process” dynamic.

Once the synthetic customer performance data are generated we can alter firm parameters to understand the distribution of customer outcomes. Relying on unsupervised machine learning algorithms, we reconstruct sets of customers and determine the Success Ratio and the Likelihood of Success and closely appraise its sensitivity to a putative Selection threshold. We show also the effect on outcomes of improvements in the firm’s decision-making.

² Commentators have closely examined the responses that firms make to regulatory and policy changes and the impact on operational performance (Hiatt, Grandy, & Lee, 2015) (Kozhikode, 2016).

2. A Principal-Agent Perspective of Onboarding Governance

Agency theory postulates that functional and administrative unit managers in corporations are hired by upper management to operate the business units. Put differently, they are hired because the firm wants to harness their specialized operational know-how (Eisenhardt, 1989) (Blair, 1995). Analytically, agency theory serves to examine instances where the principal and agent are at odds with respect to their desired outcomes. Agents and principals have asymmetric interests, and the agents may have a propensity to emphasize their self-interest.

Any ensuing conflict and attendant agency loss may be a result of varying degrees of cost-benefit burdens. Moral hazard, for instance, arises where an agent has no incentive to embrace firm-wide objectives because they bear only a portion of the potential costs of decision-making. Hold-up costs may proliferate when a parties' bargaining power in a relationship increases due to obligations undertaken by one party but not the other. And adverse selection considerations derive from transactions where one party has relevant information that the other one does not have. As a result, agents, department heads, may place their own interests before the principal's, senior management, when making decisions.

In the case of onboarding processes, the task of senior management is to align corporate-wide interests in efficient operations with the incentives of department heads who may pursue self-interested decisions characterized by a predilection for risk-averseness or career prospects. Difficulties in moderating inherent interests in self-preservation arise as due to the increased specialized information proffered by departments to underscore their actions. Simply put, upper-level management cannot discern whether the caution advocated, and inherent in agent's decision-making lags is true (Brudney, 1985).

A more immediate solution is to internalize the free-rider problem that besets the fragmented decision-making causing excessive decision-making lags. This can be accomplished by consolidating the onboarding approval process onto one decision-maker. This alternative makes sense if the firm can embrace the cost-benefits of a trade-off in enhanced or improved decision-making agility with possible dissipations of specialized know-how.

A more difficult problem sets forth the broader question as a logical predicate: how do we determine whether a company's onboarding process is working? A proper assessment requires knowledge beset by data limitations. That is to say, it is impossible to fully understand the quality of the onboarding process unless we know the fortunes of rejected potential customers; barring that we are unable to fully appraise our decision-making process.

3. Replicating the Onboarding Process

It is conceivable that a prospective customer's compliance variables and compliance performance be reduced to a compliance profile. Similarly, a prospective customer firm's reputation, goodwill, and other intangibles can be reduced to a reputational quality index. Assuming right skewed distributions to describe both prospective customer variables reflects the typical reality in competitive markets. The correlation between the variables is set arbitrarily to represent a complementarity between a firm's reputational score and the quality of its compliance performance. The correlation among features proxies for the cooperation among administrative personal handling customer vetting. Put differently, a higher correlation describes less "bite" from principal-agent distortions and vice-versa.

Our interest is in understanding both the individual influence of each particular attribute and the tradeoffs among these attributes. But more specifically, we are interested in the impact on the firm of a change in its reputational quality threshold for varying levels of feature covariance and varying levels of resources devoted to the success of the newly-formed relationship.

The examination of the impact of altering approval thresholds occurs over two periods (“0” and “1”). In both periods the firm examines prospective customers and onboards those meeting a threshold condition. Whether a prospective firm is accepted as a customer is uncertain; this decision is modeled via a stochastic binary process. Once accepted, whether a firm succeeds is also uncertain and similarly determined stochastically.

Onboarding propensity and the associated Propensity to Succeed are determined by both variables in the first period and second period. Presumably, Success is masked in the 2nd period and is therefore the variable that needs to be identified via the One-Class algorithms.

3.1 Model Structure

Let binary Q represent a prospective customer’s true status for a feature of interest: performance. There exists a relationship between Q and customer-level information, X . X may constitute financial variables, legal considerations, managerial quality, or any other feature of interest.

Let S indicate whether a particular customer is selected by the firm where the probability of a customer being selected depends on a firm’s reputational quality X_1 , along with other firm-level covariates: X_2 . The mapping showing this structure can be seen in Figure 1.

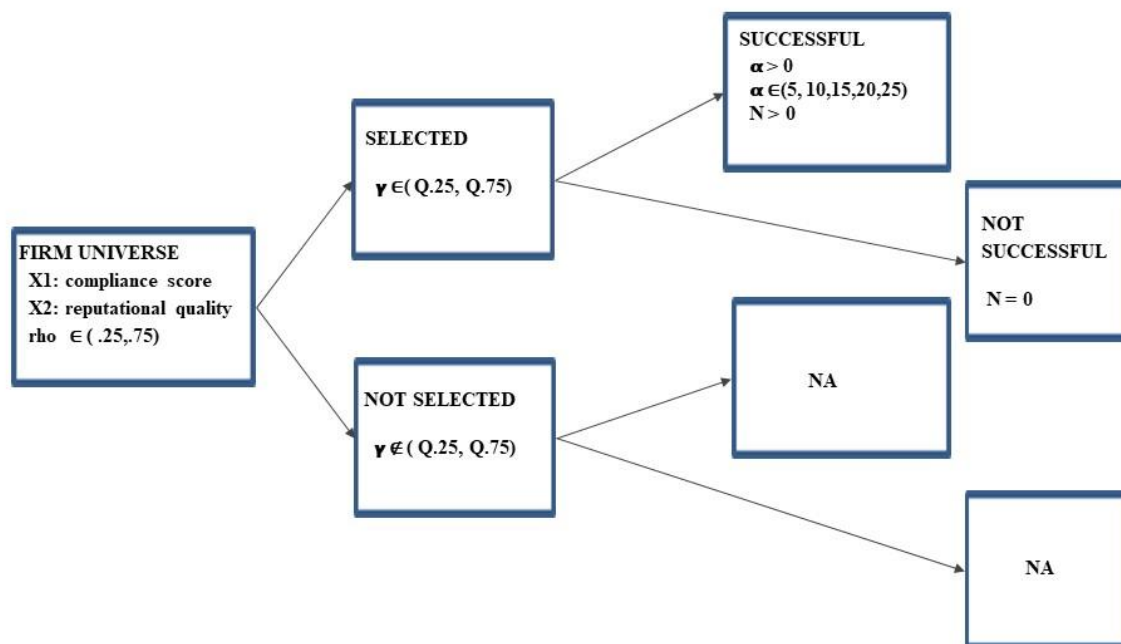


Figure 1

3.2 Simulation

We simulated data for a situation in which there were two baseline covariates: X_1 , X_2 , representing a compliance score and a reputational quality metric, for each firm. These covariates were drawn from a multivariate non-normal distributions random number generator with the following mean and covariance structure:

$$(X_1, X_2) \sim \varphi \begin{bmatrix} 1 & rho \\ rho & 1 \end{bmatrix}$$

Where rho proxies for administrative synergies and varies from between 0.25 and 0.75. Reputational Quality is rescaled to an inverse scale from 1 to 4, where 4 is ascribes to the highest quality candidates. Similarly, the Compliance Score variable is rescaled to range from 1 to 100 whereby a fuller compliance profile is positively increasing.

Initial selection occurs at a threshold set at the 3rd quartile of the Reputational Quality variable subsequently reduced to the 1st quartile – representing an ameliorating of the onboarding acceptance threshold. Similarly, the initial covariance structure is set at 0.25 – representing the situation of a firm exhibiting a dysfunctional customer onboarding process. We explore the impact of variations of the covariance structures up to 0.75.

The likelihood of Success is established by a probabilistic binary process.

$$Success = \text{expit}(a_0 + \alpha + a_1 * X_1 + a_2 * X_2)$$

Being accepted incorporates the customer into a firm's managerial ecosystem. Those customers accepted continue on to establish relationships, normative arrangements, and understandings with their supplier firm who now has a keen interest in the commercial success of their newly-minted partner (Hilton, Hajihashemi, Henderson, & Palmatier, 2020). The enhanced performance conveyed by the firm's success ecosystem is represented by the parameter alpha. The parameter alpha enhances managerial quality and thus augments the likelihood of the customer's success. We examine variation of α across the following vector:

$$\alpha = c(5, 6, 8, 10)$$

In each of the scenarios, we generated 10000 datasets.

4. Repairing Dataset Distortions

In an initially perplexing feature – we found that all the firm's customers were "Successful." There were no "Fails;" only One Class of customers. Close scrutiny reveals that there are two reasons for this - for the most part. First, a firm's "Success Ecosystem" is one of the intangibles at the root of many a managerial success. Both parties to the budding relationship want to make sure that the partnership works; companies invest considerable funds in building and maintaining relationships aimed at ensuring the longevity and soundness of their business relationships. This process allows for the quick identification of those companies who are "not going to make it."

Second, ongoing customer relationships are periodically scrutinized through a cost-benefit lens. Those customers who came out on the wrong side were politely "let go." The reason is simple: the cost of servicing the individual account is not surpassed by the gains or even prospective gains.

To sort out the impact of these two managerial tweaks we would have to reconstruct the current-customers portfolio to reflect, at the very least, "Performing" from "Under-Performing" firms. Only then would we be able to determine the impact of any changes in onboarding.

It is possible to identify those customer firms likely to succeed using grouping algorithms known in unsupervised learning as One-Class algorithms (Khan & Madden, 2004). In the terminology of machine learning, statistical techniques for classification are referred to as "supervised" and "unsupervised" learning methodologies. Supervised learning can occur when the data supplies both dependent (target group or class membership identification) and independent (predictor) variable observations. Unsupervised learning classification algorithms, on the other hand, derive classification results using "unlabeled data" where no known and verified dependent variable is available.

The use and capabilities of two such unsupervised learning algorithms is demonstrated: Support Vector Machines and Isolation Forests. One-class algorithms step in when one has no information on the classification label: put differently, whether the customer firm is likely to Succeed.

4.1 One Class Support Vector Machines

When data are all grouped into one class as is the case here it is necessary to find natural clustering of the data to group or classify into – at least – two classes. Support vector machine is an unsupervised learning algorithm derived from statistical learning theory.

A SVM One Class classifier learns the boundary between groups by maximizing the margin, or distance, between class members. SVM sets forth all available data as members of its first group, C_1 and the origin as the sole member of the second group, C_2 . The hyperparameter, ν , constitutes a penalty applied to the trade-off between groups one and two. The chosen kernel is key – and they include linear (inner-product), polynomial and sigmoid. Kernels help to determine the shape of the vector, plane or hyperplane and decision boundary between the groups. In this instance the fitting process is simple and relies on a line to separate our two Classes. The basic linear kernel function is selected:

$$K(x_i, x_j) = x_i * x_j + c$$

where x and y are input vectors and "*" represents the dot product and c is a constant.

The prediction utilizes the two known covariates. The classification results are compared to the previous generated Success classification scores obtained from the simulated propensity to Succeed; Success was generated via a randomized Bernoulli process.

The results return an Accuracy score of 75 percent a relatively acceptable performance in identifying the customers likely to be successful.

4.2 Isolation Forests

Isolation Forest is an ensemble learning method applicable to One Class problems. It explicitly isolates outliers rather than learn a model for normal instances.

A normal sample is hard to isolate from other samples. An outlier is more easily detected from other samples. Isolation Forests is composed of a fixed number of isolation trees each one built on a random selection of samples from the training set, the forest (Molnar, 2022). From this subset of samples, an isolation tree is constructed by a random recursive partitioning, until all the samples are isolated or until a stop criterion is reached.

The partitioning into classes is realized by the random selection of an attribute and the random choice of a pivot value in the range of the selected attribute. For an isolation tree, the sample scores are computed as the distance between the leaf node containing the sample and the root node of the tree.

The algorithm uses the number of tree splits to identify minority classes in an imbalanced data set. Intuitively, outliers take fewer splits because the density around the outliers is low. And again, caret's confusion matrix function is used to appraise the accuracy of the isolation forest algorithm.

The results return a classification accuracy of 91 percent indicating a reasonable alternative to classify customers or assign them to a constructed classification scheme.

5. Results

Figure 2 shows the impact on the firm's Success Ratio of reducing the acceptance threshold. The results are intuitive – the lower the criteria for accepting a new customer the lower the number of firms Succeeding. What is instructive it to see that closer attention to eliminating the principal-agent distortions improves this relationship considerably.

Average Success Ratio v. Acceptance Threshold

for High & Low Managerial Intensity Levels

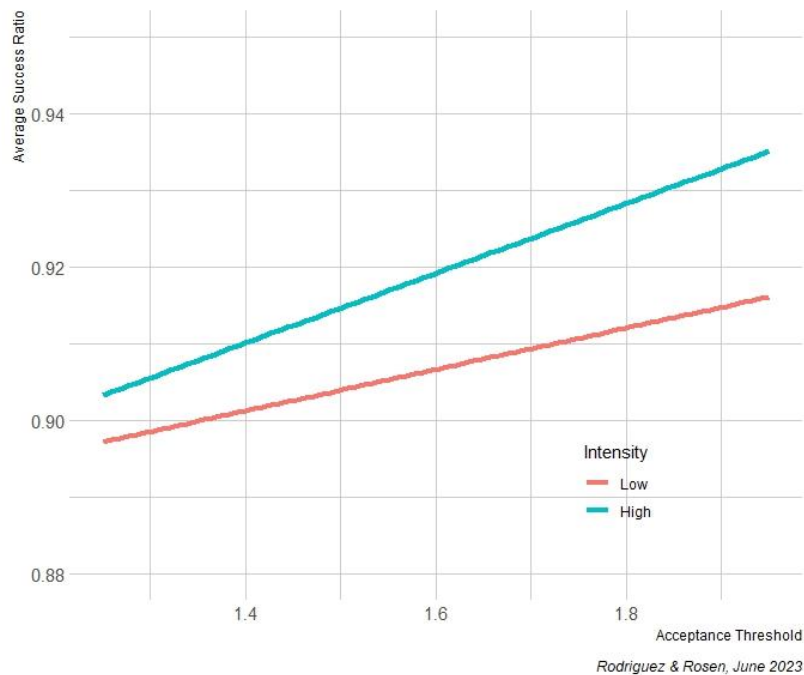


Figure 2

In fact, strategically, the firm can increase its acceptance pool and maintain its Success Ratio if it reduces the number of approval veto points.

Aside from enhancing managerial discipline, the firm has another option. It can devote resources to enhancing the Success EcoSystem whose aim is to enhance customer's performance by enhancing customer relationships. Figure 3 provides estimates of the relationship between the Average Probability of a firm's Success and increasing the level of resources devoted to the Success EcoSystem. Note also that the riskiness associated with the likelihood of Success diminishes as resources are improved.

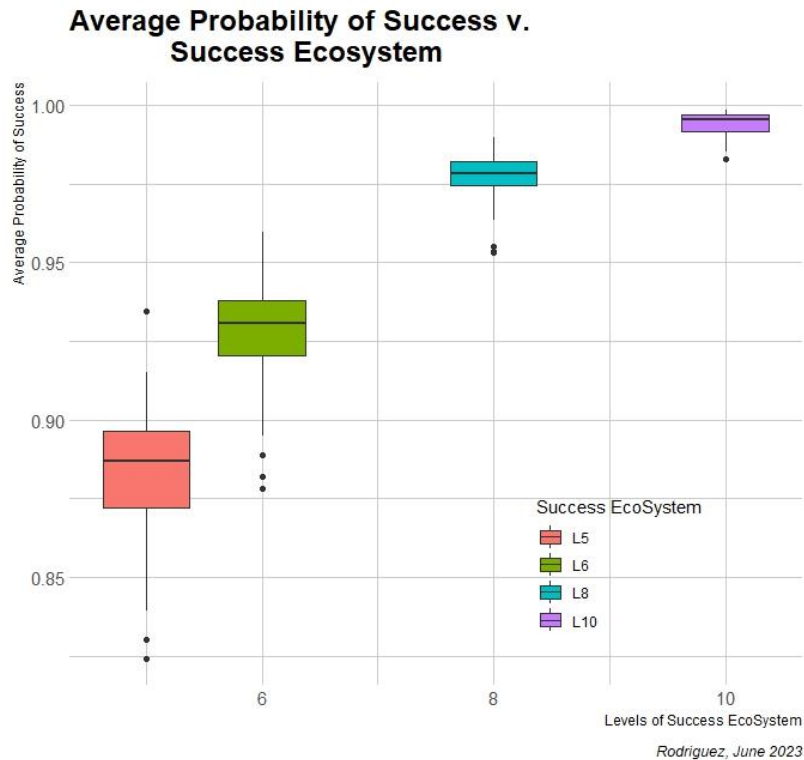


Figure 3

7. Discussion

Subjecting a firm’s onboarding process to a plausible understanding of its counterfactual outcomes can convey considerable benefits to a firm. Scrutinized decision-making lags may be better understood and thereby shortened. Operational cost and risk may be reduced; improved transparency of end-to-end process may be improved; share of customer spending may be increased; customer satisfaction and retention may be enhanced; consistency and conformity across business lines may be insured; and customer referrals may be increased.

First, we identify the existence of principal-agent based distortions of varying severity. Careful attention to this artifact can ameliorate most of any negative impact associated with the reducing of the firm’s customer acceptance threshold. Second, we recognize the value of the natural efforts of business partners to jointly strive for the success of a relationship. This built-in Success EcoSystem can also serve to mitigate the firms losses when opting to enhance their customer intake by resetting their acceptance thresholds.

7.1 Limitations

Simulations can be pliable, dependent on key assumptions, specified variables, chosen parameters, and other, similar idiosyncratic elements. We cannot purport to establish generalities aside from broad interpretations. Nonetheless, the analysis, the models, the simulations, create a rich platform for discussions of not only the possible but also countless scenarios that would otherwise be obscured.

8. Concluding Comments

The results presented here highlight the interconnectedness between established business acceptance thresholds and business practices that can be articulated to minimize any concerns or impact on performance measures. We show here that the impact on two performance metrics: the Success Ratio of the firm and the Average Probability of Success of any individual customers can be managed by careful fine-tuning of two tools at the disposal of the firm: managerial discipline and the Success Ecosystem.

References

- Barki, H., & Hartwick, J. (1994, March). Measuring User Participation, User Involvement, and User Attitude. *MIS Quarterly*, 18(1), 59-82. doi:<https://doi.org/10.2307/249610>
- Bazerman, M. H., & Chugh, D. (2006, January). Decisions Without Blinders. *Harvard Business Review*.
- Blair, M. M. (1995). *Ownership and Control: Rethinking Corporate Governance for the Twenty-First Century*. Washington, D.C.: Brookings Institution.
- Brudney, V. (1985). Corporate Governance, Agency Costs, and the Rhetoric of Contract. *Columbia Law Review*, 1403-1444. doi:<https://doi.org/10.2307/1122518>
- Chen, L., & Lee, H. L. (2022). Supply Chain Fairness. *Production and Operations Management*, 31(12), 4304-4318. doi:<https://doi.org/10.1111/poms.13849>
- Eisenhardt, K. M. (1989, Jan). Agency Theory: An Assessment and Review. *The Academy of Management Review*, 14(1), 57-74. doi:<https://doi.org/10.5465/amr.1989.4279003>
- Halvorsrud, R., Kvale, K., & Folstad, A. (2016). Improving Service Quality Through Customer Journey Analysis. *Journal of Service Theory and Practice*, 26(6), 840-867. doi:<http://dx.doi.org/10.1108/JSTP-05-2015-0111>
- Hiatt, S. R., Grandy, J. B., & Lee, B. H. (2015). Organizational Responses to Public and Private Politics: An Analysis of Climate Change Activists and US Oil and Gas Firms. *Organization Science*, 26(6), 1769-1786. doi:<https://doi.org/10.1287/orsc.2015.1008>
- Hilton, B., Hajihashemi, B., Henderson, C. M., & Palmatier, R. W. (2020). Customer Success Management: the next evolution in customer management practice? *Industrial Marketing Management*, 90, 360-369. doi:<https://doi.org/10.1016/j.indmarman.2020.08.001>
- Khan, S. S., & Madden, M. G. (2004). One Class Classification: Taxonomy of Study and Review of Techniques. *The Knowledge Engineering Review*, 1-24. doi:<https://doi.org/10.1017/S026988891300043X>
- Khan, S. S., & Madden, M. G. (2014). One-Class Classification: Taxonomy of Study and Review of Techniques. *The Knowledge Engineering Review*, 29(3), 345-374. doi:<https://doi.org/10.1017/S026988891300043X>

- Kozhikode, R. K. (2016). Dormancy as a Strategic Response to Detrimental Public Policy. *Organization Science*, 27(1), 189-206.
- Lee, H. L., & Tang, C. S. (2017). Socially and Environmentally Responsible Value Chain Innovations: New Operations Management Research Opportunities. *Management Science*, 64(3).
- Leite, W. (2017). *Practical Propensity Score Methods Using R*. Thousand Oaks, CA: Sage Publications, Inc.
- Molnar, C. (2022). *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable (2nd ed)*. Creative Commons. Retrieved from crispdm.github.io/interpretable-ml-book
- Tett, G. (2015). *The Silo Effect: The Peril of Expertise and the Promise of Breaking Down Barriers*. New York: Simon & Schuster.
- Zhou, Z.-H. (2018). A Brief Introduction to Weakly Supervised Learning. *National Science Review*, 44-53.