A Theoretical Analysis of the Lean Startup Method

Onesun Steve Yoo

UCL School of Management, University College London, London E14 5AB, United Kingdom, onesun.yoo@ucl.ac.uk

Tingliang Huang

Business Analytics Department, Carroll School of Management, Boston College, Chestnut Hill, MA 02467, USA, tingliang.huang@bc.edu

Kenan Arifoğlu

UCL School of Management, University College London, London E14 5AB, United Kingdom, k.arifoglu@ucl.ac.uk

The lean startup method (LSM) advocates an iterative and adaptive product development and testing approach to innovation. It recommends firms to build test products, use them to learn about consumer preferences, and modify (or 'pivot') the product design accordingly. It is less straightforward to understand how effective LSM can be however, not least because consumers' responses to the test product are influenced by its quality, price, and design – that is, learning is endogenous to the features of the test product. This paper analyzes the build-test-learn cycle of LSM using an analytical model to understand its micro-foundation and how best to implement it. We find that an optimal test product that maximizes learning should aim either to confirm a more likely product design or to rule out a less likely product design as being the most desired by consumers; have a vertical quality that is neither too high nor too low; and have a higher quality when aiming to confirm than to rule out. We also identify the product–market conditions for which the LSM is more effective. Conceptualizing the LSM via a formal model may help to improve its implementation in practice and to advance further academic research.

Key words: Lean Startup method, entrepreneurship, customer discovery, new product development, Bayesian learning

1. Introduction

Recent decades witnessed many remarkably successful startups disrupting old industries or creating new ones (e.g., Google, Facebook, Airbnb, Uber). These startups take advantage of modern technology shocks (e.g., open-source software, cloud-based computing, mobile platforms) which enable the creation of new business models and give birth to a wide category of novel products and services that are conceptually new to consumers (Economist 2019a, Economist 2019b). When developing such disruptive innovations, it is often challenging for firms to learn about consumers' preferences prior to product launch because consumers would seldom know what they want (Hoeffler 2003, Nijssen 2017). Moreover, with the accompanying decrease in the cost of development, testing, and distribution of the innovative product (Kerr et al. 2014, Nanda and Rhodes-Kropf

2016), the key challenge to creating a successful consumer-oriented innovation has shifted from tackling development challenges to tackling market uncertainty and identifying what appeals to its users (von Hippel 2005). Coupled with a significant increase in the access to financing, e.g., in the form of venture capital, angel investors, and crowdsourcing (Hellmann and Thiele 2015, Ewens et al. 2018), talented entrepreneurs and their startup firms are playing an increasingly prominent role in the innovation economy (Gompers and Lerner 2001, Gompers and Lerner 2006).

This emerging community of innovative startups had found that many traditional guidances on innovation (originally intended for large firm research and development (R&D) or market research for incremental innovation) do not apply well to their contexts in practice (Blank 2005). The Lean Startup method (LSM) was introduced to fill this gap (Ries 2011, Blank 2013). Originating from the software startup community, it promotes an experimentation-based approach to customer discovery and product development. Its build-test-learn loop takes advantage of the prevalent start-up features (e.g., low cost of development and adjustment, agile organizational structure, easy access to customers) to test hypotheses about consumers that are difficult to investigate using traditional market research methods, and to adjust the product development goals accordingly after receiving consumer feedback.

A single iteration of LSM starts when the entrepreneur launches a *test product*, which is a version of the product that communicates the unique selling point of the product but without all the 'bells and whistles' (e.g., a consumer website that has 10% of functionality that founders ultimately envision).[1] Next, by observing consumers' actual purchase behavior or willingness to cross the 'penny gap,'[2] the entrepreneur obtains feedback about consumer preferences. Finally, after assessing the feedback, the entrepreneur modifies the development goals, or *pivots* to better meet consumer needs (Blank 2013). This iterative experimentation approach allows entrepreneurs to test unproven hypotheses about consumers that are otherwise difficult to test (e.g., surveys or market research) and, as a result, has gained significant acceptance among a broad range of entrepreneurs.[3]

---

[1] Throughout this paper, a "test product" refers to the product that is used to learn quickly and cost-effectively about consumer preferences, which might be interpreted as the "minimum viable product (MVP)" in practice (Merryweather 2020). The LSM literature vaguely uses MVP and does not specify what "minimum" or "viable" means (e.g., see Ambler 2017). Due to this vagueness, the term MVP represents a broad notion ranging from a rather prototype-like product to a fully-fledged and marketable product (Ambler 2017). We use "test product" instead of "MVP" to avoid confusion.

[2] The concept of penny gap refers to the huge difference in customer behaviors between taking a product that is free and buying a product that costs a penny. Crossing the penny gap indicates the consumer's willingness to engage in an exchange, which need not be monetary, such as time, effort, or providing information (Patel 2011).

[3] The essence of LSM has been applied in several large firms (e.g., General Electric) in their new product development processes. However, the implementation is more about fostering collaboration between sales and engineering teams, rather than iteration of the build-test-learn loop (Power 2014, Harris 2017).

For example, in its early days, Brian Chesky and Joe Gebbia, the founders of Airbnb believed that people would rent out their rooms to strangers and people would pay to sleep in strangers' rooms. Potential investors and customers they sought after initially all rejected the concept as unrealistic for mainstream market (Hoffman and Yeh 2018). To test their hypothesis, they posted pictures of their lofts on a WordPress blog as a test product; they soon discovered that large segments of population would pay to stay in their rooms, and so they launched a webpage which evolved to become Airbnb (Aberant 2018, Gracia 2019).[4]

Despite its wide adoption, there is limited theoretical understanding about how LSM works. The consumer responses to a test product serve as an imperfect signal about the product-market fit. As in all experiments, due to the random selection of customers, a positive outcome (e.g., sales) does not necessarily indicate that the test product has the right features (i.e., false positive) while a negative outcome (e.g., failure to sell) does not necessarily indicate that the test product has the wrong features (i.e., false negative). The quality, price, and design of the test product affects the consumer feedback and influences both the learning from the test product and the likelihood of attaining false positive and false negative results. In other words, the learning is *endogenous* to the features of the test product. Understanding how accurately an entrepreneur can learn from a test product and benefit from adopting LSM is complex and merits a formal enquiry.

In this paper, we develop a stylized analytical model of LSM to better understand its micro-foundation. Using the model, we aim to explore how entrepreneurs who seek to adopt LSM *should* implement it in practice. Specifically, we address the following research questions: (i) How does a test product's design, quality and price influence the entrepreneur's learning? (ii) Which attribute should the entrepreneur seek to choose or to avoid in a test product? And how should the entrepreneur pivot? (iii) In which product–market environments is LSM more effective, and why?

To study these questions, we consider a profit-maximizing entrepreneur who aims to develop and launch a new product under high market uncertainty. Customers are homogeneous in their valuations for *product quality*, but they are heterogeneous in their valuation for *product design* (i.e., the location in the attribute space). The consumers' tastes are clustered around an *unknown* location in the attribute space. The entrepreneur identifies *two* candidate locations as the epicenter of consumer tastes, and seeks to learn and develop the product design that will sell to most customers (i.e., 'ideal product'). Due to significant market uncertainty, market research helps the entrepreneur form an initial belief about the ideal product location but does not completely reveal it. To learn, he experiments with the consumers by using a test product.

---

[4] Many other similar examples abound, including those of Dropbox, Groupon, Zappos (Aberant 2018), CarsDirect (Allen 2015), PayPal and LinkedIn (Hoffman and Yeh 2018) just to name a few. Similar iterative learning can be found in the nuclear energy sector, where startups such as TerraPower take advantage of advances in supercomputing to learn about the viability of potential nuclear technology (Nanda and Rhodes-Kropf 2016).

We model a single iteration of the LSM as a two-stage decision process as follows. In the first stage, based on his initial belief, the entrepreneur decides what test product to develop (i.e., its quality and price, and design location) and presents it to a single customer randomly selected from the population. The consumer decides to purchase the test product if its quality and price and the fit with his/her idiosyncratic taste provides positive utility. Based on the realized sales outcome, the entrepreneur updates his belief about the ideal product location in a Bayesian manner. In the second stage, using his updated belief, the entrepreneur 'pivots' to a new design location, enhances the product's quality, and launches it to the entire market, and obtains profit.

Our analysis reveals that the optimal test product maximizes learning and minimizes the probability of developing the wrong product, in line with the experts' suggestions (e.g., Ries 2011). Also, it strikes the balance between minimizing false positives and false negatives. We present the following nuanced insights on the optimal implementation of LSM.

First, in terms of the vertical quality and price, the test product should have an intermediate level of quality. If the test product's quality is too high (too low), everybody (nobody) is expected to purchase the test product, and observing the obvious sale (no sale) outcome provides little information about consumer tastes. Second, in terms of design location, the test product should be located on the *exterior* of the two candidate locations (e.g., contains attributes unique to one candidate ideal product) aimed to either confirm a more likely design or rule out a less likely design as ideal. A design location that is between the two candidates (e.g., contains attributes from both candidate designs) achieves little learning and should be avoided. The vertical quality and design location of the test product interact: if the entrepreneur seeks to confirm a design as ideal, he should launch it at a higher vertical quality than when he seeks to rule out a less likely design. Both test products lead to an equal amount of learning, indicating that there are multiple optimal test products and, in particular, that failure to sell a test product can be equally informative as success in selling it.

The outcome of learning influences the pivoting decision. To maximize expected profit, the entrepreneur should pivot to a product design represented by location *interior* to the two candidate locations. This is in contrast to optimal design of the test product. Yet, the pivoting location gets closer to one of two product designs as the entrepreneur learns more from the test product. The accurate learning reduces the need to hedge against the risk of developing the wrong product, and allows the entrepreneur to develop products that are more appealing to customer tastes.

Finally, our parsimonious model provides insights into how product–market environments influence the effectiveness of the LSM. We find that the LSM is more effective in situations with greater market uncertainty and lower heterogeneity in consumer tastes. This is because, in these situations, the risk of developing the wrong product is higher and the optimal test product is more informative

Author: *Analysis of the Lean Startup*
Article submitted to *Marketing Science*; manuscript no. MKSC-18-0058.R4

5

about the location of the ideal product. The LSM's effectiveness is, however, more sensitive to the design of the test product in situations where consumer tastes are less heterogeneous and they rely more on quality for their purchase. In other words, from a practical implementation perspective, the LSM's effectiveness can be highly variable. Thus, our results offer two explanations to the variability of the LSM's effectiveness observed in practice.

We examine several extensions to enrich the key insights. First, we consider the presence of pivoting friction. Since the entrepreneur must divert resources away from development to implement a pivot, pivoting reduces the final product's quality. We find that if the friction is small, the set of optimal test products reduces to a unique test product aimed to confirm the likelier design; if the friction is large, it is optimal not to pivot at all. Second, we assume that the entrepreneur can change the final product's price before launch to further improve profit. We find that the additional profit created by flexible pricing further increases with improved learning, suggesting that pricing and learning are complements. Third, we extend the model to allow for multiple iterations of the build-test-learn loop. The optimal test product allows for successive increase in the quality and price of the test product after each iteration. We also find that a single iteration can be optimal when it leads to sufficient learning. Finally, we examine a setting with three candidate locations for the ideal design, and report consistent insights for the optimal test product and pivot decisions.

The rest of this paper is organized as follows. After reviewing the related literature in §2, we present our model in §3, provide our analysis in §4, and examine the extensions in §5. We conclude in §6. Robustness checks are provided in Appendix A, and proofs in Appendices B and C.

2. Literature Review

The LSM shares many features studied in prior academic literature on new product development and market research. Its uniqueness stems from how those features are combined to address challenges in a different innovation context. We compare the LSM to related literature and highlight our paper's contributions.

Early literature on new product development focuses on the stage-gate method where a project must pass through a series of stages marked by gatekeepers to be launched as a new product. At each stage, a project is reviewed by an internal gatekeeper that collects information to gauge its potential for successful product launch and determines whether to continue to the next stage or stop it (Cooper 1990). The gatekeeper's decision represents an investment under uncertainty, and is formally studied using a real option framework (e.g., see Kwon and Lippman 2011). This innovation process is necessary if each stage entails large costs and lead times, for example, in the case for the developing physical products (see review by Krishnan and Ulrich 2001) and are suitable if the key uncertainty of the project in each stage can be tackled in a self-contained

manner without feedback from downstream stages. It is employed in settings with high technical uncertainty through phases (e.g., in the case of drug development), or in settings where commercial risk can be reduced through market research (e.g., for incremental innovation). Since it enables firms to run multiple projects in parallel (starting with a large set of ideas, and allocating resources to a small set of promising projects), the stage-gate approach is employed fruitfully in large firms' R&D processes to take advantage of economies of scale and reduce the cycle time of successful innovations (e.g., see Girotra et al. 2007). The LSM, however, originates from software startup community and takes advantages of smaller lead time and costs in each stage. It runs through all stages quickly, and relies on learning from feedback in downstream stages.

The innovation process that learns through experimentation has also been extensively studied as an alternative to the stage-gate approach. First, the stage-gate approach was found to be less effective for innovations that were considered more discontinuous (i.e., neither a product nor consumers exist).[5] For example, Lynn et al. (1996) examine case studies of successful disruptive innovations in large firms (e.g., GE and Motorola) and promote the 'probe and learn' innovation approach. Namely, firms should iteratively test the market during product development and alter the product goals, similar to LSM. Second, the advent of technology that enabled rapid prototyping (e.g., computer-aided design (CAD) technology) and testing (e.g., computer simulation) made experimentation feasible. Thomke (1998) studies how the innovation process can adopt these modes of experimentation, taking into consideration the trade-off between cost and accuracy. This emergence of trial-and-error learning prompted Loch et al. (2001) to formally investigate the efficacy between the experimentation approach and stage-gate parallel learning; Sommer and Loch (2004) examine how the level of complexity and uncertainty influences the choice between iteration or stage-gate innovation; while Erat and Kavadias (2008) examine how design complexity influences the number of experimentations as a way of solving technical uncertainty. The key aim of the learning through trial-and-error discussed in the literature is tackling technological development uncertainty aimed at successfully engineering a *functioning* product. In contrast, the LSM employs iteration to tackle the market uncertainty to learn about what consumers want. By studying LSM, we contribute to the literature on experiment-based innovation by formalizing the process.

Mirroring the comparison between stage-gate development and experimentation, a debate had existed regarding the value of business planning in the entrepreneurship literature. Proponents of business plans argued that they allow more effective use of resources and increase decision speed (Delmar and Shane 2003). Those who oppose business plans argue that they divert critical resources away from business operation and prevent entrepreneurs from learning on the job (Bhide 2000).

---

[5] For a more detailed look at terminologies, see Garcia and Calantone (2002).

In a qualitative study, Sarasvathy (2001) and Read et al. (2011) find that successful entrepreneurs follow a process of iteratively acquiring resources through learning and shifting goals, similar in spirit to LSM. Through their meta study, Brinckmann et al. (2010) find that business plans are valuable but their value declines when the startups face higher uncertainty. LSM argues against relying on business plans for startups altogether. Our formal analysis of LSM contributes to the discussion by showing that LSM is more effective in settings where market uncertainty is greater.

Market research plays an important part in determining an innovation's success, and various techniques have been developed to fruitfully elicit consumer preferences. One of the predominant market research techniques is conjoint analysis that asks potential customers to evaluate a set of alternatives in various incrementally innovative domains, such as consumer packaged goods, consumer electronics, or new drinks flavors (Green and Srinivasan 1978, Green and Srinivasan 1990, Mahajan and Wind 1992, Sandor and Wedel 2001, Yu et al. 2011). Other approaches such as virtual concept testing (Dahan and Srinivasan 2000) or crowdsourcing (Bayus 2013) similarly take advantage of existing customers. These techniques leverage the well-established customer base to learn how different attributes (e.g., resolution versus lens quality) should be included in a new product (e.g., camera). Such incremental innovations have only one chance to succeed because customers have many alternatives to choose from, and are often employed in an early phase of a stage-gate process (often independently from product development decisions). In contrast, for discontinuous innovations (no pre-existing products or consumers), consumers cannot conceptualize how they would use a new product (Cooper et al. 2002) and market research becomes unreliable (Hoeffler 2003, Nijssen 2017). The LSM can be viewed as a different type of market research suitable for tackling the uncertainty associated with discontinuous innovations.

Building on this literature, there are streams of papers that model market research and examine how it interacts with innovation (e.g., Ofek and Turut 2008, Lauga and Ofek 2009, Katona 2015, Kwark et al. 2018). These models assume a stage-gate innovation process with market research preceding product development. By investing in market research, firms can reduce or eliminate uncertainty about consumer preferences, which in turn influences their strategy concerning investment in product development. These papers examine how the competitive structure of the industry influence the firms' decision to invest in market research and R&D, and ultimately the innovation output. Our study also examines the interaction between learning and product development. However, our focus is on understanding the nuances of how learning via test product takes place and how LSM can be fruitfully employed to develop an innovation that maximizes the fit with consumer tastes in a noncompetitive setting.

Iteratively launching a test product is related to launching multiple versions of a product. Thus, LSM is related to the literature on strategic versioning of innovations. The aim of versioning is

8

Author: *Analysis of the Lean Startup*
Article submitted to *Marketing Science*; manuscript no. MKSC-18-0058.R4

to launch incremental innovations to maximize earnings or profits. Norton and Bass (1987) study the substitution effect of having multiple product generations in the market simultaneously, and Wilson and Norton (1989) show that product line extensions should be either introduced early in a product's life cycle or not introduced at all, depending on the degree of substitutability between the original product and its extension. Bayus (1992) analyzes the prices for the two generations that maximize the discounted profit after the second product is launched. Bhaskaran et al. (2020) study how startups should sequentially launch products to generate revenue to balance innovation and survival; Lobel et al. (2016) study how firms should behave strategically and launch new versions of existing products when consumers expect them to, while Allon et al. (2019) empirically examine mobile app developers' versioning behavior in response to customer feedback. The aim of a test product in the context of LSM is fundamentally different in that each launch aims to maximize learning about unknown consumer tastes.

Finally, in our learning model, the entrepreneur receives *censored* information from the launch of a test product that he uses to update his prior belief in a Bayesian fashion. Common iterative learning models in new product development and marketing employ search (e.g., Weitzman 1979) or the multi-armed bandit models (e.g., Hauser et al. 2009, Lin et al. 2015), or Bayesian optimization (e.g., Letham et al. 2017). In all these models, the samples used in learning are selected from exogenously given probability distributions. In contrast, in our model the probability distribution is *endogenous* to the design of test product. Moreover, we examine how the entrepreneur can learn about consumer tastes by launching test products with different qualities and designs. We are the first to analytically formalize how the *interplay* between quality and design of test products influences the learning.

3. Model

In this section, we present our model of LSM. We begin with the demand model with unknown consumer tastes (§3.1), then characterize the LSM and its process of learning based on the test product (§3.2), and finally introduce the entrepreneur's problem (§3.3).

3.1. Consumers

We consider a market where consumers are heterogeneous in their tastes of design (i.e., horizontal preferences) and homogeneous in their valuation for product quality.[6] A consumer's taste is represented by her position on the real line $x \in \mathbb{R}$, which represents a collection of attributes, similar to spatial differentiation models (e.g., Hotelling 1929). The target population is distributed according to a probability density $h(x|W)$, where the value W represents the *unknown* epicenter of consumer

---

[6] The distinction is similar to innovation in form (design) and function (Chan et al. 2018).

Author: *Analysis of the Lean Startup*
Article submitted to *Marketing Science*; manuscript no. MKSC-18-0058.R4

9

tastes. For ease of analysis and illustration, we consider finite support $[W - \epsilon, W + \epsilon]$, where the parameter ϵ represents the level of heterogeneity in consumer tastes. We normalize the population size to 1 and all costs to 0.

Suppose that a product with design (i.e., collection of attributes) $\Lambda \in \mathbb{R}$ is launched. Then the disutility experienced by a consumer with taste x due to lack of fit is given by $t \cdot |\Lambda - x|$. The parameter t denotes the strength of a preference for horizontal fit. A high (low) value of t implies that consumers will purchase the product primarily based on horizontal fit (vertical quality) of the product.[7] If V and p, respectively, denote the vertical quality and the price of the product, a consumer located at position x experiences the net surplus:

$$s(x) = V - p - t|\Lambda - x|.$$

The consumer purchases the product if the surplus is positive, i.e., $s(x) > 0$. The demand for the product (Λ, V, p) then corresponds to the fraction of consumers who would experience positive surplus, which is given by:

$$D(\Lambda, V, p) \equiv \int_{-\frac{V-p}{t} + \Lambda}^{\frac{V-p}{t} + \Lambda} h(x|W)dx. \tag{1}$$

The demand depends critically on the design location of the product Λ as illustrated in Figure 1. The left panel represents the case where $\Lambda = W$ so that most of the consumers around Λ have a positive surplus and purchase the product. The right panel represents the case where $\Lambda < W$, in which case the entrepreneur captures only a fraction of potential consumers. Thus, a product $\Lambda = W$ can be thought of as the 'ideal product,' as it would maximize the demand for the product.

3.2. Lean Startup Method (LSM)

We consider a monopolist entrepreneur who aims to introduce a new product to this consumer market.[8] Based on the initial market research, the entrepreneur has narrowed down the possible ideal product to one of two possible product designs $W \in \{0, C\}$, where $\epsilon < C < 2\epsilon$, and forms an initial belief $P(W = 0) = r$ about which of these is the ideal product.[9] The difference between the two product designs, C, represents the degree of market uncertainty (i.e., the higher the value of C,

---

[7] Note that t is not inherently connected to our model or the test product, instead it is common across all products and can be estimated through market research – e.g., efficacy of drugs implies low t whereas taste in flavor or color may have high t. Although this parameter can be uncertain, its uncertainty is lower compared to that of test product design, so we consider exogenously given to simplify the model, as consistent with the literature (e.g., Netessine and Taylor 2007, and Huang et al. 2018).

[8] Due to the novelty of the product, there is higher uncertainty compared to incrementally innovative products. This results in less competition, consistent with the assumption of the LSM (Blank 2013, Allen 2015).

[9] This range of C ensures overlap between $h(x|W = 0)$ and $h(x|W = C)$ to make learning relevant.

Figure 1   **Impact of design location (Λ) on demand. Consumer heterogeneity in taste is represented by an illustrative distribution $h(x|W)$, depicted by the dotted line. The solid lines represent the surplus for consumers with taste represented by location x, $s(x) = V - t|\Lambda - x| - p$. The left panel illustrates the case where $\Lambda = W$, and the right panel illustrates the case where $\Lambda < W$. The shaded area represents the demand.**
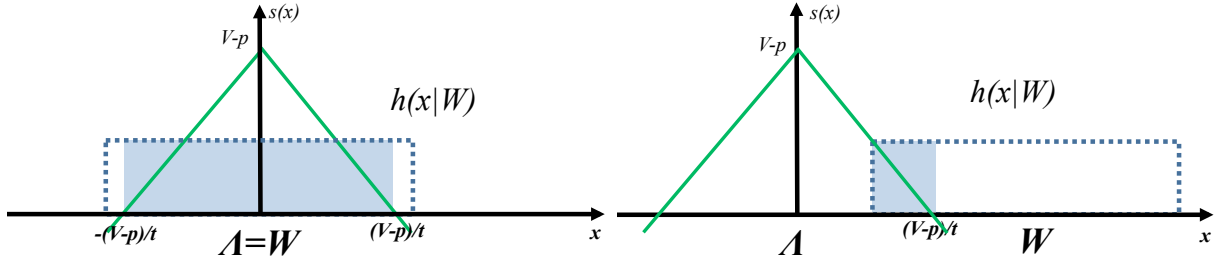
the greater downside risk of developing the wrong product). The entrepreneur seeks to implement the LSM to learn more.

We assume that the final product's quality V and price p are exogenously given, but that the value of $V - p$ is bounded above so that the final product results in a high demand only if the design location is close to the ideal product.

ASSUMPTION 1. $V - p \leq \epsilon t$.

We formalize a single build-test-learn iteration of the LSM, as illustrated in Figure 2. In the first stage, the entrepreneur decides on the test product. To be precise, we model a *test product* defined by its design location ($\lambda \in \mathbb{R}$), and its quality (v) and price (ρ) that is sold to the targeted consumers.

Next, as an abstraction of sampling,[10] we assume that the test product is launched and presented to a single consumer i with location x_i who is randomly picked from the consumer distribution $h(x|W)$. The consumer purchases the test product if her surplus is positive, that is, $s_i \equiv s(x_i) = v - \rho - t|\lambda - x_i| > 0$. Based on this feedback (sale or no sale of test product), the entrepreneur updates his belief about the ideal product's location.
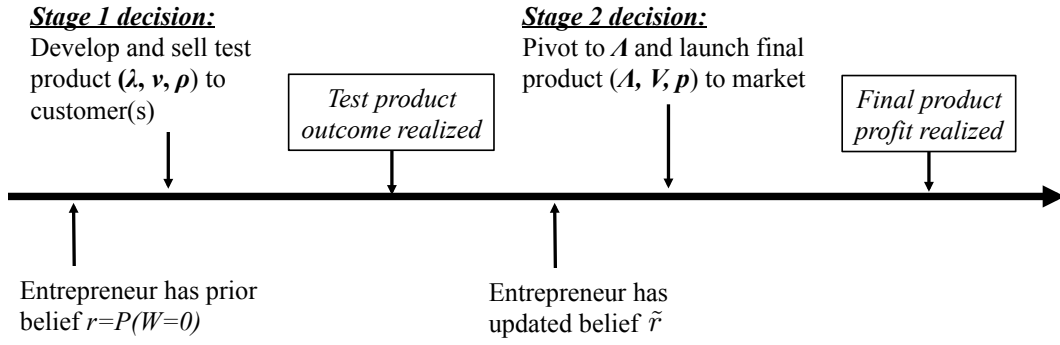
In the second stage, based on his updated beliefs, the entrepreneur pivots the product design from λ to Λ, and adds the 'bells and whistles' to develop vertical quality (V) and price (p), and launches the product. The profit is realized based on demand equation (1).

Learning with a test product. A key feature of our learning model is that the entrepreneur learns through *censored* feedback from the test product – that is, he observes the sales outcome from a purchase, but not the consumer's taste x.[11] Moreover, the test product's quality, price and design influence the consumer's probability of purchase and the entrepreneur's updated belief.

---

[10] We consider sampling multiple customers in the launch of a single test product in Appendix A.

[11] We have also considered the case when the selected consumer reveals her surplus s_i with probability $\xi < 1$. The key insights do not change however. Please see Appendix A.

Author: *Analysis of the Lean Startup*
Article submitted to *Marketing Science*; manuscript no. MKSC-18-0058.R4

11

Figure 2   Sequence of events.

We let $g(s_i|(\lambda, v, \rho), W)$ denote the probability density function of the surplus s_i of a randomly chosen consumer i. We derive $g(s_i|(\lambda, v, \rho), W)$ by using the consumer's taste distribution $h(x|W)$ as Lemma 1 shows.

LEMMA 1 (**Surplus Density Function**). *The probability density function of surplus s_i, which depends on the test product (λ, v, ρ) and $h(x|W)$, is:*

$$g(s_i|(\lambda, v, \rho), W) = \frac{1}{t} \left[h \left(\frac{-(v-\rho)+s_i}{t} + \lambda \,\bigg|\, W \right) + h \left(\frac{v-\rho-s_i}{t} + \lambda \,\bigg|\, W \right) \right].$$

Based on the density $g(s_i|(\lambda, v, \rho), W)$, the probabilities that the test product is sold or not sold can be represented by $\int_0^\infty g(s_i|(\lambda, v, \rho), W)\, ds$ and $\int_{-\infty}^0 g(s_i|(\lambda, v, \rho), W)\, ds$, respectively, which depend on the entrepreneur's choice of test product.

Depending on the outcome of the test product's launch, the entrepreneur updates his belief from $r = P(W = 0)$ to the posterior belief $\tilde{r}$ according to Bayes' rule. Intuitively, if the test product located near $\lambda = 0$ is launched and results in a sale, then the entrepreneur's belief that $W = 0$ will increase, that is, $\tilde{r} > r$; and if it results in no sale, then his belief that $W = 0$ will decrease, that is, $\tilde{r} < r$. The following lemma formalizes this intuition.

LEMMA 2 (**Posterior Belief**). *After launching the test product (λ, v, ρ), the prior belief r is updated to the posterior belief $\tilde{r}$ as follows:*

(i) If sale is observed $(s_i > 0)$, then

$$\tilde{r} = \bar{r}(\lambda, v, \rho | r) = \frac{r \int_0^\infty g(s_i|(\lambda, v, \rho), W = 0)ds}{r \int_0^\infty g(s_i|(\lambda, v, \rho), W = 0)ds + (1-r) \int_0^\infty g(s_i|(\lambda, v, \rho), W = C)ds}.$$

(ii) If no-sale is observed $(s_i \le 0)$, then

$$\tilde{r} = \underline{r}(\lambda, v, \rho | r) = \frac{r \int_{-\infty}^0 g(s_i|(\lambda, v, \rho), W = 0)ds}{r \int_{-\infty}^0 g(s_i|(\lambda, v, \rho), W = 0)ds + (1-r) \int_{-\infty}^0 g(s_i|(\lambda, v, \rho), W = C)ds}.$$

The outcome of the test product yields imperfect information due to the inherent randomness in the selection of a consumer. Launching an ideal test product ($\lambda = W$) can lead to no sale (false negative) because the selected consumer's taste can be located far away from W. Similarly, launching a test product far away from the ideal product can lead to a sale (false positive) if the selected consumer's taste is located close to λ. Since each $g(s_i|(\lambda, v, \rho), W)$ is dependent on the attributes of the test product (λ, v, ρ), the extent of learning (whether posterior $\tilde{r}$ is close to 0 or 1) depends on the choice of test product.

3.3. The Entrepreneur's Problem

We now formalize the entrepreneur's decision problem, which consists of deciding on the test product in the first stage, and pivoting in the second stage.

In the second stage, given the updated belief $\tilde{r}$, the entrepreneur aims to maximize his expected profit from product launch by determining the final pivot location Λ. Thus, the entrepreneur's problem in the second stage is as follows:

$$\max_{\Lambda} \quad \pi(\tilde{r}) \equiv p \int_{\Lambda + \frac{V-p}{t}}^{\Lambda - \frac{(V-p)}{t}} \Big[\tilde{r} \cdot h(x|W=0) + (1-\tilde{r}) \cdot h(x|W=C) \Big] dx. \tag{2}$$

We shall denote the maximum profit obtained after optimal pivot as $\pi^*(\tilde{r})$.

In the first stage, using his initial belief r, the entrepreneur must decide which test product (λ, v, ρ) to launch so as to profitably influence the distribution of the posterior $\tilde{r}$ and maximize the expected profit from the final product launch. (We do not account for the profit from the test product in the first stage because it is from a single customer and is negligible.) The entrepreneur's problem in the first stage is as follows:

$$\max_{(\lambda, v, \rho)} \quad \mathbb{E}_{s_i(\lambda, v, \rho)} \pi^*(\tilde{r}(\lambda, v, \rho|r)) \tag{3}$$
$$= \int_0^\infty \pi^* \Big(\bar{r}(\lambda, v, \rho|r) \Big) \Big[r \cdot g \Big(s_i \Big| (\lambda, v, \rho), W=0 \Big) + (1-r) \cdot g \Big(s_i \Big| (\lambda, v, \rho), W=C \Big) \Big] ds_i$$
$$+ \int_{-\infty}^0 \pi^* \Big(\underline{r}(\lambda, v, \rho|r) \Big) \Big[r \cdot g \Big(s_i \Big| (\lambda, v, \rho), W=0 \Big) + (1-r) \cdot g \Big(s_i \Big| (\lambda, v, \rho), W=C \Big) \Big] ds_i.$$

The first (second) line represents the expected profit in the event of sales (no sales). Depending on the sales outcome, the entrepreneur updates posterior to $\bar{r}$ or $\underline{r}$ as defined in Lemma 2(i)–(ii), respectively. We next characterize the optimal decisions for the entrepreneur implementing the LSM.

4. Analysis

Using backward induction, in §4.1 we first examine the optimal pivoting location Λ^*, and the expression for the optimal second-stage profit given belief $\tilde{r}$, i.e., $\pi^*(\tilde{r})$. Then in §4.2, we examine the

optimal test product decision (λ^*, v^*, ρ^*) in the first stage that maximizes the expected second-stage profit $\mathbb{E}_{s_i(\lambda,v,\rho)}\pi^*(\tilde{r}(\lambda,v,\rho))$. Finally in §4.3, we examine the optimal $\mathbb{E}_{s_i(\lambda^*,v^*,\rho^*)}\pi^*(\tilde{r}(\lambda^*,v^*,\rho^*))$ to gain insights into how the product–market features (C, ϵ and t) influence the effectiveness of the LSM.

To give clear insights, we assume that $h(x|W)$ is uniformly distributed in $[W-\epsilon, W+\epsilon]$.[12]

4.1. Second Stage: Optimal Pivot

If $h(x|W)$ is uniformly distributed within the range $[W-\epsilon, W+\epsilon]$, the entrepreneur's second-stage problem (2) reduces to:

$$
\max_{\Lambda} \; p\left\{\frac{\tilde{r}}{2\epsilon}\left[\min\left(\epsilon, \Lambda + \frac{V-p}{t}\right) - \max\left(-\epsilon, \Lambda - \frac{(V-p)}{t}\right)\right] \right.
$$
$$
\left. + \frac{(1-\tilde{r})}{2\epsilon}\left[\min\left(C+\epsilon, \Lambda + \frac{V-p}{t}\right) - \max\left(C-\epsilon, \Lambda - \frac{(V-p)}{t}\right)\right]\right\}. \tag{4}
$$

Solving the problem in (4), we characterize the optimal pivot location given the updated belief $\tilde{r}$ in Proposition 1.

PROPOSITION 1 **(Optimal Pivot)**. *For any updated belief $\tilde{r}$, the optimal pivot location is:*

$$
\Lambda^* = \begin{cases} \epsilon - \frac{V-p}{t}, & \tilde{r} \geq \frac{1}{2}, \\ C - \epsilon + \frac{V-p}{t}, & \tilde{r} < \frac{1}{2}. \end{cases}
$$

The proposition reveals that it is always optimal to pivot to one of two design locations depending on the updated belief $\tilde{r}$. On the one hand, if $\tilde{r} > 0.5$, the ideal product W is more likely to be located at 0 and it is optimal to pivot to a design location that is closer to 0, $\Lambda^* = \epsilon - \frac{V-p}{t} \geq 0$. On the other hand, if $\tilde{r} < 0.5$, the ideal product W is more likely to be located at C and it is optimal to pivot to a design location closer to C, $\Lambda^* = C - \epsilon + \frac{V-p}{t} < C$.

We observe that, since $V - p \leq \epsilon t$ (by Assumption 1), the optimal pivot location always lies *interior* to the two candidate ideal products, i.e., $\Lambda^* \in [0, C]$. This indicates that the entrepreneur should include features of both potential ideal designs in the final product to hedge against the downside of developing the wrong product. As the quality of the final product increases ($V - p \to \epsilon t$), Λ^* converges towards 0 or C, indicating the reduced need to hedge.

Substituting the optimal pivot Λ^* in (4), we obtain the optimal expected profit as:

$$
\pi^*(\tilde{r}) = p\left[\frac{V-p}{\epsilon t} - \left(\frac{V-p}{\epsilon t} - \left(1 - \frac{C}{2\epsilon}\right)\right)\min\{\tilde{r}, 1-\tilde{r}\}\right]. \tag{5}
$$

The term $\min\{\tilde{r}, 1-\tilde{r}\}$ represents the entrepreneur's probability of pivoting towards the wrong design. If the updated belief $\tilde{r}$ is equal to 0 or 1 (high level of learning from the test product), then

---

[12] We show robustness of the results to this assumption numerically by considering different distributions for $h(x|W)$ in Appendix A.

14

Author: *Analysis of the Lean Startup*
Article submitted to *Marketing Science*; manuscript no. MKSC-18-0058.R4

$\min\{\tilde{r}, 1-\tilde{r}\} = 0$ and the entrepreneur can pivot to the ideal product location to maximize his profit ($p\frac{V-p}{\epsilon t}$). On the other hand, if $\tilde{r}$ is 0.5 (no learning from the test product), then $\min\{\tilde{r}, 1-\tilde{r}\} = 0.5$, indicating that there is 0.5 probability that the entrepreneur pivots to the wrong product, which reduces his expected profit.

4.2. First Stage: Design of the Test Product

In this section, we examine the optimal design of the test product (λ, v, ρ) in the first stage that maximizes the entrepreneur's expected profit $\mathbb{E}_{s_i(\lambda,v,\rho)}\pi^*(\tilde{r}(\lambda, v, \rho))$ in (3). Since the expression (5) is linear in $\min\{\tilde{r}, 1-\tilde{r}\}$, we simplify it as follows.

LEMMA 3 **(Equivalence of First-Stage Decision)**. *The first-stage test product design*

$$\max_{(\lambda,v,\rho)} \mathbb{E}_{s_i(\lambda,v,\rho)}\pi^*(\tilde{r}(\lambda, v, \rho|r)) = \min_{(\lambda,v,\rho)} \mathbb{E}_{s_i(\lambda,v,\rho)}\min\{\tilde{r}(\lambda, v, \rho|r), 1-\tilde{r}(\lambda, v, \rho|r)\}.$$

The test product (λ, v, ρ) that maximizes the expected profit of the entrepreneur from launching the final product is equivalent to the test product that maximizes the expected learning, which is consistent with the expert opinion that test product aims to maximize learning (e.g., Ries 2011). We next present the closed-form expressions for the optimal design of the test product by solving (3).

PROPOSITION 2 **(Optimal Test Product)**.
(i) if $r > 0.5$, then any $(\lambda, v, \rho) \in S_{(i)}$ is optimal, where

$$S_{(i)} \triangleq \Big\{(\lambda, v, \rho) : v - \rho = (\epsilon - \lambda)t, \ \forall \lambda \leq 0\Big\} \cup \Big\{(\lambda, v, \rho) : v - \rho = (-\epsilon + \lambda)t, \ \forall \lambda \geq C/2 + \epsilon\Big\};$$

(ii) if $r < 0.5$, then any $(\lambda, v, \rho) \in S_{(ii)}$ is optimal, where

$$S_{(ii)} \triangleq \Big\{(\lambda, v, \rho) : v - \rho = (C - \epsilon - \lambda)t, \ \forall \lambda \leq C/2 - \epsilon\Big\} \cup \Big\{(\lambda, v, \rho) : v - \rho = -(C - \epsilon + \lambda)t, \ \forall \lambda \geq C\Big\};$$

(iii) if $r = 0.5$, then any (λ, v, ρ) in the following convex hull is optimal:

(a) $\Big\{(\lambda, v, \rho) : v - \rho = (\epsilon - \lambda)t, \ \forall \lambda \leq 0\Big\}$ *and* $\Big\{(\lambda, v, \rho) : v - \rho = (C - \epsilon - \lambda)t, \ \forall \lambda \leq C/2 - \epsilon\Big\}$; *or*

(b) $\Big\{(\lambda, v, \rho) : v - \rho = (-\epsilon + \lambda)t, \ \forall \lambda \geq C/2 + \epsilon\Big\}$ *and* $\Big\{(\lambda, v, \rho) : v - \rho = -(C - \epsilon + \lambda)t, \ \forall \lambda \geq C\Big\}$.

Figure 3 illustrates the optimal test product stated in Proposition 3 in the $(\lambda, v - \rho)$ plane. First, we observe that there are infinitely many optimal test products. Also, the maximum learning occurs when the location of test product λ is on the *exterior* of the interval $(0, C)$ in contrast to the pivoting decision. Moreover, the farther away it is, the higher the vertical quality $v - \rho$ required to elicit the same level of information (this is to compensate for the greater disutility due to lack of fit). In particular, if the prior $r > 0.5$, there are two sets of optimal test products (λ, v, ρ) indicated

Figure 3   **Optimal set of test products illustrated in the (λ, $v - \rho$) plane. The two solid lines (two dotted lines or two shaded areas, respectively) represent the set of optimal test products for $r > 0.5$ ($r < 0.5$ or $r = 0.5$, respectively).**
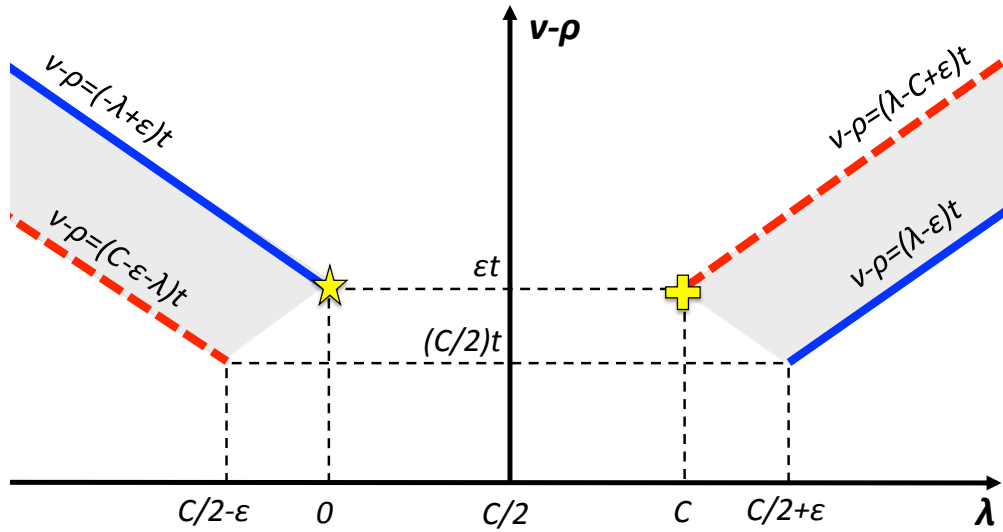
Note. The star ($r \geq 0.5$) and cross ($r \leq 0.5$) represent the optimal test product with small pivoting friction (§5.1).

by the two solid (blue) lines; if the prior $r < 0.5$, there are two sets of optimal test products (λ, v, ρ) indicated by the two dashed (red) lines. If the prior $r = 0.5$, the optimal test product corresponds to the two sets of test products indicated by the shaded area.

Suppose that $r > 0.5$ so that the entrepreneur initially believes that the ideal product W is more likely to be located at 0. The set of optimal test products depicted by the left line $(\lambda, v, \rho) \in \{(\lambda, v, \rho) : v - \rho = (-\lambda + \epsilon)t, \ \forall \lambda \leq 0\}$ represents test product designs whose position λ is closer to 0, more likely to be the ideal design. The quality $v - \rho$ is set so that the test product will *always* sell if $W = 0$ (eliminate false negative). Therefore, if the test product results in a sale (more likely event), the belief that $W = 0$ is further strengthened (i.e., $\tilde{r} > r > 0.5$); if the test product does not sell (less likely event), this reveals that the ideal product is $W = C$ (i.e., $\tilde{r} = 0$). The set of optimal test products depicted on the right line $(\lambda, v, \rho) \in \{(\lambda, v, \rho) : v - \rho = (\lambda - \epsilon)t, \ \forall \lambda \geq C/2 + \epsilon\}$ represents test product designs for which the design position is closer to C, the design less likely to be ideal. The quality $v - \rho$ is set such that the test product will never sell if $W = 0$ (eliminate false positive). Therefore, if it leads to no sales (more likely event), the belief that $W = C$ will be weakened and equivalently the belief that $W = 0$ strengthened (i.e., $\tilde{r} > r > 0.5$); if the test product results in a sale (less likely event), it reveals that the ideal product is $W = C$ (i.e., $\tilde{r} = 0$). That is, the former set of test products aims to observe sales of a likely design to confirm it as ideal, while the latter aims to observe no sales of an unlikely design to rule it out as ideal.

The vertical and horizontal attributes of the optimal test product interact. Namely, it is optimal to either launch a test product more likely to be the ideal product with a higher quality (make

sale of test product more likely), or to launch a test product less likely to be ideal with a lower quality (make no sale of test product more likely). Both strategies lead to equivalent learning for the entrepreneur, suggesting that failure to sell a test product can be just as informative as success in selling it.

Corollary 1 presents further insights into the optimal test product by examining the cross sections of $\mathbb{E}_{s_i(\lambda,v,\rho)}\pi^*(\tilde{r}(\lambda,v,\rho))$ with respect to the quality $(v-\rho)$ and the design (λ) of the test product.

COROLLARY 1 (**Properties of the Optimal Test Product**). *The entrepreneur's expected profit* $\mathbb{E}_{s_i(\lambda,v,\rho)}\pi^*(\tilde{r}(\lambda,v,\rho))$ *is (a) increasing–decreasing in quality* $v-\rho$ *of the test product; and (b) minimized at* $\lambda = C/2$.

Figure 4   **Sensitivity of** $\mathbb{E}_{s_i(\lambda,v,\rho)}\pi^*(\tilde{r}(\lambda,v,\rho))$ **with respect to quality** $v-\rho$ **evaluated at** $\lambda = 0$ **(left panel) and with respect to** λ **evaluated at** $v-\rho = \epsilon t$ **(right panel). Left panel shows that it is optimal to have an intermediate quality for the test product; right panel shows that no learning occurs when** $\lambda = C/2$.
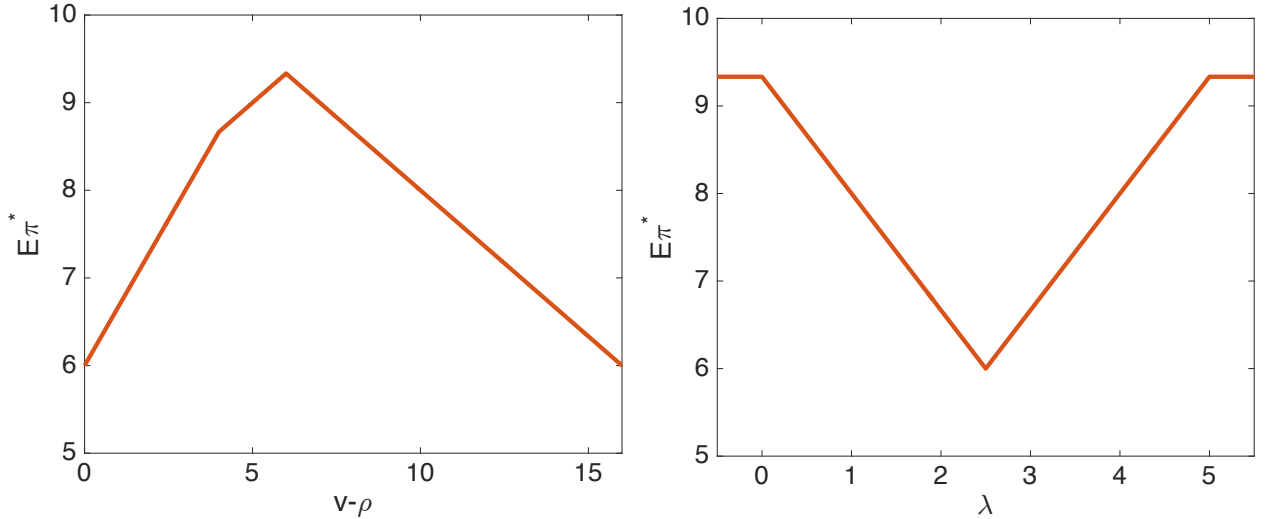
Note. Parameters: $r = 0.5$, $(C, \epsilon, t) = (5, 3, 2)$.

The left panel of Figure 4 illustrates $\mathbb{E}_{s_i(\lambda,v,\rho)}\pi^*(\tilde{r}(\lambda,v,\rho))$ evaluated at $\lambda^* = 0$ as a function of $v - \rho$. We observe that the entrepreneur should launch a test product with a quality $(v-\rho)$ that is neither too high nor too low. As $v - \rho$ increases, the probability of sale increases, reducing the false negative (no sale outcome becomes more informative) but increasing the false positive (sale outcome becomes less informative). If $v - \rho$ is high so that the test product always results in a sale (for both $W = 0$ and $W = C$), receiving the sales outcome would result in no information about the consumer tastes. Similarly, as $v - \rho$ decreases, the probability of sale decreases, increasing the false negative (no sale outcome becomes less informative) but reducing the false positives (sale outcome becomes more informative). If $v - \rho$ is low so that test product never sells (for both $W = 0$

and $W = C$), receiving the 'no sales' outcome does not give information about consumer tastes. The optimal value $(v^* - \rho^*)$ is an intermediate value that balances these two opposing effects. The right panel of Figure 4 illustrates $\mathbb{E}_{s_i(\lambda,v,\rho)}\pi^*(\tilde{r}(\lambda, v, \rho))$ evaluated at $v^* = \epsilon t$ as a function of λ. We observe that learning deteriorates as λ approaches $C/2$. If $\lambda = C/2$ is chosen, no learning occurs regardless of the value $v - \rho$ because sales are equally likely for both $W = 0$ and $W = C$.

4.3. Effectiveness of the LSM

So far, we have identified the optimal design of the test product (λ^*, v^*, ρ^*) and optimal pivoting decision Λ^*. We now investigate the effectiveness of the LSM (peak benefit and sensitivity to implementation) and how it is influenced by the product–market environments.

As the benchmark, we assume that, without LSM, the entrepreneur will develop a product based on his prior belief r. This results in the profit:

$$\pi_0^*(r) = p\left[\frac{V - p}{\epsilon t} - \left(\frac{V - p}{\epsilon t} - \left(1 - \frac{C}{2\epsilon}\right)\right)\min\{r, 1 - r\}\right].$$

The difference between the expression for $\mathbb{E}_{s_i}\pi^*(\tilde{r}(\lambda, v, \rho|r))$ and $\pi_0^*(r)$ boils down to the term

$$\beta(\lambda, v, \rho|r) \triangleq \min\{r, 1 - r\} - \mathbb{E}_{s_i(\lambda,v,\rho)}\min\{\tilde{r}(\lambda, v, \rho|r), 1 - \tilde{r}(\lambda, v, \rho|r)\}.$$

This term $\beta(\lambda, v, \rho|r)$ represents the reduction in probability of developing the wrong product (value of imperfect information) after implementing the LSM. We will refer to this term as *benefit of the LSM* and employ it to gauge the effectiveness of the LSM. Since the LSM is used in settings with high market uncertainty,[13] we focus on its effect when $r = 0.5$, the entrepreneur's prior belief that represents the greatest uncertainty. Correspondingly, we focus on the optimal values $(\lambda^*, v^* - \rho^*) = (0, \epsilon t)$. The following expressions reveal how the contextual parameters (C, ϵ, t) influence the effectiveness of the LSM.

COROLLARY 2 (**Effectiveness of the LSM**). *The effectiveness of the LSM is characterized by its: (i) peak benefit; and (ii) sensitivity to implementation, which are given respectively:*

$$(i) \quad \beta(\lambda^*, v^*, \rho^*|0.5) = \frac{C}{4\epsilon},$$

$$(ii) \quad \left.\frac{\partial\beta(\lambda, v, \rho|0.5)}{\partial(v - \rho)}\right|_{(\lambda,v-\rho)=(0,\epsilon t)} = \begin{cases} \dfrac{1}{4\epsilon t}, & if \ v - \rho < \epsilon t, \\ -\dfrac{1}{4\epsilon t}, & if \ v - \rho > \epsilon t, \end{cases}$$

$$\left.\frac{\partial\beta(\lambda, v, \rho|0.5)}{\partial\lambda}\right|_{(\lambda,v-\rho)=(0,\epsilon t)} = \begin{cases} 0, & if \ \lambda < 0, \\ \dfrac{1}{\epsilon}, & if \ \lambda > 0. \end{cases}$$

---

[13] One can show that $\beta(\lambda, v, \rho|r)$ is unimodal in r and maximized at $r = 0.5$ for any λ, v, and ρ. This is intuitive since $r = 0.5$ represents the prior belief where the entrepreneur has the most uncertainty about the location of the ideal product W, and the benefit of the LSM increases with the entrepreneur's level of uncertainty.

Part (i) shows how the effectiveness of the LSM is influenced by the product–market environments. We observe that the peak benefit of the LSM increases in C and decreases in ϵ. Recall that the parameter C denotes the degree of market uncertainty. A higher C indicates less overlap between the two distributions, and therefore, larger risk of developing the wrong product. The LSM shows greater peak effectiveness when there is a higher C because it reduces the probability of making the wrong choice. The parameter ϵ denotes the extent of consumer heterogeneity. When ϵ is larger, consumers are more spread out so that the risk of developing the wrong product will decrease. Since the benchmark is greater, the additional improvement by the LSM is reduced.

Part (ii) shows how the effectiveness of the LSM is influenced by the design of test product. We observe that the benefit of the LSM is more sensitive to the test product's quality $(v - \rho)$ when the quantity $\epsilon \cdot t$ is low, and to horizontal design λ when ϵ is low. A lower t implies that consumers will purchase primarily based on the vertical quality of a product. In such a case, a small deviation from the optimal quality $v^* - \rho^*$ of the test product could lead to a large fluctuation in the probability of test product's sale. Similarly, a lower ϵ indicates that more consumer tastes are concentrated around a product design. Thus, a small deviation in $v - \rho$ or λ leads to a large deviation in the probability of test product's sale. In other words, when either ϵ or t is low, deviation from the optimal vertical value $v^* - \rho^*$ or design λ^* can quickly increase the false negatives or false positives and reduce the informativeness of the test product. So, the effectiveness of the LSM is sensitive to test product design. The dependence of effectiveness of the lean approach is summarized in Table 1.

	Δ Peak benefit	Δ Sensitivity to implementation
Market uncertainty (C) ↑ :	+	0
Heterogeneity in consumer tastes (ϵ) ↑ :	−	−
Strengths of horizontal preference (t) ↑ :	0	−

Table 1   **Benefit and sensitivity of the LSM, depending on** C, ϵ, **and** t.

Practitioners often find a high variance in the reported effectiveness of the LSM (e.g., Bosch et al. 2013, Dennehy et al. 2016, Hokkanen et al. 2016a, Hokkanen et al. 2016b). Our model helps explain that this heterogeneity may be driven by differences in situations (e.g., one situation has higher C or lower ϵ than another) or due to implementation difficulties (i.e., entrepreneurs operate in the same product–market environment but with low ϵ or t).

5. Extensions

In this section, we examine four essential extensions of the model: (i) pivoting friction; (ii) endogenous pricing of the final product; (iii) multiple iterations; and (iv) more than two potential ideal products.

Author: *Analysis of the Lean Startup*
Article submitted to *Marketing Science*; manuscript no. MKSC-18-0058.R4

19

5.1. Impact of Pivoting Friction

Thus far, we have considered settings (e.g., software) where the pivoting friction is low (i.e., the cost of both pivoting and development are similarly low, hence normalized to zero). In many cases, however, pivoting friction is not negligible and, to overcome and implement the pivot, the firm would need to divert significant resources (e.g., time, money, people) away from development, impacting the final product quality V (Teece et al. 2016, Edison et al. 2017).

We assume that the resource diverted away for pivoting is proportional to the extent of the pivot, $|\Lambda^* - \lambda|$. Then, the quality of the final product is represented by,

$$V^* = V - f|\Lambda^* - \lambda|, \tag{6}$$

where $f > 0$ indicates the magnitude of the friction. (Observe that $f = 0$ eliminates such a trade-off and reduces to the earlier analysis.)

The expression (6) characterizes the trade-off between the final product's quality (V^*) and fit (Λ^*) in the presence of pivoting friction. It endogenizes the final product quality V in the entrepreneur's second-stage problem (2), and replaces the V to V^* (in the interval defining the integral), complicating the pivoting decision Λ^* in the second stage. Moreover, since the extent of the pivot depends on where you pivot from, the optimal second-stage profit depends on λ, so that $\pi^*(\tilde{r})$ becomes $\pi^*(\tilde{r}, \lambda)$. This also influences the entrepreneur's design of the test product in the first stage, given by (3).

We first examine the optimal pivoting decision Λ^*, which depends on the location of the test product λ, the updated belief $\tilde{r}$, and the magnitude of pivot friction f as illustrated next.

PROPOSITION 3 **(Optimal Pivoting given Test Product)**. *Given posterior belief $\tilde{r}$ and λ, Λ^* is as follows:*

(i) If $f < t$, we have $V^* = V - f|\Lambda^* - \lambda|$, where

(a) if $\lambda \in \left(C - \epsilon - \frac{V-p}{t}, \; \epsilon - \frac{V-p}{t}\right)$, then

$$\Lambda^* = \begin{cases} \left(\frac{1}{1+f/t}\right)\left(\frac{V-p}{t} - \epsilon + \lambda\frac{f}{t}\right), & \tilde{r} > \frac{1-f/t}{1+f/t}, \; \lambda - \frac{V-p}{t} < -\epsilon, \\ \lambda, & \tilde{r} > \frac{1-f/t}{1+f/t}, \; \lambda - \frac{V-p}{t} > -\epsilon, \\ \left(\frac{1}{1-f/t}\right)\left(\epsilon - \frac{V-p}{t} + \lambda\frac{f}{t}\right), & \tilde{r} \in \left(\frac{1}{2}\left(1 - \frac{f}{t}\right), \; \frac{1-f/t}{1+f/t}\right), \\ \left(\frac{1}{1+f/t}\right)\left(C - \epsilon + \frac{V-p}{t} + \lambda\frac{f}{t}\right), & \tilde{r} < \frac{1}{2}\left(1 - \frac{f}{t}\right); \end{cases}$$

(b) if $\lambda \in \left(\epsilon - \frac{V-p}{t}, C - \epsilon + \frac{V-p}{t}\right)$, then

$$\Lambda^* = \begin{cases} \left(\frac{1}{1+f/t}\right)\left(C - \epsilon + \frac{V-p}{t} + \lambda\frac{f}{t}\right), & \tilde{r} < \frac{1}{2}\left(1 - \frac{f}{t}\right), \\ \lambda, & \frac{1}{2}\left(1 - \frac{f}{t}\right) < \tilde{r} < \frac{1}{2}\left(1 + \frac{f}{t}\right), \\ \left(\frac{1}{1+f/t}\right)\left(\epsilon - \frac{V-p}{t} - \lambda f/t\right), & \tilde{r} > \frac{1}{2}\left(1 + \frac{f}{t}\right); \end{cases}$$

(c) if $\lambda \in \left(C - \epsilon + \frac{V-p}{t}, \, \epsilon + \frac{V-p}{t}\right)$, then

$$\Lambda^* = \begin{cases} \left(\frac{1}{1+f/t}\right)\left(C + \epsilon - \frac{V-p}{t} - \lambda\frac{f}{t}\right), & \tilde{r} < \frac{2f/t}{1+f/t}, \; \lambda + \frac{V-p}{t} > C + \epsilon, \\ \lambda, & \tilde{r} < \frac{2f/t}{1+f/t}, \; \lambda + \frac{V-p}{t} < C + \epsilon, \\ \left(\frac{1}{1-f/t}\right)\left(C - \epsilon + \frac{V-p}{t} - \lambda\frac{f}{t}\right), & \tilde{r} \in \left(\frac{2f/t}{1+f/t}, \, \frac{1}{2}\left(1+\frac{f}{t}\right)\right), \\ \left(\frac{1}{1+f/t}\right)\left(\epsilon - \frac{V-p}{t} - \lambda\frac{f}{t}\right), & \tilde{r} > \frac{1}{2}\left(1+\frac{f}{t}\right); \end{cases}$$

(ii) If $f > t$, it is optimal to not pivot, and we have $V^* = V$ and $\Lambda^* = \lambda$.

Figure 5   **Illustration of optimal pivoting from a given test product (λ, v) to (Λ^*, V^*). In both panels, the solid (green) path indicates pivoting when $f = 0$ (i.e., no pivoting friction); the red (dotted) path indicates optimal pivoting when $f < t$ (i.e., small pivoting friction); and the blue (dashed) path indicates optimal pivoting when $f > t$ (i.e., large pivoting friction).**
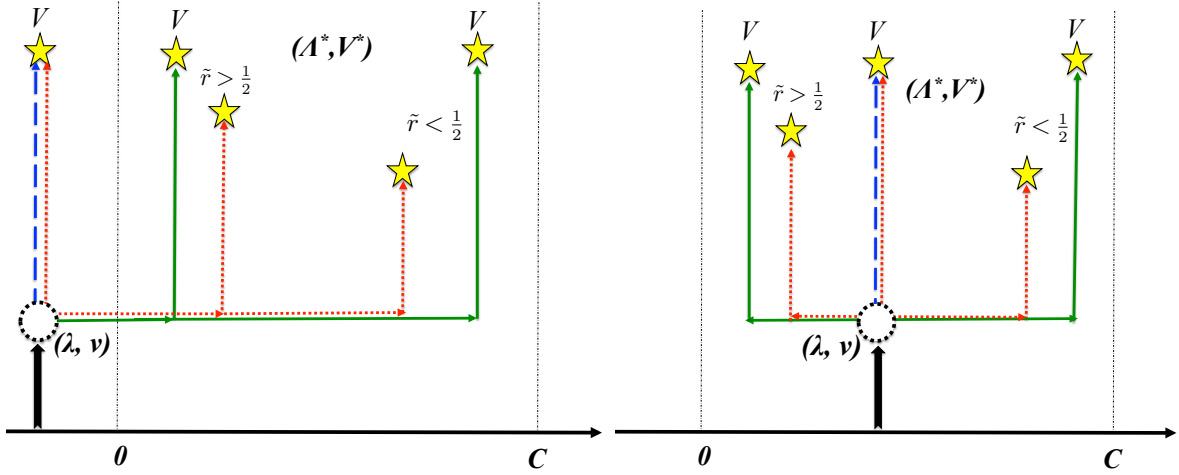

Figure 5 illustrates Proposition 3. Parts (i-a) and (i-b) are illustrated in the left and right panels respectively (part (i-c) is a mirror image of part (i-a)). We first examine the two extreme cases of pivoting cost in both panels. First, if pivoting is frictionless ($f = 0$), then it is optimal to pivot to Λ^* in the interior of $(0, C)$ regardless of the starting location λ as defined in Proposition 1. The entrepreneur will then develop the product Λ^* to quality $V^* = V$. If the pivoting friction is large ($f > t$), any gain from improved fit cannot offset the reduction in quality. Thus, it is optimal not to pivot at all and instead develop the existing product with the highest possible quality, V. This is indicated by the blue (dashed) path in both panels.

When the pivoting friction is moderate ($f \in (0, t)$), the optimal pivoting path is a compromise between the two extreme cases, and is influenced by the location of λ, as indicated by the red (dotted) paths in both panels. The left panel represents the case with $\lambda < 0$ (Proposition 1(i-a)). In this case, if $\tilde{r}$ is sufficiently high (likely that $W = 0$) and the pivoting cost is high, it is optimal to not pivot at all and instead develop the highest-quality product V at the current location λ.

Otherwise, it is optimal to pivot to Λ^* which lies in the interior of $(0, C)$. If $\tilde{r} > 0.5$ then Λ^* is closer to 0 and if $\tilde{r} < 0.5$ it is closer to C. Observe that the quality V^* is reduced according to the extent of pivot $|\Lambda - \lambda|$. The right panel represents the hybrid test product, $\lambda \in (0, C)$ (Proposition 1(i-b)). In this case, it is optimal to pivot towards 0 or C only if the belief $\tilde{r}$ is sufficiently strong, that is, less than $0.5(1 - f/t)$ or greater than $0.5(1 + f/t)$, respectively. Otherwise, it is optimal to not pivot at all and to develop the highest quality at $\Lambda^* = \lambda$. Observe that, in both panels, the optimal pivoting location Λ^* is interior to those with zero pivoting cost (green solid path), due to the increased need to hedge to account for the reduced value offering $V^* - p$ to customers.

We next examine the impact of pivoting cost on the optimal test product.

PROPOSITION 4 **(Optimal Test Product)**.
(i) Suppose $f \to 0$. Then,

$$(\lambda^*, v^* - \rho^*) = \begin{cases} (0, \ \epsilon t), & \text{if } r \geq 0.5, \\ (C, \ \epsilon t), & \text{if } r \leq 0.5. \end{cases}$$

*Furthermore, if $\frac{\partial f}{\partial v} > 0$ then $\rho^* = 0$ and $v^* = \epsilon t$.*
(ii) Suppose $f > t$. Then,

$$(\lambda^*, v^*, \rho^*) = \begin{cases} \left(\epsilon - \frac{V-p}{t}, \ V, \ p\right), & \text{if } r \geq 0.5, \\ \left(C - \epsilon + \frac{V-p}{t}, \ V, \ p\right), & \text{if } r \leq 0.5. \end{cases}$$

First, we observe that, compared to Proposition 3, the presence of small pivoting friction reduces the optimal set of test products to one of two points (star and cross) in Figure 3. These two points minimize the pivot distance compared to other test products that achieve the same learning. Moreover, if the pivoting friction increases in the quality of the test product, then it is optimal to reduce its price $\rho \to 0$ so as to minimize v (while maintaining the same $v - \rho$) to save resources diverted away to pivoting. Next, if the pivoting friction is high ($f > t$), there is no pivoting in the second stage, and the optimal test product choice is identical to the second-stage final product launch decision.

Due to the complexity, we numerically examine the optimal test product for the intermediate range of pivoting costs, $f \in (0, t)$. Figure 6 illustrates the contour plots for $\mathbb{E}_{s_i} \pi^*(\tilde{r}, \lambda, v)$ and the optimal test product (represented by the stars) when $r = 0.5$ as a function of the test product (λ, v, ρ) for $f = 0.01$ (upper left), $f = .99$ (upper right), $f = 0.3$ (lower left), and $f = 0.7$ (lower right). First, we observe that the numerical results confirm our analytical findings in the extreme cases of f. Namely, when $f \to 0$, the optimal position of the test product λ is either 0 or C (Proposition 6(i)) and the optimal quality is $v^* = \epsilon t$; and when $f > t$, the contour plots are vertical

Figure 6 **Contour plots for expected revenue** $\mathbb{E}_{s_i(\lambda,v,\rho)}\pi^*(\tilde{r}(\lambda,v,\rho),\lambda)$ **based on test product's attribute** (λ,v) **for** $f=0.01$ **(upper left),** $f=0.99$ **(upper right),** $f=0.3$ **(lower left), and** $f=0.7$ **(lower right). The stars represent the optimal location of the test product.**
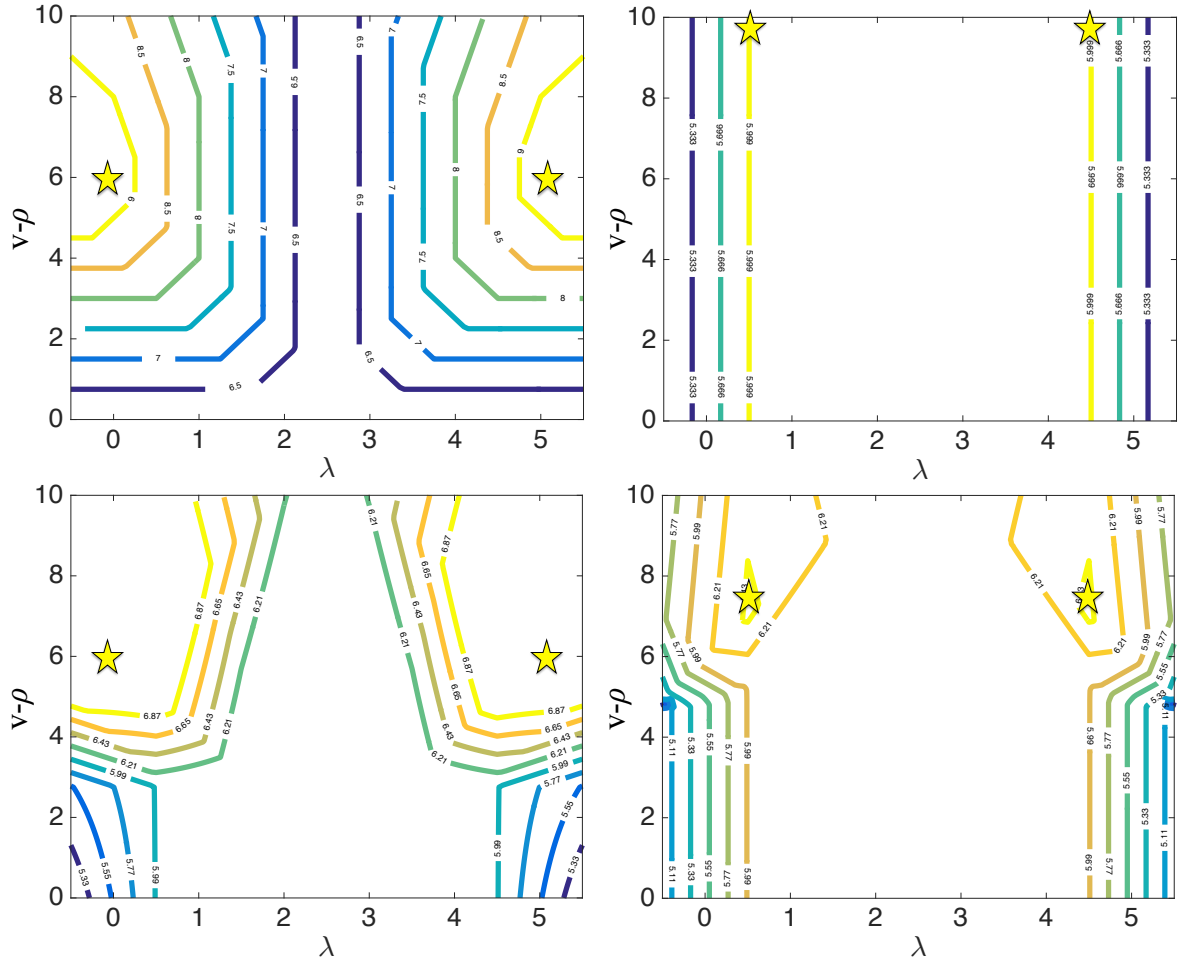
Note. Parameters: $r=0.5$, $(C,\epsilon,t)=(5,3,1)$.

lines and the optimal position of test product is either $\epsilon - \frac{V-p}{t} \in (0,C)$ or $C - \epsilon + \frac{V-p}{t} \in (0,C)$ (Proposition 6(ii)).

Next, we observe that, as the pivoting cost increases from $f=0.01$ (upper left) to $f=0.3$ (lower left) to $f=0.7$ (lower right) to $f=0.99$ (upper right), the shape of the contour plots undergoes a continuous transformation. The optimal test product location, however, does not change continuously. Instead, based on our numerical analysis, there appears to be a threshold value for the pivoting cost f below which the optimal test product corresponds to the case for $f=0$ (i.e., $\lambda^* \in \{0,C\}$) and above which it corresponds to the location for $f>t$, i.e., ($\lambda^* \in \{\epsilon - \frac{V-p}{t}, C - \epsilon + \frac{V-p}{t}\}$).

Author: *Analysis of the Lean Startup*
Article submitted to *Marketing Science*; manuscript no. MKSC-18-0058.R4

23

5.2. Endogenous Pricing

We next examine the case where the final product's price p is endogenously determined by the entrepreneur. Specifically, in the second stage after updating his belief to $\tilde{r}$, the entrepreneur determines the price and pivot location Λ simultaneously. We examine how this additional degree of freedom in the second stage impacts the design of test product in the first stage. Mirroring Assumption 1, we consider the setting where $V \leq Ct$ to ensure that $V - p^* < \epsilon t$ so that finding the product–market fit remains key to generating demand.

The second-stage problem (2) becomes:

$$
\begin{aligned}
\pi^*(\tilde{r}) &\equiv \max_{\Lambda,p} \; p \int_{\Lambda + \frac{V-p}{t}}^{\Lambda - \frac{(V-p)}{t}} \left[\tilde{r} \cdot h(x|W=0) + (1-\tilde{r}) \cdot h(x|W=C) \right] dx \\
&= \max_p \left[\max_\Lambda \; p \int_{\Lambda + \frac{V-p}{t}}^{\Lambda - \frac{(V-p)}{t}} \left[\tilde{r} \cdot h(x|W=0) + (1-\tilde{r}) \cdot h(x|W=C) \right] dx \right] \\
&= \max_p \; p \left[\frac{V-p}{\epsilon t} - \left(\frac{V-p}{\epsilon t} - \left(1 - \frac{C}{2\epsilon} \right) \right) \min\{\tilde{r}, 1-\tilde{r}\} \right].
\end{aligned}
$$

where the first equality is due to joint concavity of the profit function, and the second equality follows from Proposition 2. The optimal price p^* of the final product is characterized in Lemma 4.

LEMMA 4 **(Optimal Price).**

$$
p^* = \frac{V}{2} + \frac{\min\{\tilde{r}, 1-\tilde{r}\}}{\max\{\tilde{r}, 1-\tilde{r}\}} \frac{\epsilon t}{2} \left(1 - \frac{C}{2\epsilon} \right).
$$

Figure 7 (left panel) illustrates Lemma 4 and shows that it depends on the amount of learning. In particular, the entrepreneur reduces the price of the final product if the learning from the test product improves (i.e., $\tilde{r} \to 0$ or $\tilde{r} \to 1$). The intuition is as follows. Suppose that the entrepreneur completely learned the ideal product design so that he pivots to develop it. Then demand will reduce more if he increases the price. In contrast, if the entrepreneur learned nothing, then he can expect his pricing to influence only a small fraction of the consumer population. As such, the demand curve becomes steeper with increased learning. Hence, the entrepreneur would reduce the price to capture higher profits as he learns more.

The objective in the first period is to find the optimal test product that would maximize the optimal expected second-stage profit. The expression for the objective is presented next.

LEMMA 5 **(Expected Profit).** *Let $s_i \equiv s_i(\lambda, v, \rho)$ and $\tilde{r} \equiv \tilde{r}(\lambda, v, \rho | r)$.*

$$
\mathbb{E}_{s_i} \pi^*(\tilde{r}) = \mathbb{E}_{s_i} \frac{\epsilon t}{4} \left[\frac{V}{\epsilon t} + \frac{\min\{\tilde{r}, 1-\tilde{r}\}}{\max\{\tilde{r}, 1-\tilde{r}\}} \left(1 - \frac{C}{2\epsilon} \right) \right] \left[\frac{V}{\epsilon t} - \left[\frac{V}{\epsilon t} - \left(1 - \frac{C}{2\epsilon} \right) \right] \min\{\tilde{r}, 1-\tilde{r}\} \right]. \quad (7)
$$

Figure 7 **Optimal price as a function of learning (left panel) and profit comparison of $\mathbb{E}_{s_i}\pi^*(\tilde{r},\lambda,v)$ with and without pricing (right panel).**
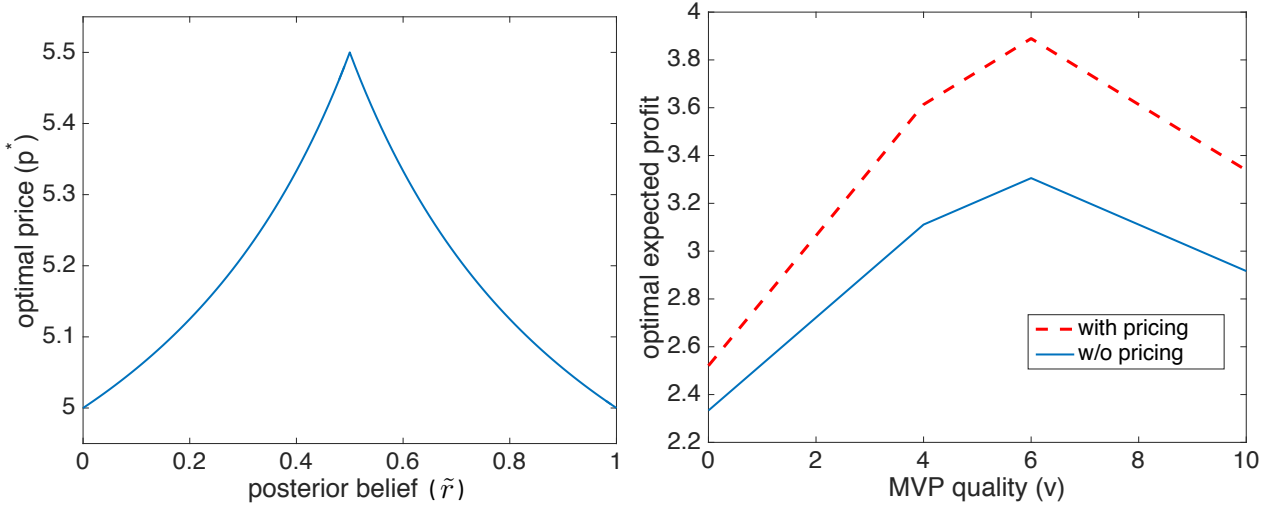
Note. We assume $\rho = 0$ for simplicity. Parameters: $V = 10$, $p = 7$ (right panel), $r = 0.5$, $(C, \epsilon, t) = (5, 3, 2)$.

Unlike the expression in Proposition 2, the expression for expected profit is no longer a linear function of $\mathbb{E}_{s_i}\min\{\tilde{r}, 1 - \tilde{r}\}$. Thus, examining the optimal test product $(\lambda,\ v,\ \rho)$ that maximizes (7) is analytically difficult and we resort to numerical examples. Figure 7 (right panel) illustrates the profit for the case where the prior $r = 0.5$ so that the location of test product is symmetric and we can focus on test product's quality as the measure for learning. The right panel of the figure plots the profit without pricing (solid curve) and with pricing (dashed curve) with respect to test product's quality v.

First, we find that the optimal test product does not change as a result of endogenizing pricing; in other words, $v^* = 6$ for both endogenous and exogenous price. Moreover, comparing the dashed curve (endogenous price) with the solid curve (exogenous price), we find that the difference between the two curves also increases as test product's quality v approaches to the optimal quality $v^* = 6$ (i.e., as learning from the test product increases). This indicates that learning by the test product accentuates the impact of flexible pricing, suggesting a complementary relationship between flexible pricing and learning via the test product.

5.3. Multiple Iterations of the Test Product Launch

Thus far, we assumed a single iteration of the build-test-learn loop. In practice, the entrepreneur has the option of launching the test product and updating his belief multiple times. In such a setting, what should the optimal test product be? To gain insight into the impact of multiple iterations, we assume that the test product can be launched n times. For tractability, however, we restrict the entrepreneur's choice of test product and final product location and assume that $\lambda, \Lambda \in \{0, C\}$.

The optimal design and quality of the test product with n periods to go is presented next.

PROPOSITION 5 (**Optimal Policy for n Iterations**). *Let $\lambda_n \in \{0, C\}$, v_n and ρ_n respectively denote the optimal design location for the test product, its quality and price with n periods to go. If $h(x|W)$ is uniformly distributed in $[W - \epsilon, W + \epsilon]$, $\epsilon > C/2$, then the optimal test product is as follows for all n,*

$$(\lambda_n^*, v_n^* - \rho_n^*) = \begin{cases} (0, \epsilon t), & r \geq 0.5, \\ (C, \epsilon t), & r < 0.5. \end{cases}$$

Proposition 5 reveals that, even with multiple additional iterations, it is optimal to first launch the test product with the design that the entrepreneur believes to be ideal, and with value $v_n - \rho_n = \epsilon t$. Yet, this also implies that it is optimal to launch a test product that increases quality v_n and price ρ_n after each successive iteration, provided that the $v_n - \rho_n$ remains the same.

Figure 8   **Test product is launched over two stages.**
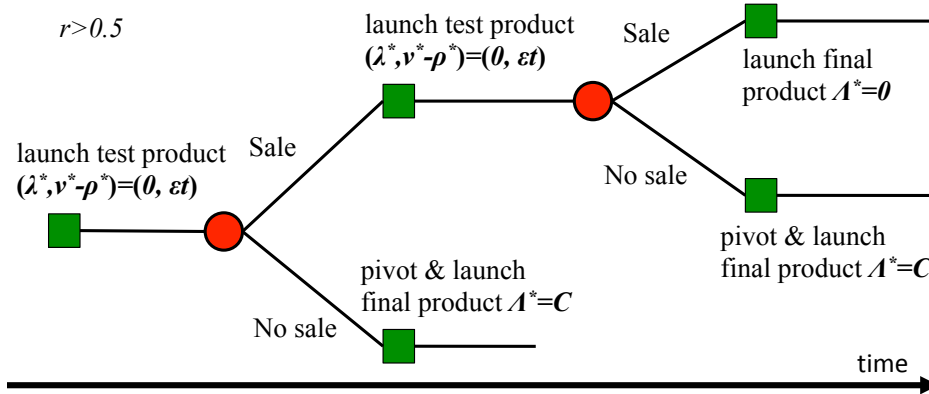
Furthermore, to understand how the policy is implemented dynamically, we examine the optimal policy for launching the test product over two periods, as illustrated in Figure 8. Suppose that $r > 0.5$. In period 1, the entrepreneur launches the test product $(\lambda, v) = (0, \epsilon t)$. If the test product results in no sale (less likely outcome), it is revealed that the ideal product $W = C$ ($\tilde{r} = 0$), pivots and launches the final product with design $\Lambda = C$ without the need to launch an additional test product. If the test product results in a sale (more likely outcome), then the entrepreneur strengthens his belief that $W = 0$ so that $\tilde{r} > r$. In the second build-test-learn iteration, the entrepreneur launches the test product with the same position $\lambda = 0$, but has flexibility in v and ρ. If a sale does not occur (less likely outcome), it is revealed that the ideal product is $W = C$ ($\tilde{r} = 0$) and so the entrepreneur pivots and launches it; if a sale occurs (more likely outcome) he strengthens his belief that $W = 0$ is the ideal product even further, stays the course and launches $\Lambda^* = 0$. The dynamic illustration shows that, if the test product is optimally selected, a single iteration could be optimal if the test

26

Author: *Analysis of the Lean Startup*
Article submitted to *Marketing Science*; manuscript no. MKSC-18-0058.R4

product results in no sale, suggesting that a single iteration may be sufficient if the entrepreneur learns sufficiently well.

We next return to the benefit function $\beta(\lambda, v, \rho|0.5)$ to examine the effectiveness of the LSM that allows for multiple iterations. The following corollary reveals the peak benefit of the LSM involving multiple iterations to go when the entrepreneur has a prior belief r.

COROLLARY 3 (**Effectiveness of the LSM with n Iterations to Go**). *The peak benefit of the LSM with n iterations to go is,*

$$\beta_n(\lambda^*, v^*, \rho^*|0.5) = \frac{1}{2}\left(1 - \left(1 - \frac{C}{2\epsilon}\right)^{n+1}\right).$$

Recall that, in the single iteration setting, the benefit of the LSM $\beta(0.5)$ is $\frac{C}{4\epsilon}$. Thus, the additional iteration of a test product launch increases the benefit of the LSM, but it does so in a marginally decreasing manner.

5.4. More than two candidate ideal products

Thus far, we have assumed that the ideal product was one of two locations $W \in \{0, C\}$. We now extend it to more than two and consider $W \in \{-C, 0, C\}$. As before, we consider $h(x|W)$ to be uniformly distributed between $[W - \epsilon, W + \epsilon]$, $\epsilon \in [C/2, C]$, as illustrated in Figure 9. Observe that

Figure 9   **Three overlapping consumer distributions** $h(x|W)$.
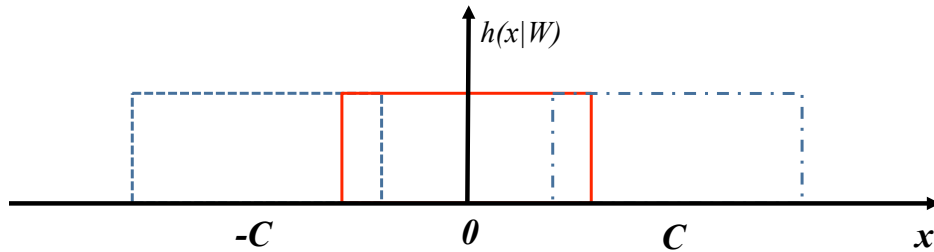
0 is symmetrically positioned between $-C$ and C. To gain analytical insights, again, we assume a discrete set of design locations $\lambda, \Lambda \in \{-C, 0, C\}$. Let us denote the prior probabilities with $r_0 = P(W = 0)$ and $r_C = P(W = C)$, so that $1 - r_0 - r_C = P(W = -C)$. To simplify notations, let

$$D_1 \triangleq \int_{-\frac{V-p}{t}+\Lambda}^{\frac{V-p}{t}+\Lambda} h(x|W = \Lambda)dx, \qquad D_2 \triangleq \int_{-\frac{V-p}{t}+\Lambda}^{\frac{V-p}{t}+\Lambda} h(x| |W - \Lambda| = C)dx,$$

so that D_1 and D_2 are demand from launching the ideal product and adjacent-to-ideal product, respectively. (If $\Lambda = C$ and $W = -C$, then the demand is zero.) The entrepreneur's second-stage problem in (2) becomes,

$$\pi^*(\tilde{r}_0, \tilde{r}_C) = \max\{\tilde{r}_0 D_1 + (1 - \tilde{r}_0)D_2, \ \tilde{r}_C D_1 + \tilde{r}_0 D_2, \ (1 - \tilde{r}_0 - \tilde{r}_C)D_1 + \tilde{r}_0 D_2\}.$$

The following lemma characterizes the entrepreneur's optimal pivoting decision.

PROPOSITION 6 **(Optimal Pivoting)**. *Suppose that after pivoting, we have $\tilde{r}_0$ and $\tilde{r_C}$. Then,*

$$\Lambda^*(\tilde{r}_0, \tilde{r}_C) = \begin{cases} C, & r_C > \frac{D_2}{D_1} + \left(1 - 2\frac{D_2}{D_1}\right)r_0, \ r_C > \frac{1}{2} - \frac{1}{2}r_0 \\ 0, & r_C \leq \frac{D_2}{D_1} + \left(1 - 2\frac{D_2}{D_1}\right)r_0, \ r_C \geq \left(1 - \frac{D_2}{D_1}\right) - 2\left(1 - \frac{D_2}{D_1}\right)r_0 \\ -C, & r_C < \left(1 - \frac{D_2}{D_1}\right) - 2\left(1 - \frac{D_2}{D_1}\right)r_0, \ r_C < \frac{1}{2} - \frac{1}{2}r_0 \end{cases}.$$

The left panel of Figure 10 illustrates the optimal pivot location as a function of posterior r_0 and r_C. We observe that when the posterior is such that $r_0 = r_C = 1/3$ (highest uncertainty), the optimal pivoting location is $\Lambda^* = 0$ as a way to hedge and take advantage of the fact that 0 overlaps with both C and $-C$. This is consistent with our finding in the base model.

Figure 10   Optimal pivot location (left panel) and test product location (right panel) in (r_0, r_C) **space.**
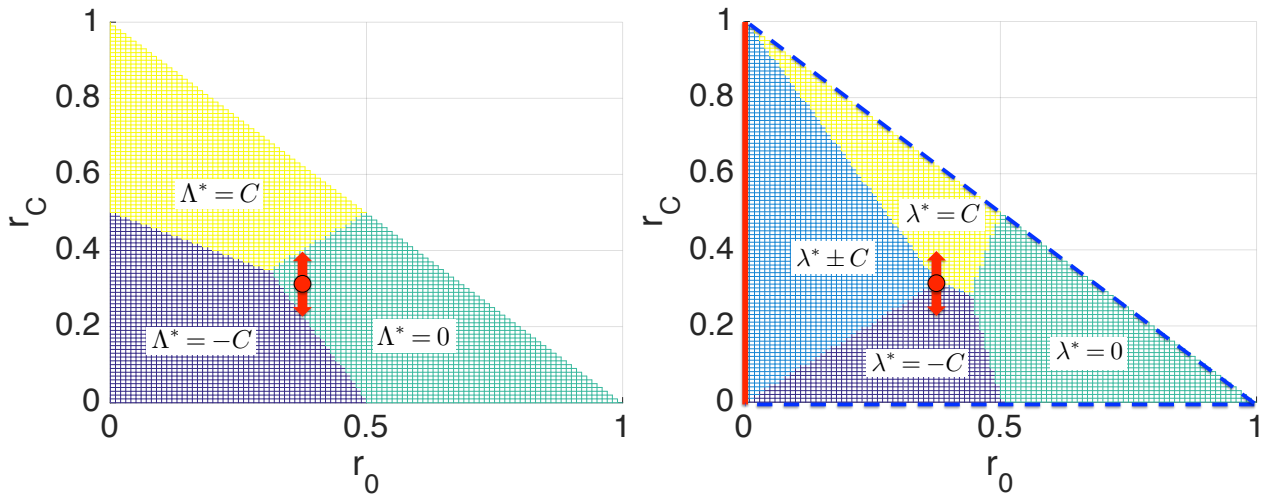
Note. Parameters: $V = 10$, $p = 7.5$, $(C, \epsilon, t) = (5, 3, 1)$.

We next examine the first-stage optimal design of the test product to maximize the expected profit in the final period launch $\mathbb{E}_{s_i} \pi^*(\tilde{r}_0, \tilde{r}_C)$. Suppose that the prior belief is (r_0, r_C) and the entrepreneur launches the test product (λ, v). Let $\bar{r}_0$ and $\bar{r}_C$ (respectively, $\underline{r}_0$ and $\underline{r}_C$) denote the Bayesian posterior beliefs after observing sale (respectively, no sale) of test product. The first-stage problem of the entrepreneur (3) becomes:

$$\max_{\lambda \in \{-C, 0, C\}, v} P(\text{Sale of } (\lambda, v)) \pi^*(\bar{r}_0, \bar{r}_C) + (1 - P(\text{Sale of } (\lambda, v))) \pi^*(\underline{r}_0, \underline{r}_C).$$

The following characterizes the optimal test product for key settings. (The remaining settings are complemented numerically.)

PROPOSITION 7 **(Optimal Test Product)**. *Suppose $h(x|W)$ is uniformly distributed between* $[W - \epsilon, W + \epsilon]$ *where $W \in \{-C, 0, C\}$, $\epsilon \in [C/2, C]$.*
(i) If $r_0 = 0$, then

$$(\lambda^*, v^*) = (\pm C, \epsilon t).$$

(ii) If $r_C = 0$ or $1 - r_0 - r_C = 0$, then respectively,

$$(\lambda^*, v^*) = \begin{cases} (0, \epsilon t) & r_0 \geq 0.5 \\ (-C, \epsilon t) & r_0 \leq 0.5 \end{cases} \quad or \quad (\lambda^*, v^*) = \begin{cases} (0, \epsilon t) & r_0 \geq 0.5, \\ (C, \epsilon t) & r_0 \leq 0.5 \end{cases}.$$

(iii) If $r_0 = 1/3$, then

$$(\lambda^*, v^*) = \begin{cases} (C, \epsilon t) & r_C \geq 1/3 \\ (-C, \epsilon t) & r_C \leq 1/3 \end{cases}.$$

The right panel of Figure 10 illustrates the optimal location of the test product. Proposition 7(i) shows the case when $r_0 = 0$ (vertical solid line) where the ideal product $W \in \{-C, C\}$. In this case, since there is no overlap between $h(x|W = C)$ and $h(x|W = -C)$, launching a test product with quality $v = \epsilon t$ at either location will reveal the ideal product location. If the entrepreneur launches a test product in C (or $-C$) and it sells, he confirms that C (or $-C$) is ideal; if it does not sell, he confirms that instead $-C$ (or C) is ideal. The entrepreneur can then pivot appropriately to launch the ideal product. Proposition 7(ii) shows the case when $r_C = 0$ or $r_0 + r_C = 1$ (two dashed lines on the edges of the triangle) where there is zero probability that C or $-C$ is the ideal product. In this case, the setting reduces to the original setting with two potential ideal products, and it is optimal to launch a test product that is more likely to be the ideal product. For example, if $r_0 \geq 0.5$, then it is optimal to launch $\lambda = 0$. If a sale occurs, it is optimal to stay the course and launch $\Lambda = 0$; if a sale does not occur, it is optimal to pivot and launch C (or $-C$).

Finally, Proposition 7(iii) shows the case when $r_0 = 1/3$ (circle in the middle of the triangle). If $r_C > 1/3$, then C is more likely than $-C$ to be the ideal product. In this case, it is optimal to launch a test product at location $\lambda = C$. If it results in sales, it is optimal to stay the course and launch the final product at $\Lambda = C$; if the test product does not result in sales, it is optimal to pivot to $-C$ to launch the final product in $\Lambda = -C$. The reason the pivot does not happen to 0 is that, if sales do not occur, it becomes more likely that it is because the ideal product is in $-C$ rather than 0. Interestingly, we observe that, when $(r_0, r_C) = (1/3, 1/3)$ (which is the point with the highest uncertainty), the optimal location of the test product $\lambda \neq 0$. This is consistent with our earlier analysis in §4 that the test product should be in the exterior (C or $-C$) to maximize learning.

6. Conclusion

The LSM is widely followed by entrepreneurs for their early product development. Its key paradigm is experimentation using build-test-learn cycles, which involves engaging early with customers to learn what they want from using test products, and then pivoting product development objectives according to its feedback. Despite the influence of this approach, to date there has been no formal characterization of how learning occurs and why LSM is effective. To the best of our knowledge, our paper is the first to critically examine the build-test-learn cycle of the LSM using an analytical

Author: *Analysis of the Lean Startup*
Article submitted to *Marketing Science*; manuscript no. MKSC-18-0058.R4

29

model. Conceptualizing the LSM via a formal model has allowed us to investigate how the features of the test product influence learning, reveal insight into the optimal attributes of the test product (e.g., design, quality, and price), and pivot decision. We have also been able to investigate how the product–market conditions influence LSM's effectiveness.

Our paper can inform practice and future research relating to LSM. Practitioners often find it challenging to gather useful feedback through early versions of a product (Hokkanen et al. 2016a, Hokkanen et al. 2016b), or have difficulty progressing the test product towards the final product that consumers seek (Bosch et al. 2013, Dennehy et al. 2016). Our insights can help entrepreneurs better implement LSM, and to set expectations about its effectiveness when they deploy them. On the theoretical front, Table 1 presents an empirically testable hypothesis about how the product–market environment (i.e., market uncertainty, consumer heterogeneity, and strength of horizontal preference) influences the observed variance in the effectiveness of the LSM. Empirically confirming or disproving such a relationship would be a fruitful direction for further research to enhance our understanding about the LSM.

References

Aberant, J. 2018. The power of minimum viable products (and the key to their success). *CMS Wire, January 23*, https://www.cmswire.com/digital-experience/the-power-of-minimum-viable-products-and-the-key-to-their-success.

Allen, K.R. 2015. *Launching New Ventures: An Entrepreneurial Approach.* Cengage Learning, Boston, MA 7th Edition.

Allon, G., G. Askalidis, R. Berry, N. Immorlica, K. Moon, A. Singh. 2019. When to be agile: Ratings and version updates in mobile apps. *Working Paper available on SSRN.*

Ambler, S. 2017. Defining MVP, MMF, MMP, and MMR. The Disciplined Agile (DA) Framework. Available at https://www.projectmanagement.com/blog-post/61937/Defining-MVP–MMF–MMP–and-MMR, accessed on August 5, 2020.

Bayus B.L. 1992. The dynamic pricing of next generation consumer durables. *Marketing Science*, 11(3):251–265.

Bayus, B.L. 2013. Crowdsourcing new product ideas over time: An analysis of the Dell IdeaStorm community. *Management Science*, 59(1), 226–244.

Bhaskaran, S., S. Erzurumlu, K. Ramachandran. 2020. Sequential innovation by start-ups: Balancing survival and profitability. Forthcoming in *Manufacturing & Service Operations Management.*

Bhide, A.V. 2000. *The Origin and Evolution of New Businesses.* Oxford University Press, New York, NY.

Blank, S. 2005. *The Four Steps to Ephiphany.* Third Edition. K&S Ranch Consulting Inc.

Blank, S. 2013. Why the lean start-up changes everything. *Harvard Business Review*, May, 1-9.

Bosch, J., H. Holmstrom Olsson, J. Bjork, J. Ljungblad. 2013. The early stage software startup development model: A framework for operationalizing lean principles in software startups. In: Fitzgerald B., Conby K., Power K., Valerdi R., Morgan L., Stol KJ. (eds) Lean Enterprice Software and Systems. *Lecture Notes in Business Information Processing*, vol 167. Springer, Berlin, Heidelberg.

Brinckmann, J., D. Grichnik, D. Kapsa. 2010. Should entrepreneurs plan or just storm the castle? A meta-analysis on contextual factors impacting the business planning-performance relationship in small firms. *Journal of Business Venturing*, 25(1), 24–40.

Chan, T.H., J. Mihm, M. Sosa. 2018. On styles in product design: An analysis of US design patents. *Management Science*, 64(3), 1230–1249.

Cooper, R.G. 1990. Stage-gate systems: A new tool for managing new products. *Business Horizons*, 33(3), 44–54.

Cooper, R.G., S.J. Edgett, E.J. Kleinschmidt. 2002 Optimizing the stage-gate process: What best-practice companies do. *Research-Technology Management*, 45(5), 21–27.

Dahan, E., V. Srinivasan. 2000. The predictive power of internet-based product concept testing using visual depiction and animation. *Journal of Product Innovation Management*, 17(2), 99–109.

Delmar, F., S. Shane. 2003. Does business planning facilitate the development of new ventures? *Strategic Management Journal*, 24(12), 1165–1185.

Dennehy, D., L. Kasraian, P. O'Raghallaigh, K. Conboy. 2016. Product market fit frameworks for lean product development. R&D Management Conference 2016, "From Science to Society: Innovation and Value Creation." 3–6 July 2016, Cambridge, UK.

Economist. 2019a. The wave of unicorn IPOs reveals Silicon Valley's groupthink. *The Economist* April 17th Edition.

Economist. 2019b. The entrepreneur's new clothes. *The Economist* September 28th Edition.

Edison, H., N.M. Smorsgard, X. Wang, P. Abrahamsson. 2017. Lean internal startups for software product innovation in large companies: Enablers and inhibitors. *Journal of Systems and Software*, 53(9), 1452–1466.

Erat, S., S. Kavadias. 2008. Sequential testing of product designs: Implications for learning *Management Science*, 54(5), 956–968.

Ewens, M., R. Nanda, M. Rhodes-Kropf. 2018. Cost of experimentation and the evolution of venture capital. *Journal of Financial Economics*, 128, 422–442.

Garcia, R., R. Calantone. 2002. A critical look at technological innovation typology and innovativeness terminology: a literature review. *Journal of Product Innovation Management*, 19(2), 110–132.

Girotra, K., C. Terwiesch, K.T. Ulrich. 2007. Valuing R&D projects in a portfolio: Evidence from the pharmaceutical industry. *Management Science*, 53(9), 1452–1466.

Gompers, P.A., J. Lerner. 2001. The venture capital revolution. *Journal of Economic Perspectives*, 15(2), 145–168.

Gompers, P.A., J. Lerner. 2006. *The Venture Capital Cycle*, 2nd Edition. MIT Press, Cambridge MA.

Gracia, C. 2019 How to build a minimum viable platform. Sharetribe (July 8) https://www.sharetribe.com/academy/how-to-build-a-minimum-viable-platform/

Green, P.E., V. Srinivasan. 1978. Conjoint analysis in consumer research: Issues and outlook. *Journal of Consumer Research*, 5(2), 103–123.

Green, P.E., V. Srinivasan. 1990. Conjoint analysis in marketing: New developments with implications for research and practice. *Journal of Marketing*, 54(4), 3–19.

Harris, E. 2017. GE the startup. Medium.com. Feb 12. https://medium.com/startup-foundation-stories/ge-the-startup-f3317a3fdfee

Hauser, J.R., G.L. Urban, G. Liberali, M. Braun. 2009. Website morphing. *Marketing Science*, 28(2), 202–223.

Hellmann, T., V. Thiele. 2015. Friends or foes? The interrelationship between angel and venture capital markets. *Journal of Financial Economics*, 115(2015), 639–653.

Hoeffler, S. 2003. Measuring preferences for really new products. *Journal of Marketing Research*, 40(4), 406–420.

Hoffman, R., C. Yeh. 2018. *Blitzscaling: The Lightning-Fast Path to Building Massively Valuable Companies.* Penguin Random House, New York.

Hokkanen, L., K. Kuusinen, K. Vaananen. 2016. Early product design in startups: Towards a UX Strategy. In: Abbrahamsson P., Corral L., Olivo M., Russo B. (eds) Product-focused Software Process Improvement. *Lecture Notes in Computer Science*, vol 9459. Springer, Cham.

Hokkanen, L., K. Kuusinen, K. Vaananen. 2016. Minimum viable user experience: A framework for supporting product design in startups. In: Sharp H., Hall T. (eds) Agile Processes, in Software Engineering, and Extreme Programming. *Lecture Notes in Business Information Processing*, vol 251. Springer, Cham.

Hotelling, H. 1929. Stability in competition. *The Economic Journal*, 153(39), 41–57.

Huang, T., C. Liang, and J. Wang. 2018. The value of "bespoke": Demand learning, preference learning, and customer behavior. *Management Science*, 64(7), 3129–3145.

Katona, Z. 2015. Democracy in product design: Consumer participation and differentiation strategies. *Quantitative Marketing and Economics*, 13(12), 359–394.

Kerr, W.R., R. Nanda, M. Rhodes-Kropf. 2014. Entrepreneurship as experimentation. *Journal of Economic Perspectives*, 28(3), 25–48.

Krishnan, V., K.T. Ulrich. 2001. Product development decisions: A review of the literature. *Management Science*, 47(1), 1–21.

Kwark, Y., J. Chen, S. Raghunathan. 2018. User-generated content and competing firms' product design. *Management Science*, 64(10), 4608–4628.

Kwon, H.D., S.A. Lippman. 2011. Acquisition of project-specific assets with Bayesian updating. *Operations Research*, 59(5), 1119–1130.

Lauga, D.O., E. Ofek. 2009. Market research and innovation strategy in a duopoly. *Marketing Science*, 28(2), 373–396.

Letham, B., B. Karrer, G. Ottoni, E. Bakshy. 2017. Constrained Bayesian optimization with noisy experiments. *Technical Report*, arXiv:1706.07094, ArXiv.org. https://arxiv.org/abs/1706.07094.

Lin, S., J. Zhang, J.R. Hauser. 2015. Learning from experience, simply. *Marketing Science*, 34(1), 1–19.

Lobel, I., J. Patel, G. Vulcano, J. Zhang. 2016. Optimizing product launches in the presence of strategic consumers. *Management Science*, 62(6), 1778–1779.

Loch, C.H., C. Terwiesch, S. Thomke. 2001. Parallel and sequential testing of design alternatives. *Management Science*, 47(5), 663–678.

Lynn, G.S., J.G. Morone, A.S. Paulson. 1996. Marketing and discontinuous innovation: The probe and learn process. *California Management Review*, 38(3), 8–37.

Mahajan, V., J. Wind. 1992. New product models: Practice, shortcomings and desired improvements. *Journal of Product Innovation Management*, 9(2), 128–139.

Merryweather, E. 2020. The difference: Prototype vs MVP. https://productschool.com/blog/product-management-2/difference-prototype-mvp/, accessed on August 15, 2020

Nanda, R., M. Rhodes-Kropf. 2016. Financing entrepreneurial experimentation. *Innovation Policy and the Economy*, 16, 1-23.

Netessine, S., T.A. Taylor. 2007. Product line design and production technology. *Marketing Science*, 26(1), 101–117.

Nijssen, E. 2017. *Entrepreneurial Marketing: An Effectual Approach.* Routledge, New York, NY.

Norton J., F. Bass. 1987. A diffusion theory model of adoption and substitution for successive generations of high-technology products. *Management Science*, 33(9):1069–1086.

Ofek, E., O. Turut. 2008. To innovate or imitate? Entry strategy and the role of market research. *Journal of Marketing Research*, 45(5), 575–592.

Patel, N. 2011. Four questions for developing a minimum viable product. *Geekwire.* November 22. https://www.geekwire.com/2011/questions-developing-minimum-viable-product/

Power, B. 2014. How GE Applies Lean Startup Practices. April 23. HBR.org. https://hbr.org/2014/04/how-ge-applies-lean-startup-practices

Read, S., S Sarasvathy, N. Dew, R. Wiltbank, A.V. Ohlsson. 2011. *Effectual Entrepreneurship.* Routledge, New York, NY.

Ries, E. 2011. *The Lean Startup.* Penguin Group, New York, NY.

Sandor, Z., M. Wedel. 2001. Designing conjoint choice experiments using managers' prior beliefs. *Journal of Marketing Research*, 38(4), 430–444.

Sarasvathy S.D. 2001. Causation and effectuation: Toward a theoretical shift from economic inevitability to entrepreneurial contingency. *Academy of Management Review* 26(2), 243–263.

Sommer, S.C., C.H. Loch. 2004. Selectionism and learning in projects with complexity and unforeseeable uncertainty. *Management Science*, 50(10), 1334–1347.

Teece, D., M. Peteraf, S. Leih. 2016. Dynamic capabilities and organizational agility: Risk, uncertainty, and strategy in the innovation economy. *California Management Review*, 58(4), 13–35.

Thomke, S.H. 1998. Managing experimentation in the design of new products. *Management Science*, 44(6), 743–762.

von Hippel, E. 2005. *Democratizing Innovation*. MIT Press, Cambridge, MA.

Weitzman, M.L. 1979. Optimal search for the best alternative. *Econometrica*, 47(3), 641–654.

Wilson L.O., J.A., Norton. 1989. Optimal entry timing for a product line extension. *Marketing Science*, 8(1):1–17.

Yu, J., P. Goos, M. Vandebroek. 2011. Individually adapted sequential Bayesian conjoint-choice designs in the presence of consumer heterogeneity. *International Journal of Research in Marketing*, 28(4), 378–388.

Appendix A: Robustness Checks

We examine robustness checks along three fronts: (i) nonuniform distributions of $h(x|W)$, (ii) sampling multiple customers, and (iii) asymmetric responses to test product with consumer revealing their surplus with probability ξ.

A.1.   Nonuniform distributions

In this section, we examine the robustness of our results to the distributional assumption regarding $h(x|W)$ by considering a shifted and scaled beta distribution with parameters (α, β). Recall that distribution $h(x|W)$ influences both the design of the test product in the first stage and the pivoting decision in the second stage.

First, for a given $\tilde{r}$, the optimal pivoting decision Λ^* for second stage problem (2) satisfies

$$\Lambda^* = \arg\max_{\Lambda} \left\{ \int_{\Lambda - \frac{V-p}{t}}^{\Lambda + \frac{V-p}{t}} \left[\tilde{r} h(x|W=0) + (1-\tilde{r}) h(x|W=C) \right] \right\}.$$

For general distributions $h(x|W)$, this Λ^* does not have a simple closed form, and must be computed numerically. It can be observed that the optimal final product has design $\Lambda^* \in [0, C]$. The complexity is compounded in the optimal design of the test product in the first stage, and the problem (3) is analytically intractable. Therefore the optimal test product is computed numerically.

Recall that the beta distribution with parameter $(\alpha, \beta) = (1,1)$ corresponds to a uniform distribution, which we have examined analytically in §4. Using this case as a benchmark, we will compare two cases of beta distributions, with parameters $(\alpha, \beta) = (2,2)$ and $(\alpha, \beta) = (1,2)$, plotted in the upper left and upper right panels of Figure A-1), respectively. The former reveals the impact of the shape of $h(x|W)$, while the second reveals the impact of asymmetry of $h(x|W)$. The contour plots for expected revenue $\mathbb{E}_{s_i(\lambda,v,\rho)} \pi^*(\tilde{r}(\lambda,v,\rho))$ as a function of test product $(\lambda, v - rho)$ are given in Figure A-1 for $(\alpha, \beta) = (2,2)$ (bottom left panel) and $(\alpha, \beta) = (1,2)$ (bottom right panel).

34

Author: *Analysis of the Lean Startup*
Article submitted to *Marketing Science*; manuscript no. MKSC-18-0058.R4

Figure A-1 **Nonuniform distributions of** $h(x|W)$ **and corresponding contour plots of** $\mathbb{E}_{s_i(\lambda,v,\rho)}\pi^*(\tilde{r}(\lambda,v,\rho))$ **on** $(\lambda, v - \rho)$ **plane.**
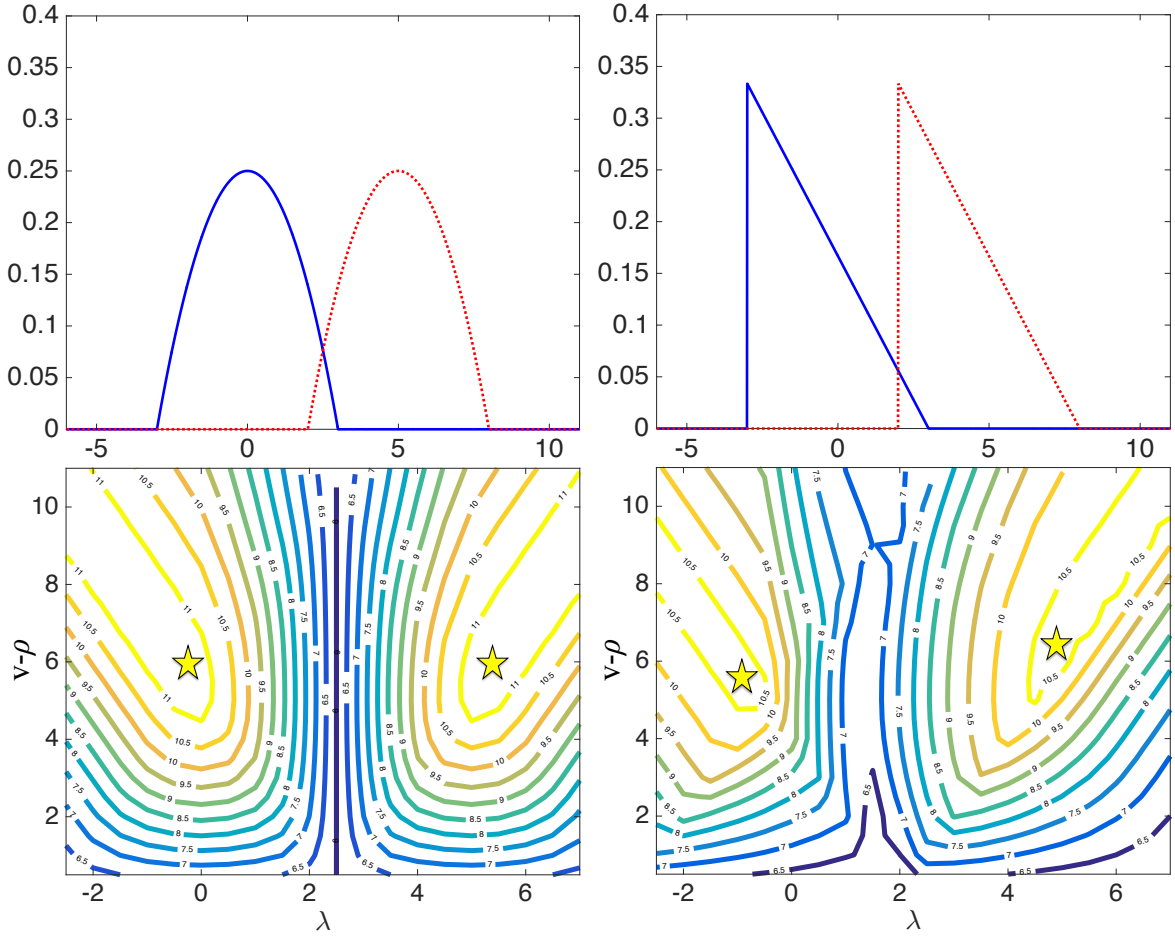
Note. Parameters: $r = 0.5$, $(C, \epsilon, t) = (5, 3, 2)$.

We observe that the contour plot is less "rectangular" (left column) or is shifted to the left to account for greater mass on the left (right column). Apart from these differences, the contour plots and the optimal locations of the test product are similar to those found in §4. Namely, the optimal test products $(\lambda, v - \rho)$ do not change much as a result of change in the distributions $h(x|W)$, the expected revenue $\mathbb{E}_{s_i}\pi^*(\tilde{r})$ is increasing–decreasing in the quality $v - \rho$ of the test product, and learning does not occur when λ is in the interior. In other words, our insights in §4 are robust to the assumptions regarding the underlying distribution $h(x|W)$.

A.2. Sampling Multiple Customers

Our main model in §4 examined learning that results from sampling a single customer. In this section, we examine the impact of sampling multiple independent customers.

Let $\tilde{r}(n)$ denote the updated posterior belief given the outcome from sampling n customers. Let us denote the sales outcome (sale or no sale) of n customers by a vector **s**. If the entrepreneur increases the sample of customers, or as $n \to \infty$, we know that $\tilde{r}(n) \to 0$ or $\tilde{r}(n) \to 1$, resulting in a lower expected probability of

launching the wrong product, $\mathbb{E}_{\mathbf{s}} \min\{\tilde{r}(n), 1 - \tilde{r}(n)\}$. The distribution of the posterior $\tilde{r}(n)$ depends on 2^n combination of each customer's sales outcomes. Hence, it is infeasible to analytically examine the signal for a general $n > 1$, so we resort to numerical analysis.

The top panels of Figure A-2 compare the contour curves for $\mathbb{E}_{\mathbf{s}} \pi^*(\tilde{r}(n))$ on the $(\lambda, v - \rho)$ space for $n = 1$ (upper left panel) and $n = 4$ (upper right panel). We observe that the overall shapes of the contour curves remain the same, and that the optimal test product that maximizes the $\mathbb{E}_{\mathbf{s}} \pi^*(\tilde{r}(n))$ are equivalent for both values of n.

Figure A-2   **Impact of sampling multiple customers. Top row shows the impact of sampling on contours of $\mathbb{E}_{\mathbf{s}} \pi^*(\tilde{r}(n))$ as a function of test product's attribute and optimal test product on $(\lambda, v - \rho)$ plane ($n = 1$ in left panel and $n = 4$ in right panel). Bottom row examines the cross section of $\mathbb{E}_{\mathbf{s}} \pi^*(\tilde{r}(n))$ evaluated at $\lambda = 0$ with respect to $v - \rho$ (left panel) and evaluated at $v - \rho = 6$ with respect to λ (right panel).**
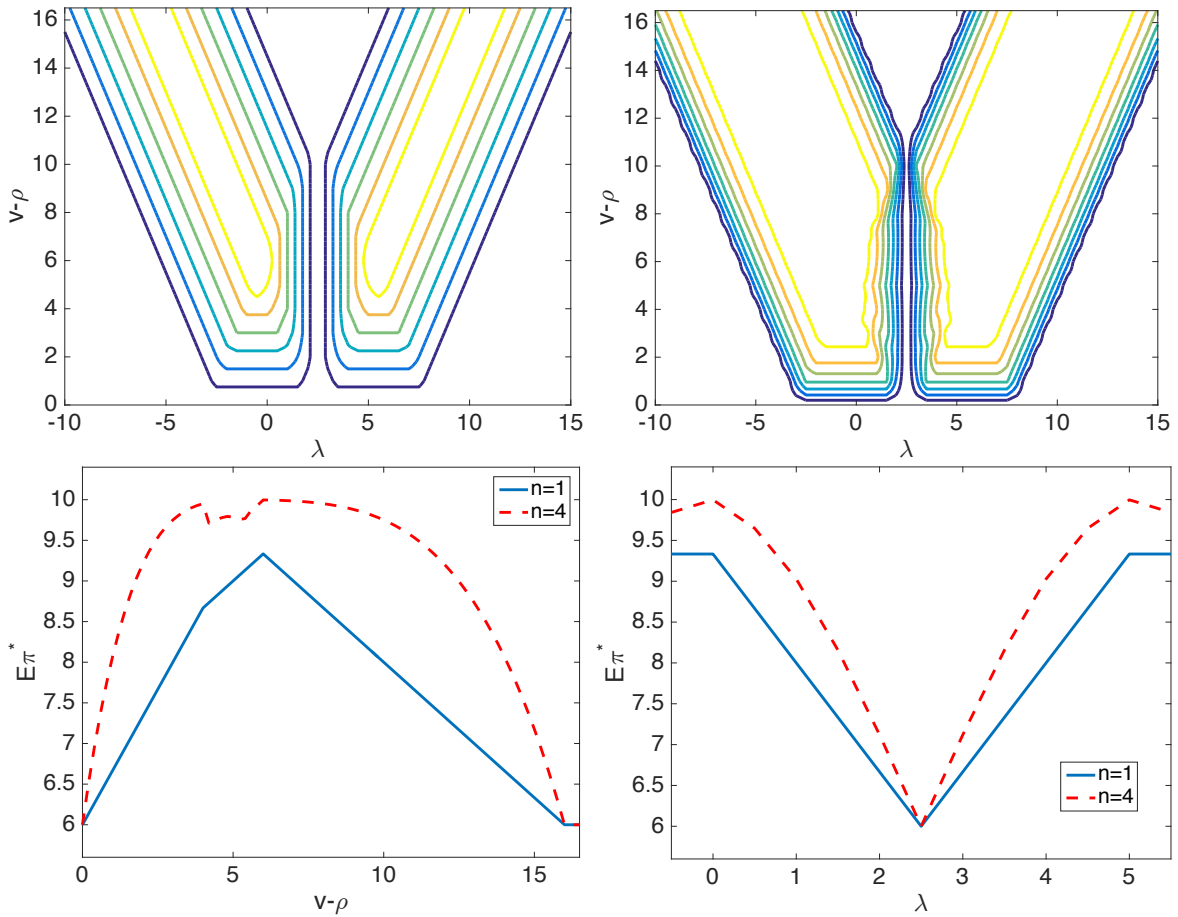
Note. Parameters: $r = 0.5$, $(C, \epsilon, t) = (5, 3, 2)$.

The bottom panels of Figure A-2 compare the cross section of the two contour curves with respect to v (lower left panel) and λ (lower right panel). The solid (blue) curves represent the case where $n = 1$, and the dotted (red) curves represent the case where $n = 4$. Examining the values of $\mathbb{E}_{\mathbf{s}} \pi^*(\tilde{r}(4))$ more closely, we

Author: *Analysis of the Lean Startup*
Article submitted to *Marketing Science*; manuscript no. MKSC-18-0058.R4

36

observe that it is (largely) unimodal in the vertical quality $v - \rho$ and also decreases as $\lambda \to C/2$. However, the value $\mathbb{E}_{\mathbf{s}}\pi^*(\tilde{r}(4))$ is greater than the value of $\mathbb{E}_{\mathbf{s}}\pi^*(\tilde{r}(1))$ and appears less sensitive to the test product's attributes λ and $v - \rho$ at the optimal. Ultimately, the number of samples n influences the distribution of posterior distribution $\tilde{r}$ and improves the effectiveness of LSM. However, the key insights of our analysis regarding the optimal test product and its key features remain same as in §4.

A.3. Asymmetric consumer responses $\xi > 0$

So far, we have assumed that a sampled customer will either purchase the test product or not purchase the test product. We now assume that consumers can provide more detailed feedback (e.g., through comments or posting texts) by revealing their surplus information s_i. To incorporate the asymmetry between feedback after purchase and after no purchase, we assume that consumers reveal surplus with probability ξ only after purchase.

Thus, conditional on sales, the consumer reveals surplus information with probability ξ and fails to reveal it with probability $1 - \xi$. If the surplus is negative, she does not purchase the product and therefore does not reveal her surplus. In other words, the entrepreneur's sales effort may lead to one of three outcomes: (i) no sales (and no surplus information), (ii) sales without surplus information, and (iii) sales with surplus information. This updates the entrepreneur's learning and updates Lemma 2 as follows.

LEMMA A.1 **(Posterior Belief)**. *After launching the test product (λ, v, ρ), the prior belief r is updated to the posterior belief $\tilde{r}$ as follows:*

(i) If the surplus s_i is revealed by the customer, then

$$\tilde{r}(s_i|\lambda, v, \rho) \triangleq \frac{rg(s_i|(\lambda, v, \rho), W = 0)}{rg(s_i|(\lambda, v, \rho), W = 0) + (1 - r)g(s_i|(\lambda, v, \rho), W = C)}.$$

(ii) If only the sale event $s_i > 0$ is observed, then

$$\tilde{r}(s_i > 0|\lambda, v, \rho) \triangleq \frac{r \int_0^\infty g(s_i|(\lambda, v, \rho), W = 0)ds}{r \int_0^\infty g(s_i|(\lambda, v, \rho), W = 0)ds + (1 - r) \int_0^\infty g(s_i|(\lambda, v, \rho), W = C)ds}.$$

(iii) If only the no-sale event $s_i \leq 0$ is observed, then

$$\tilde{r}(s_i \leq 0|\lambda, v, \rho) \triangleq \frac{r \int_{-\infty}^0 g(s_i|(\lambda, v, \rho), W = 0)ds}{r \int_{-\infty}^0 g(s_i|(\lambda, v, \rho), W = 0)ds + (1 - r) \int_{-\infty}^0 g(s_i|(\lambda, v, \rho), W = C)ds}.$$

The first stage problem (3) then becomes:

$$
\max_{(\lambda, v, \rho)} \mathbb{E}_{s_i(\lambda, v, \rho)} \pi^*(\tilde{r}(\lambda, v, \rho))
$$
$$
= \int_0^{v - \rho} \left[\xi \pi^* \left(\tilde{r}(s_i|\lambda, v, \rho), \lambda, v \right) + (1 - \xi)\pi^* \left(\tilde{r}(s_i > 0|\lambda, v, \rho), \lambda, v \right) \right] \times
$$
$$
\left[r \cdot g \left(s_i \middle| (\lambda, v, \rho), W = 0 \right) + (1 - r) \cdot g \left(s_i \middle| (\lambda, v, \rho), W = C \right) \right] ds_i
$$
$$
+ \int_{-\infty}^0 \pi^* \left(\tilde{r}(s_i < 0|\lambda, v, \rho), \lambda, v \right) \left[r \cdot g \left(s_i \middle| (\lambda, v, \rho), W = 0 \right) + (1 - r) \cdot g \left(s_i \middle| (\lambda, v, \rho), W = C \right) \right] ds_i.
$$

Observe that if $\xi = 0$ the entrepreneur's problem and the learning reduces to our original problem in §4.

The following result reveals that ξ has no impact on the optimal choice of the test product.

Author: *Analysis of the Lean Startup*
Article submitted to *Marketing Science*; manuscript no. MKSC-18-0058.R4

37

PROPOSITION A.1 **(Impact of ξ on Optimal Test Product)**. *The set of optimal test products (λ^*, v^*, ρ^*) and $\mathbb{E}_{s_i(\lambda^*, v^*, \rho^*)} \pi^*(\tilde{r}(\lambda^*, v^*, \rho^*))$ when $\xi \in (0, 1)$ are equivalent to those when $\xi = 0$ in Proposition 2.*

In other words, when the test product is optimally designed, the level of learning does not improve from obtaining additional surplus information from consumers. When sales occur and a surplus is revealed, the surplus s_i can sometimes reveal the location of the ideal product (i.e., $\tilde{r} = 1$); however it can also sometimes lead to no updating at all (i.e., $\tilde{r} = r$), thus offsetting the benefits of additional information. When test product is chosen optimally, the benefit of additional surplus information does not manifest in additional profit.

Nevertheless, the additional surplus information does help entrepreneur's profit when he chooses test product suboptimally.

COROLLARY A.1 **(Effectiveness of the LSM)**. *The effectiveness of the LSM is characterized by its: (i) peak benefit; and (ii) sensitivity to implementation, which are given respectively:*

$$(i) \quad \beta(\lambda^*, v^*, \rho^* | 0.5) = \frac{C}{4\epsilon},$$

$$(ii) \quad \left.\frac{\partial \beta(\lambda, v, \rho | 0.5)}{\partial(v - \rho)}\right|_{(\lambda, v - \rho) = (0, \epsilon t)} = \begin{cases} \dfrac{1}{4\epsilon t}, & \text{if } v - \rho < \epsilon t, \\ -\dfrac{\xi}{4\epsilon t}, & \text{if } v - \rho > \epsilon t, \end{cases}$$

$$\left.\frac{\partial \beta(0.5)}{\partial \lambda}\right|_{(\lambda, v - \rho) = (0, \epsilon t)} = \begin{cases} 0, & \text{if } \lambda < 0, \\ \dfrac{1}{\epsilon}, & \text{if } \lambda > 0. \end{cases}$$

Specifically, having surplus information makes LSM more forgiving to overshooting on quality of test product. This is illustrated in Figure A-3. We observe that the expected profit $\mathbb{E}_{s_i(\lambda, v, \rho)} \pi^*(\tilde{r}(\lambda, v, \rho))$ when the entrepreneur overshoots the optimal quality $(v^* - \rho^*)$ increases as $\xi \to 1$.

Appendix B: Probability distribution of surplus s_i and posterior belief $\tilde{r}$, and $\mathbb{E}_{s_i} \min\{\tilde{r}, 1 - \tilde{r}\}$ as a function of test product's attributes (λ, v, ρ).

We consider five ranges of $\lambda \in (-\infty, C/2]$: (1) $\lambda < -\epsilon$, (2) $\lambda \in [-\epsilon, C/2 - \epsilon]$, (3) $\lambda \in [C/2 - \epsilon, 0]$, (4) $\lambda \in [0, C - \epsilon]$, and (5) $\lambda \in [C - \epsilon, C/2]$. (The other five ranges in $\lambda \in [C/2, \infty)$ can be found by interchanging λ and $C - \lambda$ and interchanging $W = 0$ and $W = C$.) These five ranges of λ correspond to five different settings of surplus density functions $g(s_i | W = 0)$ and $g(s_i | W = C)$, as illustrated in Figure B-1.

Recall that consumers reveal their surplus information s_i (with probability ξ) only if the test product results in a sale, i.e., $s_i > 0$. To account for this asymmetry, for each range of λ we must also consider five different ranges of vertical attribute $v - \rho$ to account for five ways in which surplus can be less than 0:

For case (1), we have the following 5 ranges of $(v - \rho)$: (A) $(v - \rho) > (C - \lambda + \epsilon)t$, (B) $(v - \rho) \in [(-\lambda + \epsilon)t, (C - \lambda + \epsilon)t]$, (C) $(v - \rho) \in [(C - \lambda - \epsilon)t, (-\lambda + \epsilon)t]$, (D) $(v - \rho) \in [(-\epsilon - \lambda)t, (C - \epsilon - \lambda)t]$, and (E) $(v - \rho) \in (0, (-\epsilon - \lambda)t]$.

For case (2), we have the following ranges of $(v - \rho)$: (A) $(v - \rho) > (C - \lambda + \epsilon)t$, (B) $(v - \rho) \in [(-\lambda + \epsilon)t, (C - \lambda + \epsilon)t]$, (C) $(v - \rho) \in [(C - \lambda - \epsilon)t, (-\lambda + \epsilon)t]$, (D) $(v - \rho) \in [(\epsilon + \lambda)t, (C - \epsilon - \lambda)t]$, and (E) $(v - \rho) \in (0, (\epsilon + \lambda)t]$.

For case (3), we have the following ranges of $(v - \rho)$: (A) $(v - \rho) > (C - \lambda + \epsilon)t$, (B) $(v - \rho) \in [(-\lambda + \epsilon)t, (C - \lambda + \epsilon)t]$, (C) $(v - \rho) \in [(\lambda + \epsilon)t, (-\lambda + \epsilon)t]$, (D) $(v - \rho) \in [(C - \epsilon - \lambda)t, (\epsilon + \lambda)t]$, and (E) $(v - \rho) \in (0, (C - \epsilon - \lambda)t]$.

Figure A-3   **Sensitivity of $\mathbb{E}_{s_i(\lambda,v,\rho)}\pi^*(\tilde{r}(\lambda,v,\rho))$ with respect to quality $v - \rho$ (left panel) and with respect to λ (right panel). Left panel shows that increasing the chance of receiving surplus information ξ does not improve the optimum but it mitigates the cost of overshooting on quality. Right panel shows that no learning occurs when $\lambda = C/2$.**
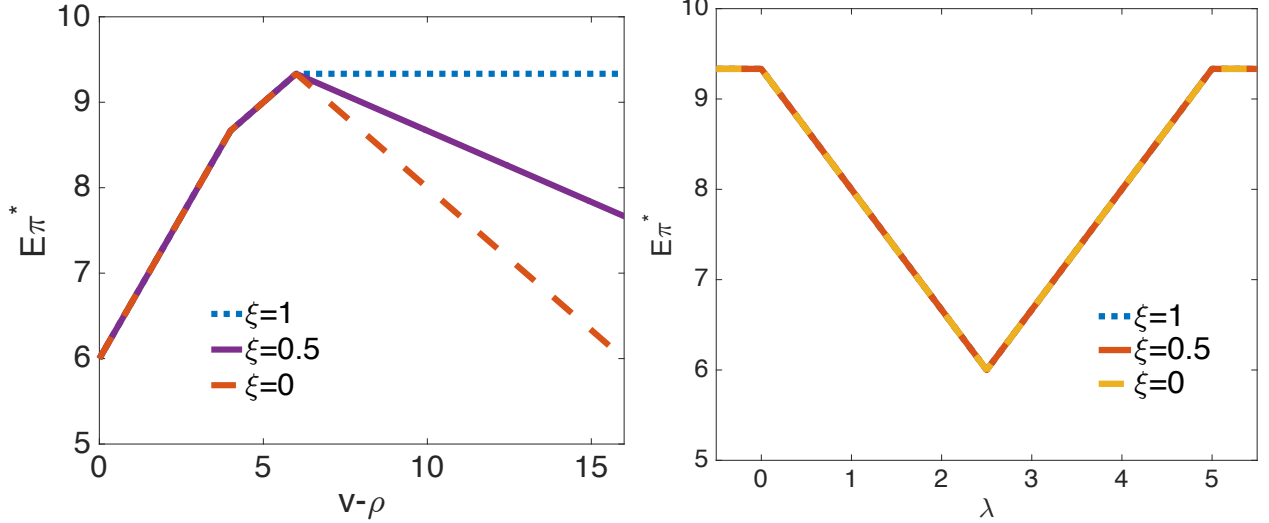
Note. Parameters: $r = 0.5$, $(C, \epsilon, t) = (5, 3, 2)$.

For case (4), we have the following ranges of $(v - \rho)$: (A) $(v - \rho) > (C - \lambda + \epsilon)t$, (B) $(v - \rho) \in [(\lambda + \epsilon)t, (C - \lambda + \epsilon)t]$, (C) $(v - \rho) \in [(-\lambda + \epsilon)t, (\lambda + \epsilon)t]$, (D) $(v - \rho) \in [(C - \epsilon - \lambda)t, (-\lambda + \epsilon)t]$, and (E) $(v - \rho) \in (0, (C - \epsilon - \lambda)t]$.

For case (5), we have the following ranges of $(v - \rho)$: (A) $(v - \rho) > (C - \lambda + \epsilon)t$, (B) $(v - \rho) \in [(\lambda + \epsilon)t, (C - \lambda + \epsilon)t]$, (C) $(v - \rho) \in [(-\lambda + \epsilon)t, (\lambda + \epsilon)t]$, (D) $(v - \rho) \in [(\lambda - C + \epsilon)t, (-\lambda + \epsilon)t]$, and (E) $(v - \rho) \in (0, (\lambda - C + \epsilon)t]$.

This leads to 50 discrete regions in $(\lambda, v - \rho)$-plane as illustrated Figure B-2. a test product with attributes (λ, v, ρ) in a region leads to the same closed form expressions for the probability distribution of s_i and updated posterior belief $\tilde{r}$, and $\mathbb{E}_{s_i}\min\{\tilde{r}, 1 - \tilde{r}\}$. We provide them next for the 25 regions that are labeled. (The other 5 cases are given after changes of values $\lambda \leftrightarrow C - \lambda$ and variables $W = 0 \to W = C$ due to symmetry.)

$\underline{(\lambda, v - \rho) \in (1A)}$

$P(s_i < 0) = 0;$

(i) with probability ξ, surplus information is given if $s_i > 0$,

$P(s_i > 0, 1n2) = (1 - r)\dfrac{C}{2\epsilon}, \quad \tilde{r} = 0,$

$P(s_i > 0, 2n3) = 1 - \dfrac{C}{2\epsilon}, \quad \tilde{r} = r,$

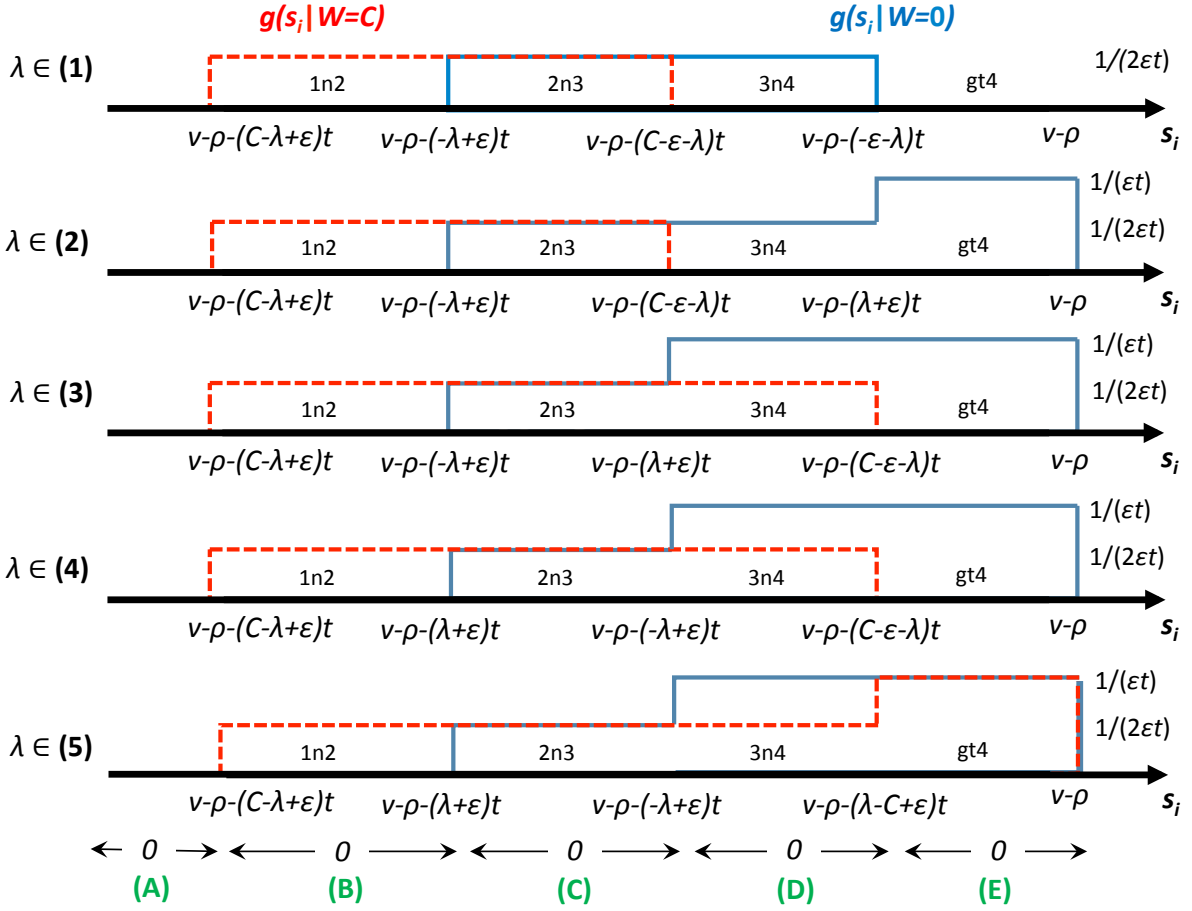$P(s_i > 0, 3n4) = r\dfrac{C}{2\epsilon}, \quad \tilde{r} = 1;$

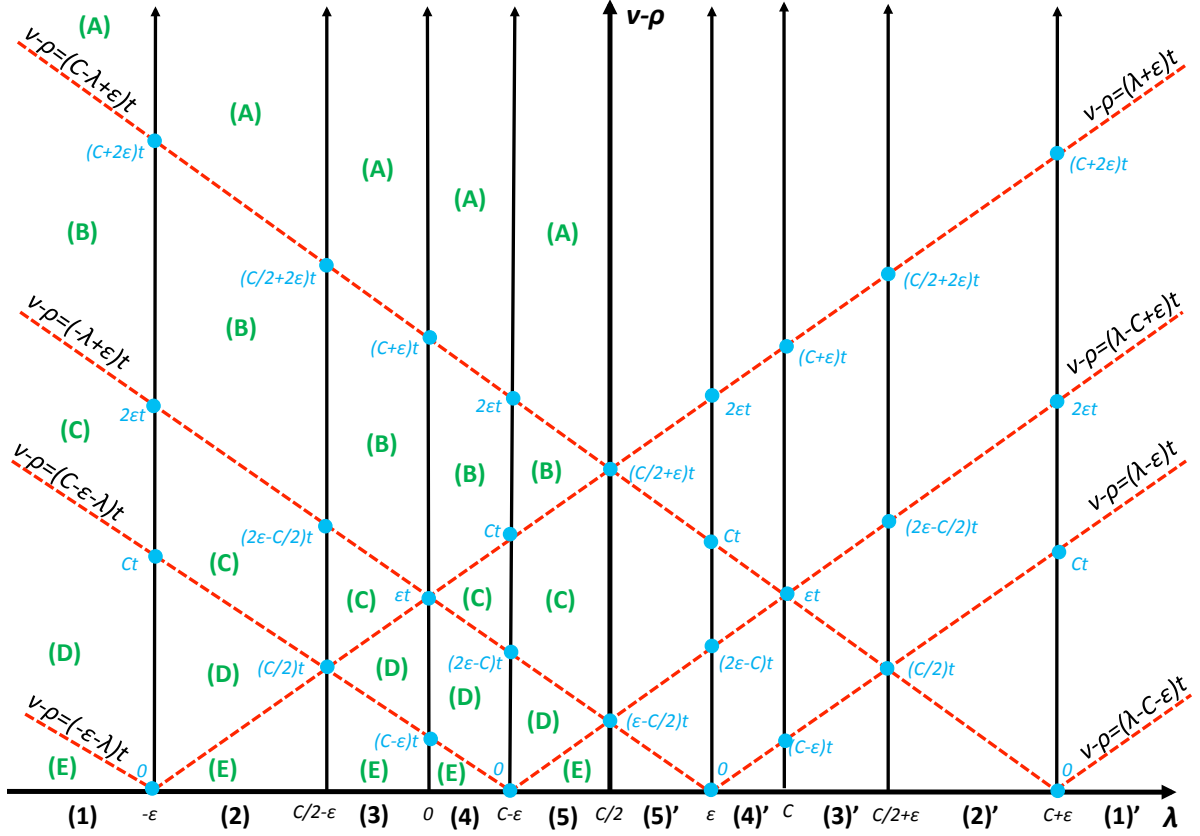(ii) with probability $1 - \xi$, surplus information is NOT given if $s_i > 0$,

$P(s_i > 0) = 1, \quad \tilde{r} = r.$

$\mathbb{E}\min\{\tilde{r}, 1 - \tilde{r}\} = \xi\left(1 - \dfrac{C}{2\epsilon}\right)\min\{r, 1 - r\} + (1 - \xi)\min\{r, 1 - r\}.$

$\underline{(\lambda, v - \rho) \in (1B)}$

Author: *Analysis of the Lean Startup*
Article submitted to *Marketing Science*; manuscript no. MKSC-18-0058.R4

39

Figure B-1 **Illustration of two different probability density functions** $g(s_i|W=0)$ **and** $g(s_i|W=C)$ **for 5 different scenarios based on position of** λ. **(The other 5 cases are given after changes of values** $\lambda \leftrightarrow C-\lambda$ **and variables** $W=0 \rightarrow W=C$ **due to symmetry.)**

$$P(s_i < 0) = (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{1}{2} + \frac{C-\lambda}{2\epsilon}\right), \quad \tilde{r}=0;$$

(i) with probability ξ, surplus information is given if $s_i > 0$,

$$P(s_i > 0, 1n2) = (1-r)\left(\frac{v-\rho}{2\epsilon t} - \frac{1}{2} + \frac{\lambda}{2\epsilon}\right), \quad \tilde{r}=0,$$

$$P(s_i > 0, 2n3) = 1 - \frac{C}{2\epsilon}, \quad \tilde{r}=r,$$

$$P(s_i > 0, 3n4) = r\frac{C}{2\epsilon}, \quad \tilde{r}=1;$$

(ii) with probability $1-\xi$, surplus information is NOT given if $s_i > 0$,

$$P(s_i > 0) = 1 - (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{1}{2} + \frac{C-\lambda}{2\epsilon}\right), \quad \tilde{r} = \frac{r}{1-(1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{1}{2} + \frac{C-\lambda}{2\epsilon}\right)}.$$

$$\mathbb{E}\min\{\tilde{r}, 1-\tilde{r}\} = \xi\left(1 - \frac{C}{2\epsilon}\right)\min\{r, 1-r\} + (1-\xi)\min\left\{r, 1-(1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{1}{2} + \frac{C-\lambda}{2\epsilon}\right) - r\right\}.$$

$(\lambda, v-\rho) \in (1C)$

$$P(s_i < 0) = r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon}\right) + (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(C-\lambda+\epsilon)}{2\epsilon}\right)$$

40

Author: *Analysis of the Lean Startup*
Article submitted to *Marketing Science*; manuscript no. MKSC-18-0058.R4

Figure B-2   50 discrete regions. We analyze 25 regions that are labeled.

$$= \left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda + \epsilon)}{2\epsilon} \right) + (1-r)\frac{C}{2\epsilon}, \quad \tilde{r} = \frac{r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda + \epsilon)}{2\epsilon} \right)}{r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda + \epsilon)}{2\epsilon} \right) + (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(C-\lambda + \epsilon)}{2\epsilon} \right)};$$

(i) with probability ξ, surplus information is given if $s_i > 0$,

$$P(s_i > 0, 2n3) = \frac{v-\rho}{2\epsilon t} - \frac{(C-\epsilon - \lambda)}{2\epsilon}, \quad \tilde{r} = r,$$

$$P(s_i > 0, 3n4) = r\frac{C}{2\epsilon}, \quad \tilde{r} = 1;$$

(ii) with probability $1 - \xi$, surplus information is NOT given if $s_i > 0$,

$$P(s_i > 0) = 1 - \left[r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda + \epsilon)}{2\epsilon} \right) + (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(C-\lambda + \epsilon)}{2\epsilon} \right) \right]$$

$$\tilde{r} = \frac{r\left[1 - \left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda + \epsilon)}{2\epsilon} \right) \right]}{1 - \left[r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda + \epsilon)}{2\epsilon} \right) + (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(C-\lambda + \epsilon)}{2\epsilon} \right) \right]}.$$

$$\mathbb{E}\min\{\tilde{r}, 1 - \tilde{r}\} = \xi\left[\min\left\{ r\left(-\frac{v-\rho}{2\epsilon t} + \frac{(-\lambda + \epsilon)}{2\epsilon} \right), (1-r)\frac{C}{2\epsilon} \right\} + \left(\frac{v-\rho}{2\epsilon t} - \frac{C-\epsilon - t}{2\epsilon} \right)\min\{r, 1-r\} \right]$$

$$+ (1-\xi)\left[\min\left\{ r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda + \epsilon)}{2\epsilon} \right), (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(C-\lambda + \epsilon)}{2\epsilon} \right) \right\} + $$

$$\min\left\{ r\left[1 - \left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda + \epsilon)}{2\epsilon} \right) \right], (1-r)\left[1 - \left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(C-\lambda + \epsilon)}{2\epsilon} \right) \right] \right\} \right].$$

$\underline{(\lambda, v-\rho) \in (1D)}$

Author: *Analysis of the Lean Startup*
Article submitted to *Marketing Science*; manuscript no. MKSC-18-0058.R4

41

$$P(s_i < 0) = r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon}\right) + (1-r), \quad \tilde{r} = \frac{r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon}\right)}{\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon}\right) + (1-r)};$$

(i) with probability ξ, surplus information is given if $s_i > 0$,

$$P(s_i > 0, 3n4) = r\left(\frac{v-\rho}{2\epsilon t} + \frac{-\epsilon-\lambda}{2\epsilon}\right), \quad \tilde{r} = 1;$$

(ii) with probability $1 - \xi$, surplus information is NOT given if $s_i > 0$,

$$P(s_i > 0) = 1 - \left[r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon}\right) + (1-r)\right], \quad \tilde{r} = 1.$$

$$\mathbb{E}\min\{\tilde{r}, 1-\tilde{r}\} = \xi\min\left\{r\left(-\frac{v-\rho}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon}\right), 1-r\right\} + (1-\xi)\min\left\{r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon}\right), (1-r)\right\}.$$

$\underline{(\lambda, v - \rho) \in (1E)}$

$$P(s_i < 0) = 1, \quad \tilde{r} = r.$$

$$\mathbb{E}\min\{\tilde{r}, 1-\tilde{r}\} = \min\{r, 1-r\}.$$

$\underline{(\lambda, v - \rho) \in (2A)}$

$$P(s_i < 0) = 0;$$

(i) with probability ξ, surplus information is given if $s_i > 0$,

$$P(s_i > 0, 1n2) = (1-r)\frac{C}{2\epsilon}, \quad \tilde{r} = 0,$$

$$P(s_i > 0, 2n3) = r\left(1 - \frac{C}{2\epsilon}\right) + (1-r)\left(1 - \frac{C}{2\epsilon}\right) = 1 - \frac{C}{2\epsilon}, \quad \tilde{r} = r,$$

$$P(s_i > 0, 3n4) = r\frac{C - 2\lambda - 2\epsilon}{2\epsilon}, \quad \tilde{r} = 1,$$

$$P(s_i > 0, gt4) = r\frac{\lambda + \epsilon}{\epsilon}, \quad \tilde{r} = 1;$$

(ii) with probability $1 - \xi$, surplus information is NOT given if $s_i > 0$,

$$P(s_i > 0) = 1, \quad \tilde{r} = r.$$

$$\mathbb{E}\min\{\tilde{r}, 1-\tilde{r}\} = \xi\left(1 - \frac{C}{2\epsilon}\right)\min\{r, 1-r\} + (1-\xi)\min\{r, 1-r\}.$$

$\underline{(\lambda, v - \rho) \in (2B)}$

$$P(s_i < 0) = (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C - \lambda + \epsilon}{2\epsilon}\right), \quad \tilde{r} = 0;$$

(i) with probability ξ, surplus information is given if $s_i > 0$,

$$P(s_i > 0, 1n2) = (1-r)\left(\frac{v-\rho}{2\epsilon t} - \frac{(-\lambda+\epsilon)}{2\epsilon}\right), \quad \tilde{r} = 0,$$

$$P(s_i > 0, 2n3) = 1 - \frac{C}{2\epsilon}, \quad \tilde{r} = r,$$

$$P(s_i > 0, 3n4) = r\frac{C - 2\lambda - 2\epsilon}{2\epsilon}, \quad \tilde{r} = 1,$$

$$P(s_i > 0, gt4) = r\frac{\lambda + \epsilon}{\epsilon}, \quad \tilde{r} = 1;$$

(ii) with probability $1 - \xi$, surplus information is NOT given if $s_i > 0$,

$$P(s_i > 0) = 1 - (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C - \lambda + \epsilon}{2\epsilon}\right), \quad \tilde{r} = \frac{r}{1 - (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon}\right)}.$$

42

Author: *Analysis of the Lean Startup*
Article submitted to *Marketing Science*; manuscript no. MKSC-18-0058.R4

$$\mathbb{E}\min\{\tilde{r}, 1-\tilde{r}\} = \xi\left(1 - \frac{C}{2\epsilon}\right)\min\{r, 1-r\} + (1-\xi)\min\left\{r, 1-(1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon}\right) - r\right\}.$$

$\underline{(\lambda, v-\rho) \in (2C)}$

$$P(s_i < 0) = r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon}\right) + (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(C-\lambda+\epsilon)}{2\epsilon}\right)$$

$$= \left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon}\right) + (1-r)\frac{C}{2\epsilon}, \quad \tilde{r} = \frac{r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon}\right)}{r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon}\right) + (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(C-\lambda+\epsilon)}{2\epsilon}\right)};$$

(i) with probability ξ, surplus information is given if $s_i > 0$,

$$P(s_i > 0, 2n3) = \frac{v-\rho}{2\epsilon t} - \frac{(C-\epsilon-\lambda)}{2\epsilon}, \quad \tilde{r} = r,$$

$$P(s_i > 0, 3n4) = r\left(\frac{C-2\lambda-2\epsilon}{2\epsilon}\right), \quad \tilde{r} = 1,$$

$$P(s_i > 0, gt4) = r\left(\frac{\lambda+\epsilon}{\epsilon}\right), \quad \tilde{r} = 1;$$

(ii) with probability $1-\xi$, surplus information is NOT given if $s_i > 0$,

$$P(s_i > 0) = 1 - \left[r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon}\right) + (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(C-\lambda+\epsilon)}{2\epsilon}\right)\right],$$

$$\tilde{r} = \frac{r\left[1 - \left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon}\right)\right]}{1 - \left[r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon}\right) + (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(C-\lambda+\epsilon)}{2\epsilon}\right)\right]}.$$

$$\mathbb{E}\min\{\tilde{r}, 1-\tilde{r}\}$$

$$= \xi\left[\min\left\{r\left(-\frac{v-\rho}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon}\right), (1-r)\left(-\frac{v-\rho}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon}\right)\right\} + \left(\frac{v-\rho}{2\epsilon t} - \frac{C-\epsilon-\lambda}{2\epsilon}\right)\min\{r, 1-r\}\right]$$

$$+ (1-\xi)\left[\min\left\{r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon}\right), (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(C-\lambda+\epsilon)}{2\epsilon}\right)\right\} + \right.$$

$$\left.\min\left\{r\left[1 - \left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon}\right)\right], (1-r)\left[1 - \left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(C-\lambda+\epsilon)}{2\epsilon}\right)\right]\right\}\right].$$

$\underline{(\lambda, v-\rho) \in (2D)}$

$$P(s_i < 0) = r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon}\right) + (1-r), \quad \tilde{r} = \frac{r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon}\right)}{r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon}\right) + (1-r)};$$

(i) with probability ξ, surplus information is given if $s_i > 0$,

$$P(s_i > 0, 3n4) = r\left(\frac{v-\rho}{2\epsilon t} - \frac{\epsilon+\lambda}{2\epsilon}\right), \quad \tilde{r} = 1,$$

$$P(s_i > 0, gt4) = r\left(\frac{\epsilon+\lambda}{\epsilon}\right), \quad \tilde{r} = 1;$$

(ii) with probability $1-\xi$, surplus information is NOT given if $s_i > 0$,

$$P(s_i > 0) = 1 - \left[r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon}\right) + (1-r)\right], \quad \tilde{r} = 1.$$

$$\mathbb{E}\min\{\tilde{r}, 1-\tilde{r}\} = \min\left\{r\left(-\frac{v-\rho}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon}\right), 1-r\right\}.$$

$\underline{(\lambda, v-\rho) \in (2E)}$

$$P(s_i < 0) = r\left(1 - \frac{(v-\rho)}{\epsilon t}\right) + (1-r), \quad \tilde{r} = \frac{r\left(1 - \frac{(v-\rho)}{\epsilon t}\right)}{r\left(1 - \frac{(v-\rho)}{\epsilon t}\right) + (1-r)};$$

(i) with probability ξ, surplus information is given if $s_i > 0$,

Author: *Analysis of the Lean Startup*
Article submitted to *Marketing Science*; manuscript no. MKSC-18-0058.R4

43

$$P(s_i > 0, gt4) = r\left(\frac{v-\rho}{\epsilon t}\right), \quad \tilde{r} = 1;$$

(ii) with probability $1 - \xi$, surplus information is NOT given if $s_i > 0$,

$$P(s_i > 0) = 1 - \left[r\left(1 - \frac{(v-\rho)}{\epsilon t}\right) + (1-r)\right], \quad \tilde{r} = 1.$$

$$\mathbb{E}\min\{\tilde{r}, 1-\tilde{r}\} = \xi\min\left\{r\left(1 - \frac{v-\rho}{\epsilon t}\right), 1-r\right\} + (1-\xi)\min\left\{r\left(1 - \frac{(v-\rho)}{\epsilon t}\right), (1-r)\right\}.$$

$\underline{(\lambda, v - \rho) \in (3A)}$

$$P(s_i < 0) = 0;$$

(i) with probability ξ, surplus information is given if $s_i > 0$,

$$P(s_i > 0, 1n2) = (1-r)\frac{C}{2\epsilon}, \quad \tilde{r} = 0,$$

$$P(s_i > 0, 2n3) = r\left(-\frac{\lambda}{\epsilon}\right) + (1-r)\left(-\frac{\lambda}{\epsilon}\right) = -\frac{\lambda}{\epsilon}, \quad \tilde{r} = r,$$

$$P(s_i > 0, 3n4) = r\left(\frac{-C+2\lambda+2\epsilon}{\epsilon}\right) + (1-r)\left(\frac{-C+2\lambda+2\epsilon}{2\epsilon}\right), \quad \tilde{r} = \frac{r\left(\frac{-C+2\lambda+2\epsilon}{\epsilon}\right)}{r\left(\frac{-C+2\lambda+2\epsilon}{\epsilon}\right) + (1-r)\left(\frac{-C+2\lambda+2\epsilon}{2\epsilon}\right)},$$

$$P(s_i > 0, gt4) = r\left(\frac{C-\lambda-\epsilon}{\epsilon}\right), \quad \tilde{r} = 1;$$

(ii) with probability $1 - \xi$, surplus information is NOT given if $s_i > 0$,

$$P(s_i > 0) = 1, \quad \tilde{r} = r.$$

$$\mathbb{E}\min\{\tilde{r}, 1-\tilde{r}\} = \xi\left[\left(-\frac{\lambda}{\epsilon}\right)\min\{r, 1-r\} + \left(\frac{-C+2\epsilon+2\lambda}{2\epsilon}\right)\min\{2r, 1-r\}\right] + (1-\xi)\min\{r, 1-r\}.$$

$\underline{(\lambda, v - \rho) \in (3B)}$

$$P(s_i < 0) = (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon}\right), \quad \tilde{r} = 0;$$

(i) with probability ξ, surplus information is given if $s_i > 0$,

$$P(s_i > 0, 1n2) = (1-r)\left(\frac{v-\rho}{2\epsilon t} - \frac{(-\lambda+\epsilon)}{2\epsilon}\right), \quad \tilde{r} = 0,$$

$$P(s_i > 0, 2n3) = -\frac{\lambda}{\epsilon}, \quad \tilde{r} = r,$$

$$P(s_i > 0, 3n4) = r\left(\frac{-C+2\lambda+2\epsilon}{\epsilon}\right) + (1-r)\left(\frac{-C+2\lambda+2\epsilon}{2\epsilon}\right), \quad \tilde{r} = \frac{r\left(\frac{-C+2\lambda+2\epsilon}{\epsilon}\right)}{r\left(\frac{-C+2\lambda+2\epsilon}{\epsilon}\right) + (1-r)\left(\frac{-C+2\lambda+2\epsilon}{2\epsilon}\right)},$$

$$P(s_i > 0, gt4) = r\left(\frac{C-\lambda-\epsilon}{\epsilon}\right), \quad \tilde{r} = 1;$$

(ii) with probability $1 - \xi$, surplus information is NOT given if $s_i > 0$,

$$P(s_i > 0) = 1 - \left[(1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon}\right)\right], \quad \tilde{r} = \frac{r}{1 - \left[(1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon}\right)\right]}.$$

$$\mathbb{E}\min\{\tilde{r}, 1-\tilde{r}\} = \xi\left[\left(-\frac{\lambda}{\epsilon}\right)\min\{r, 1-r\} + \left(\frac{-C+2\epsilon+2\lambda}{2\epsilon}\right)\min\{2r, 1-r\}\right]$$

$$+ (1-\xi)\min\left\{r, 1 - \left[(1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon}\right)\right] - r\right\}.$$

$\underline{(\lambda, v - \rho) \in (3C)}$

$$P(s_i < 0) = r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon}\right) + (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(C-\lambda+\epsilon)}{2\epsilon}\right)$$

$$= \left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon}\right) + (1-r)\frac{C}{2\epsilon}, \quad \tilde{r} = \frac{r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon}\right)}{r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon}\right) + (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(C-\lambda+\epsilon)}{2\epsilon}\right)};$$

(i) with probability ξ, surplus information is given if $s_i > 0$,

$$P(s_i > 0, 2n3) = \frac{v-\rho}{2\epsilon t} - \frac{(\epsilon+\lambda)}{2\epsilon}, \quad \tilde{r} = r,$$

$$P(s_i > 0, 3n4) = r\left(\frac{-C+2\lambda+2\epsilon}{\epsilon}\right) + (1-r)\left(\frac{-C+2\lambda+2\epsilon}{2\epsilon}\right), \quad \tilde{r} = \frac{r\left(\frac{-C+2\lambda+2\epsilon}{\epsilon}\right)}{r\left(\frac{-C+2\lambda+2\epsilon}{\epsilon}\right) + (1-r)\left(\frac{-C+2\lambda+2\epsilon}{2\epsilon}\right)},$$

$$P(s_i > 0, gt4) = r\left(\frac{C-\lambda-\epsilon}{\epsilon}\right), \quad \tilde{r} = 1;$$

(ii) with probability $1-\xi$, surplus information is NOT given if $s_i > 0$,

$$P(s_i > 0) = 1 - \left[r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon}\right) + (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(C-\lambda+\epsilon)}{2\epsilon}\right)\right],$$

$$\tilde{r} = \frac{r\left[1 - \left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon}\right)\right]}{1 - \left[r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon}\right) + (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(C-\lambda+\epsilon)}{2\epsilon}\right)\right]}.$$

$$\mathbb{E}\min\{\tilde{r}, 1-\tilde{r}\} = \xi\left[\min\left\{r\left(-\frac{v-\rho}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon}\right), (1-r)\left(-\frac{v-\rho}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon}\right)\right\}\right.$$

$$\left. + \left(\frac{v-\rho}{2\epsilon t} - \frac{\lambda+\epsilon}{2\epsilon}\right)\min\{r, 1-r\} + \left(\frac{-C+2\epsilon+2\lambda}{2\epsilon}\right)\min\{2r, 1-r\}\right]$$

$$+ (1-\xi)\left[\min\left\{r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon}\right), (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(C-\lambda+\epsilon)}{2\epsilon}\right)\right\} + \right.$$

$$\left. \min\left\{r\left[1 - \left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon}\right)\right], (1-r)\left[1 - \left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(C-\lambda+\epsilon)}{2\epsilon}\right)\right]\right\}\right].$$

$\underline{(\lambda, v-\rho) \in (3D)}$

$$P(s_i < 0) = r\left(1 - \frac{(v-\rho)}{\epsilon t}\right) + (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon}\right), \quad \tilde{r} = \frac{r\left(1 - \frac{(v-\rho)}{\epsilon t}\right)}{r\left(1 - \frac{(v-\rho)}{\epsilon t}\right) + (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon}\right)};$$

(i) with probability ξ, surplus information is given if $s_i > 0$,

$$P(s_i > 0, 3n4) = r\left(\frac{v-\rho}{\epsilon t} - \frac{(C-\epsilon-\lambda)}{\epsilon}\right) + (1-r)\left(\frac{v-\rho}{2\epsilon t} - \frac{(C-\epsilon-\lambda)}{2\epsilon}\right),$$

$$\tilde{r} = \frac{r\left(\frac{v-\rho}{\epsilon t} - \frac{(C-\epsilon-\lambda)}{\epsilon}\right)}{r\left(\frac{v-\rho}{\epsilon t} - \frac{(C-\epsilon-\lambda)}{\epsilon}\right) + (1-r)\left(\frac{v-\rho}{2\epsilon t} - \frac{(C-\epsilon-\lambda)}{2\epsilon}\right)},$$

$$P(s_i > 0, gt4) = r\left(\frac{C-\epsilon-\lambda}{\epsilon}\right), \quad \tilde{r} = 1;$$

(ii) with probability $1-\xi$, surplus information is NOT given if $s_i > 0$,

$$P(s_i > 0) = 1 - \left[r\left(1 - \frac{(v-\rho)}{\epsilon t}\right) + (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon}\right)\right],$$

$$\tilde{r} = \frac{r\left[1 - \left(1 - \frac{(v-\rho)}{\epsilon t}\right)\right]}{1 - \left[r\left(1 - \frac{(v-\rho)}{\epsilon t}\right) + (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon}\right)\right]}.$$

$\mathbb{E}\min\{\tilde{r}, 1-\tilde{r}\}$

$$= \xi\left[\min\left\{r\left(1 - \frac{v-\rho}{\epsilon t}\right), (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon}\right)\right\} + \left(\frac{v-\rho}{2\epsilon t} - \frac{(C-\epsilon-\lambda)}{2\epsilon}\right)\min\{2r, 1-r\}\right]$$

$$+ (1-\xi)\left[\min\left\{r\left(1 - \frac{(v-\rho)}{\epsilon t}\right), (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon}\right)\right\} + \right.$$

Author: *Analysis of the Lean Startup*
Article submitted to *Marketing Science*; manuscript no. MKSC-18-0058.R4

45

$$\min\left\{r\left[1-\left(1-\frac{(v-\rho)}{\epsilon t}\right)\right], (1-r)\left[1-\left(-\frac{(v-\rho)}{2\epsilon t}+\frac{C-\lambda+\epsilon}{2\epsilon}\right)\right]\right\}\right].$$

$(\lambda, v-\rho) \in (3E)$

$$P(s_i < 0) = r\left(1-\frac{(v-\rho)}{\epsilon t}\right)+(1-r), \quad \tilde{r}=\frac{r\left(1-\frac{(v-\rho)}{\epsilon t}\right)}{r\left(1-\frac{(v-\rho)}{\epsilon t}\right)+(1-r)};$$

(i) with probability ξ, surplus information is given if $s_i > 0$,

$$P(s_i > 0, gt4) = r\left(\frac{v-\rho}{\epsilon t}\right), \quad \tilde{r}=1;$$

(ii) with probability $1-\xi$, surplus information is NOT given if $s_i > 0$,

$$P(s_i > 0) = 1-\left[r\left(1-\frac{(v-\rho)}{\epsilon t}\right)+(1-r)\right], \quad \tilde{r}=1.$$

$$\mathbb{E}\min\{\tilde{r}, 1-\tilde{r}\} = \xi\min\left\{r\left(1-\frac{v-\rho}{\epsilon t}\right), 1-r\right\}+(1-\xi)\min\left\{r\left(1-\frac{(v-\rho)}{\epsilon t}\right), (1-r)\right\}.$$

$(\lambda, v-\rho) \in (4A)$

$$P(s_i < 0) = 0;$$

(i) with probability ξ, surplus information is given if $s_i > 0$,

$$P(s_i > 0, 1n2) = (1-r)\left(\frac{C-2\lambda}{2\epsilon}\right), \quad \tilde{r}=0,$$

$$P(s_i > 0, 2n3) = \frac{\lambda}{\epsilon}, \quad \tilde{r}=r,$$

$$P(s_i > 0, 3n4) = r\left(\frac{-C+2\epsilon}{\epsilon}\right)+(1-r)\left(\frac{-C+2\epsilon}{2\epsilon}\right), \quad \tilde{r}=\frac{r\left(\frac{-C+2\epsilon}{\epsilon}\right)}{r\left(\frac{-C+2\epsilon}{\epsilon}\right)+(1-r)\left(\frac{-C+2\epsilon}{2\epsilon}\right)},$$

$$P(s_i > 0, gt4) = r\left(\frac{C-\lambda-\epsilon}{\epsilon}\right), \quad \tilde{r}=1;$$

(ii) with probability $1-\xi$, surplus information is NOT given if $s_i > 0$,

$$P(s_i > 0) = 1, \quad \tilde{r}=r.$$

$$\mathbb{E}\min\{\tilde{r}, 1-\tilde{r}\} = \xi\left[\left(\frac{\lambda}{\epsilon}\right)\min\{r, 1-r\}+\left(\frac{-C+2\epsilon}{2\epsilon}\right)\min\{2r, 1-r\}\right]+(1-\xi)\min\{r, 1-r\}.$$

$(\lambda, v-\rho) \in (4B)$

$$P(s_i < 0) = (1-r)\left(-\frac{(v-\rho)}{2\epsilon t}+\frac{C-\lambda+\epsilon}{2\epsilon}\right), \quad \tilde{r}=0;$$

(i) with probability ξ, surplus information is given if $s_i > 0$,

$$P(s_i > 0, 1n2) = (1-r)\left(\frac{v-\rho}{2\epsilon t}-\frac{(\lambda+\epsilon)}{2\epsilon}\right), \quad \tilde{r}=0,$$

$$P(s_i > 0, 2n3) = \frac{\lambda}{\epsilon}, \quad \tilde{r}=r,$$

$$P(s_i > 0, 3n4) = r\left(\frac{-C+2\epsilon}{\epsilon}\right)+(1-r)\left(\frac{-C+2\epsilon}{2\epsilon}\right), \quad \tilde{r}=\frac{r\left(\frac{-C+2\epsilon}{\epsilon}\right)}{r\left(\frac{-C+2\epsilon}{\epsilon}\right)+(1-r)\left(\frac{-C+2\epsilon}{2\epsilon}\right)},$$

$$P(s_i > 0, gt4) = r\left(\frac{C-\lambda-\epsilon}{\epsilon}\right), \quad \tilde{r}=1;$$

(ii) with probability $1-\xi$, surplus information is NOT given if $s_i > 0$,

$$P(s_i > 0) = 1-(1-r)\left(-\frac{(v-\rho)}{2\epsilon t}+\frac{C-\lambda+\epsilon}{2\epsilon}\right) \quad \tilde{r}=\frac{r}{1-(1-r)\left(-\frac{(v-\rho)}{2\epsilon t}+\frac{C-\lambda+\epsilon}{2\epsilon}\right)}.$$

46

Author: *Analysis of the Lean Startup*
Article submitted to *Marketing Science*; manuscript no. MKSC-18-0058.R4

$$\mathbb{E}\min\{\tilde{r},\, 1-\tilde{r}\} = \xi\left[\left(\frac{\lambda}{\epsilon}\right)\min\{r, 1-r\} + \left(\frac{-C+2\epsilon}{2\epsilon}\right)\min\{2r, 1-r\}\right]$$

$$+(1-\xi)\min\left\{r, 1-(1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon}\right) - r\right\}.$$

$\underline{(\lambda, v-\rho) \in (4C)}$

$$P(s_i < 0) = r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(\lambda+\epsilon)}{2\epsilon}\right) + (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(C-\lambda+\epsilon)}{2\epsilon}\right),$$

$$\tilde{r} = \frac{r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(\lambda+\epsilon)}{2\epsilon}\right)}{r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(\lambda+\epsilon)}{2\epsilon}\right) + (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(C-\lambda+\epsilon)}{2\epsilon}\right)};$$

(i) with probability ξ, surplus information is given if $s_i > 0$,

$$P(s_i > 0, 2n3) = \frac{v-\rho}{2\epsilon t} - \frac{(-\lambda+\epsilon)}{2\epsilon}, \quad \tilde{r} = r,$$

$$P(s_i > 0, 3n4) = r\left(\frac{-C+2\epsilon}{\epsilon}\right) + (1-r)\left(\frac{-C+2\epsilon}{2\epsilon}\right), \quad \tilde{r} = \frac{r\left(\frac{-C+2\epsilon}{\epsilon}\right)}{r\left(\frac{-C+2\epsilon}{\epsilon}\right) + (1-r)\left(\frac{-C+2\epsilon}{2\epsilon}\right)},$$

$$P(s_i > 0, gt4) = r\left(\frac{C-\lambda-\epsilon}{\epsilon}\right), \quad \tilde{r} = 1;$$

(ii) with probability $1-\xi$, surplus information is NOT given if $s_i > 0$,

$$P(s_i > 0) = 1 - \left[r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(\lambda+\epsilon)}{2\epsilon}\right) + (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(C-\lambda+\epsilon)}{2\epsilon}\right)\right],$$

$$\tilde{r} = \frac{r\left[1 - \left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(\lambda+\epsilon)}{2\epsilon}\right)\right]}{1 - \left[r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(\lambda+\epsilon)}{2\epsilon}\right) + (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(C-\lambda+\epsilon)}{2\epsilon}\right)\right]}.$$

$$\mathbb{E}\min\{\tilde{r},\, 1-\tilde{r}\} = \xi\left[\min\left\{r\left(-\frac{v-\rho}{2\epsilon t} + \frac{\lambda+\epsilon}{2\epsilon}\right), (1-r)\left(-\frac{v-\rho}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon}\right)\right\}\right.$$

$$+ \left(\frac{v-\rho}{2\epsilon t} - \frac{(-\lambda+\epsilon)}{2\epsilon}\right)\min\{r, 1-r\} + \left(\frac{-C+2\epsilon}{2\epsilon}\right)\min\{2r, 1-r\}\right]$$

$$+(1-\xi)\left[\min\left\{r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(\lambda+\epsilon)}{2\epsilon}\right), (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(C-\lambda+\epsilon)}{2\epsilon}\right)\right\} + \right.$$

$$\left.\min\left\{r\left[1 - \left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(\lambda+\epsilon)}{2\epsilon}\right)\right], (1-r)\left[1 - \left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(C-\lambda+\epsilon)}{2\epsilon}\right)\right]\right\}\right].$$

$\underline{(\lambda, v-\rho) \in (4D)}$

$$P(s_i < 0) = r\left(1 - \frac{(v-\rho)}{\epsilon t}\right) + (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon}\right), \quad \tilde{r} = \frac{r\left(1 - \frac{(v-\rho)}{\epsilon t}\right)}{r\left(1 - \frac{(v-\rho)}{\epsilon t}\right) + (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon}\right)};$$

(i) with probability ξ, surplus information is given if $s_i > 0$,

$$P(s_i > 0, 3n4) = r\left(\frac{v-\rho}{\epsilon t} - \frac{(C-\epsilon-\lambda)}{\epsilon}\right) + (1-r)\left(\frac{v-\rho}{2\epsilon t} - \frac{(C-\epsilon-\lambda)}{2\epsilon}\right),$$

$$\tilde{r} = \frac{r\left(\frac{v-\rho}{\epsilon t} - \frac{(C-\epsilon-\lambda)}{\epsilon}\right)}{r\left(\frac{v-\rho}{\epsilon t} - \frac{(C-\epsilon-\lambda)}{\epsilon}\right) + (1-r)\left(\frac{v-\rho}{2\epsilon t} - \frac{(C-\epsilon-\lambda)}{2\epsilon}\right)},$$

$$P(s_i > 0, gt4) = r\left(\frac{C-\epsilon-\lambda}{\epsilon}\right), \quad \tilde{r} = 1;$$

(ii) with probability $1-\xi$, surplus information is NOT given if $s_i > 0$,

$$P(s_i > 0) = 1 - \left[r\left(1 - \frac{(v-\rho)}{\epsilon t}\right) + (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon}\right)\right],$$

Author: *Analysis of the Lean Startup*
Article submitted to *Marketing Science*; manuscript no. MKSC-18-0058.R4

47

$$\tilde{r} = \frac{r\left[1 - \left(1 - \frac{(v-\rho)}{\epsilon t}\right)\right]}{1 - \left[r\left(1 - \frac{(v-\rho)}{\epsilon t}\right) + (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon}\right)\right]}.$$

$$\mathbb{E}\min\{\tilde{r},\, 1-\tilde{r}\}$$

$$= \xi\left[\min\left\{r\left(1 - \frac{v-\rho}{\epsilon t}\right),\, (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon}\right)\right\} + \left(\frac{v-\rho}{2\epsilon t} - \frac{(C-\epsilon-\lambda)}{2\epsilon}\right)\min\{2r, 1-r\}\right]$$

$$+ (1-\xi)\left[\min\left\{r\left(1 - \frac{(v-\rho)}{\epsilon t}\right),\, (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon}\right)\right\} + \right.$$

$$\left. \min\left\{r\left[1 - \left(1 - \frac{(v-\rho)}{\epsilon t}\right)\right],\, (1-r)\left[1 - \left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon}\right)\right]\right\}\right].$$

$\underline{(\lambda, v-\rho) \in (4E)}$

$$P(s_i < 0) = r\left(1 - \frac{(v-\rho)}{\epsilon t}\right) + (1-r), \quad \tilde{r} = \frac{r\left(1 - \frac{(v-\rho)}{\epsilon t}\right)}{r\left(1 - \frac{(v-\rho)}{\epsilon t}\right) + (1-r)};$$

(i) with probability ξ, surplus information is given if $s_i > 0$,

$$P(s_i > 0, gt4) = r\left(\frac{v-\rho}{\epsilon t}\right), \quad \tilde{r} = 1;$$

(ii) with probability $1 - \xi$, surplus information is NOT given if $s_i > 0$,

$$P(s_i > 0) = 1 - \left[r\left(1 - \frac{(v-\rho)}{\epsilon t}\right) + (1-r)\right], \quad \tilde{r} = 1.$$

$$\mathbb{E}\min\{\tilde{r},\, 1-\tilde{r}\} = \min\left\{r\left(1 - \frac{v-\rho}{\epsilon t}\right), 1-r\right\}.$$

$\underline{(\lambda, v-\rho) \in (5A)}$

$$P(s_i < 0) = 0;$$

(i) with probability ξ, surplus information is given if $s_i > 0$,

$$P(s_i > 0, 1n2) = (1-r)\left(\frac{C-2\lambda}{2\epsilon}\right), \quad \tilde{r} = 0,$$

$$P(s_i > 0, 2n3) = \frac{\lambda}{\epsilon}, \quad \tilde{r} = r,$$

$$P(s_i > 0, 3n4) = r\left(\frac{C-2\lambda}{\epsilon}\right) + (1-r)\left(\frac{C-2\lambda}{2\epsilon}\right), \quad \tilde{r} = \frac{r\left(\frac{C-2\lambda}{\epsilon}\right)}{r\left(\frac{C-2\lambda}{\epsilon}\right) + (1-r)\left(\frac{C-2\lambda}{2\epsilon}\right)},$$

$$P(s_i > 0, gt4) = \frac{\lambda - C + \epsilon}{\epsilon}, \quad \tilde{r} = r;$$

(ii) with probability $1 - \xi$, surplus information is NOT given if $s_i > 0$,

$$P(s_i > 0) = 1, \quad \tilde{r} = r.$$

$$\mathbb{E}\min\{\tilde{r},\, 1-\tilde{r}\} = \xi\left[\left(\frac{\lambda}{\epsilon}\right)\min\{r, 1-r\} + \left(\frac{C-2\lambda}{2\epsilon}\right)\min\{2r, 1-r\} + \left(\frac{\lambda - C + \epsilon}{\epsilon}\right)\min\{r, 1-r\}\right]$$

$$+ (1-\xi)\min\{r, 1-r\}.$$

$\underline{(\lambda, v-\rho) \in (5B)}$

$$P(s_i < 0) = (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon}\right), \quad \tilde{r} = 0;$$

(i) with probability ξ, surplus information is given if $s_i > 0$,

$$P(s_i > 0, 1n2) = (1-r)\left(\frac{v-\rho}{2\epsilon t} - \frac{(\lambda+\epsilon)}{2\epsilon}\right), \quad \tilde{r} = 0,$$

48

Author: *Analysis of the Lean Startup*
Article submitted to *Marketing Science*; manuscript no. MKSC-18-0058.R4

$$P(s_i > 0, 2n3) = \frac{\lambda}{\epsilon}, \quad \tilde{r} = r,$$

$$P(s_i > 0, 3n4) = r\left(\frac{C - 2\lambda}{\epsilon}\right) + (1 - r)\left(\frac{C - 2\lambda}{2\epsilon}\right), \quad \tilde{r} = \frac{r\left(\frac{C - 2\lambda}{\epsilon}\right)}{r\left(\frac{C - 2\lambda}{\epsilon}\right) + (1 - r)\left(\frac{C - 2\lambda}{2\epsilon}\right)},$$

$$P(s_i > 0, gt4) = \frac{\lambda - C + \epsilon}{\epsilon}, \quad \tilde{r} = r;$$

(ii) with probability $1 - \xi$, surplus information is NOT given if $s_i > 0$,

$$P(s_i > 0) = 1 - (1 - r)\left(-\frac{(v - \rho)}{2\epsilon t} + \frac{C - \lambda + \epsilon}{2\epsilon}\right), \quad \tilde{r} = \frac{r}{1 - (1 - r)\left(-\frac{(v - \rho)}{2\epsilon t} + \frac{C - \lambda + \epsilon}{2\epsilon}\right)}.$$

$$\mathbb{E}\min\{\tilde{r}, 1 - \tilde{r}\} = \xi\left[\left(\frac{\lambda}{\epsilon}\right)\min\{r, 1 - r\} + \left(\frac{C - 2\lambda}{2\epsilon}\right)\min\{2r, 1 - r\} + \left(\frac{\lambda - C + \epsilon}{\epsilon}\right)\min\{r, 1 - r\}\right]$$

$$+ (1 - \xi)\min\left\{r, 1 - (1 - r)\left(-\frac{(v - \rho)}{2\epsilon t} + \frac{C - \lambda + \epsilon}{2\epsilon}\right) - r\right\}.$$

$\underline{(\lambda, v - \rho) \in (5C)}$

$$P(s_i < 0) = r\left(-\frac{(v - \rho)}{2\epsilon t} + \frac{(\lambda + \epsilon)}{2\epsilon}\right) + (1 - r)\left(-\frac{(v - \rho)}{2\epsilon t} + \frac{(C - \lambda + \epsilon)}{2\epsilon}\right),$$

$$\tilde{r} = \frac{r\left(-\frac{(v - \rho)}{2\epsilon t} + \frac{(\lambda + \epsilon)}{2\epsilon}\right)}{r\left(-\frac{(v - \rho)}{2\epsilon t} + \frac{(\lambda + \epsilon)}{2\epsilon}\right) + (1 - r)\left(-\frac{(v - \rho)}{2\epsilon t} + \frac{(C - \lambda + \epsilon)}{2\epsilon}\right)};$$

(i) with probability ξ, surplus information is given if $s_i > 0$,

$$P(s_i > 0, 2n3) = \frac{v - \rho}{2\epsilon t} - \frac{(-\lambda + \epsilon)}{2\epsilon}, \quad \tilde{r} = r,$$

$$P(s_i > 0, 3n4) = r\left(\frac{C - 2\lambda}{\epsilon}\right) + (1 - r)\left(\frac{C - 2\lambda}{2\epsilon}\right), \quad \tilde{r} = \frac{r\left(\frac{C - 2\lambda}{\epsilon}\right)}{r\left(\frac{C - 2\lambda}{\epsilon}\right) + (1 - r)\left(\frac{C - 2\lambda}{2\epsilon}\right)},$$

$$P(s_i > 0, gt4) = \frac{\lambda - C + \epsilon}{\epsilon}, \quad \tilde{r} = r;$$

(ii) with probability $1 - \xi$, surplus information is NOT given if $s_i > 0$,

$$P(s_i > 0) = 1 - \left[r\left(-\frac{(v - \rho)}{2\epsilon t} + \frac{(\lambda + \epsilon)}{2\epsilon}\right) + (1 - r)\left(-\frac{(v - \rho)}{2\epsilon t} + \frac{(C - \lambda + \epsilon)}{2\epsilon}\right)\right],$$

$$\tilde{r} = \frac{r\left[1 - \left(-\frac{(v - \rho)}{2\epsilon t} + \frac{(\lambda + \epsilon)}{2\epsilon}\right)\right]}{1 - \left[r\left(-\frac{(v - \rho)}{2\epsilon t} + \frac{(\lambda + \epsilon)}{2\epsilon}\right) + (1 - r)\left(-\frac{(v - \rho)}{2\epsilon t} + \frac{(C - \lambda + \epsilon)}{2\epsilon}\right)\right]}.$$

$$\mathbb{E}\min\{\tilde{r}, 1 - \tilde{r}\} = \xi\left[\min\left\{r\left(-\frac{v - \rho}{2\epsilon t} + \frac{\lambda + \epsilon}{2\epsilon}\right), (1 - r)\left(-\frac{v - \rho}{2\epsilon t} + \frac{C - \lambda + \epsilon}{2\epsilon}\right)\right\}\right.$$

$$\left. + \left(\frac{v - \rho}{2\epsilon t} - \frac{(-\lambda + \epsilon)}{2\epsilon}\right)\min\{r, 1 - r\} + \left(\frac{C - 2\lambda}{2\epsilon}\right)\min\{2r, 1 - r\} + \left(\frac{\lambda - C + \epsilon}{\epsilon}\right)\min\{r, 1 - r\}\right]$$

$$+ (1 - \xi)\left[\min\left\{r\left(-\frac{(v - \rho)}{2\epsilon t} + \frac{(\lambda + \epsilon)}{2\epsilon}\right), (1 - r)\left(-\frac{(v - \rho)}{2\epsilon t} + \frac{(C - \lambda + \epsilon)}{2\epsilon}\right)\right\} + \right.$$

$$\left.\min\left\{r\left[1 - \left(-\frac{(v - \rho)}{2\epsilon t} + \frac{(\lambda + \epsilon)}{2\epsilon}\right)\right], (1 - r)\left[1 - \left(-\frac{(v - \rho)}{2\epsilon t} + \frac{(C - \lambda + \epsilon)}{2\epsilon}\right)\right]\right\}\right].$$

$\underline{(\lambda, v - \rho) \in (5D)}$

$$P(s_i < 0) = r\left(1 - \frac{(v - \rho)}{\epsilon t}\right) + (1 - r)\left(-\frac{(v - \rho)}{2\epsilon t} + \frac{C - \lambda + \epsilon}{2\epsilon}\right), \quad \tilde{r} = \frac{r\left(1 - \frac{(v - \rho)}{\epsilon t}\right)}{r\left(1 - \frac{(v - \rho)}{\epsilon t}\right) + (1 - r)\left(-\frac{(v - \rho)}{2\epsilon t} + \frac{C - \lambda + \epsilon}{2\epsilon}\right)};$$

(i) with probability ξ, surplus information is given if $s_i > 0$,

$$P(s_i > 0, 3n4) = r\left(\frac{v - \rho}{\epsilon t} - \frac{(\lambda - C + \epsilon)}{\epsilon}\right) + (1 - r)\left(\frac{v - \rho}{2\epsilon t} - \frac{(\lambda - C + \epsilon)}{2\epsilon}\right),$$

Author: *Analysis of the Lean Startup*
Article submitted to *Marketing Science*; manuscript no. MKSC-18-0058.R4

49

$$\tilde{r} = \frac{r\left(\frac{v-\rho}{\epsilon t} - \frac{(\lambda - C + \epsilon)}{\epsilon}\right)}{r\left(\frac{v-\rho}{\epsilon t} - \frac{(\lambda - C + \epsilon)}{\epsilon}\right) + (1-r)\left(\frac{v-\rho}{2\epsilon t} - \frac{(\lambda - C + \epsilon)}{2\epsilon}\right)},$$

$$P(s_i > 0, gt4) = \frac{\lambda - C + \epsilon}{\epsilon}, \quad \tilde{r} = r;$$

(ii) with probability $1 - \xi$, surplus information is NOT given if $s_i > 0$,

$$P(s_i > 0) = 1 - \left[r\left(1 - \frac{(v-\rho)}{\epsilon t}\right) + (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C - \lambda + \epsilon}{2\epsilon}\right)\right],$$

$$\tilde{r} = \frac{r\left[1 - \left(1 - \frac{(v-\rho)}{\epsilon t}\right)\right]}{1 - \left[r\left(1 - \frac{(v-\rho)}{\epsilon t}\right) + (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C - \lambda + \epsilon}{2\epsilon}\right)\right]}.$$

$$\mathbb{E}\min\{\tilde{r}, 1 - \tilde{r}\} = \xi\left[\min\left\{r\left(1 - \frac{v-\rho}{\epsilon t}\right), (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C - \lambda + \epsilon}{2\epsilon}\right)\right\}\right.$$

$$+ \left(\frac{v-\rho}{2\epsilon t} - \frac{(\lambda - C + \epsilon)}{2\epsilon}\right)\min\{2r, 1-r\} + \left(\frac{\lambda - C + \epsilon}{\epsilon}\right)\min\{r, 1-r\}\right]$$

$$+ (1-\xi)\left[\min\left\{r\left(1 - \frac{(v-\rho)}{\epsilon t}\right), (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C - \lambda + \epsilon}{2\epsilon}\right)\right\} + \right.$$

$$\min\left\{r\left[1 - \left(1 - \frac{(v-\rho)}{\epsilon t}\right)\right], (1-r)\left[1 - \left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C - \lambda + \epsilon}{2\epsilon}\right)\right]\right\}\right].$$

$\underline{(\lambda, v - \rho) \in (5E)}$

$$P(s_i < 0) = 1 - \frac{(v-\rho)}{\epsilon t}, \quad \tilde{r} = r;$$

(i) with probability ξ, surplus information is given if $s_i > 0$,

$$P(s_i > 0, gt4) = \frac{v-\rho}{\epsilon t}, \quad \tilde{r} = r;$$

(ii) with probability $1 - \xi$, surplus information is NOT given if $s_i > 0$,
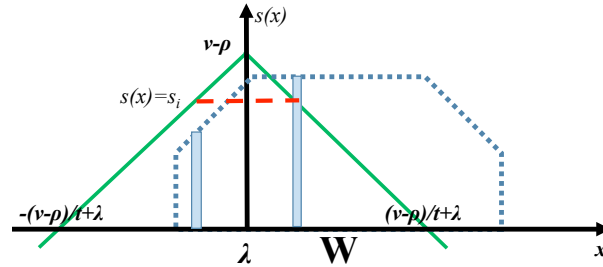
$$P(s_i > 0) = \frac{v-\rho}{\epsilon t}, \quad \tilde{r} = r.$$

$$\mathbb{E}\min\{\tilde{r}, 1 - \tilde{r}\} = \min\{r, 1-r\}.$$

Appendix C: Proofs

Proof of Lemma 1. The probability that a randomly selected customer i has surplus $[s_i, s_i + \delta]$ for some small δ is given in Figure C-1.

Figure C-1 **Illustration of probability that a randomly chosen consumer has surplus $[s_i, s_i + \delta]$ for some small δ.**

$$P(u_i \in (s_i, s_i + \delta)|(\lambda, v, \rho), W) = \int_{\frac{-v+\rho+s_i}{t} + \lambda}^{\frac{-v+\rho+s_i+\delta}{t} + \lambda} h(x|W)dx + \int_{\frac{v-\rho-s_i-\delta}{t} + \lambda}^{\frac{v-\rho-s_i}{t} + \lambda} h(x|W)dx$$

$$= h\left(\frac{-v+\rho+s_i}{t}+\lambda|W\right)\cdot\frac{\delta}{t}+h\left(\frac{v-\rho-s_i}{t}+\lambda|W\right)\cdot\frac{\delta}{t}$$

$$= \frac{1}{t}\left[h\left(\frac{-v+s_i}{t}+\lambda|W\right)+h\left(\frac{v-s_i}{t}+\lambda|W\right)\right]\cdot\delta \equiv g(s_i|(\lambda,v,\rho),W)\cdot\delta. \quad\square$$

Proof of Lemma A.1 (Lemma 2 is a special case). (i) Surplus setting: Recall that $P(\lambda, u_i \in [s_i, s_i + \delta]|W) = g(\lambda, s_i|W)\cdot\delta$. So,

$$P(W=0|\lambda, u_i = s_i) \cong P(W=0|\lambda, u_i \in [s_i, s_i+\delta])$$

$$= \frac{P(\lambda, u_i \in [s_i, s_i+\delta]|W=0)P(W=0)}{P(\lambda, u_i \in [s_i, s_i+\delta]|W=0)P(W=0) + P(\lambda, u_i \in [s_i, s_i+\delta]|W=C)P(W=C)}$$

$$= \frac{g(\lambda, s_i|W=0)\cdot\delta\cdot P(W=0)}{g(\lambda, s_i|W=0)\cdot\delta\cdot P(W=0) + g(\lambda, s_i|W=C)\cdot\delta\cdot P(W=C)}$$

$$= \frac{g(\lambda, s_i|W=0)P(W=0)}{g(\lambda, s_i|W=0)P(W=0) + g(\lambda, s_i|W=C)P(W=C)}.$$

(ii) sales and no surplus information:

$$P(W=0|1(s_i>0)) = \frac{P(1(s_i>0)|W=0)P(W=0)}{P(1(s_i>0)|W=0)P(W=0) + P(1(s_i>0)|W=C)P(W=C)};$$

(iii) No sales and no surplus information:

$$P(W=0|1(s_i<0)) = \frac{P(1(s_i<0)|W=0)P(W=0)}{P(1(s_i<0)|W=0)P(W=0) + P(1(s_i<0)|W=C)P(W=C)}. \quad\square$$

Proof of Proposition 3 (Proposition 1 is a special case). (i) We shall examine the case when $f < t$.

To avoid redundancy, we will show the result for the setting (a) $\lambda \in \left(C - \epsilon - \frac{V-p}{t}, \epsilon - \frac{V-p}{t}\right)$, and examine the derivative w.r.t. Λ in the following three cases. Setting (c) corresponding to $\lambda \in \left(C - \epsilon + \frac{V-p}{t}, \epsilon + \frac{V-p}{t}\right)$

Figure C-2   Three cases: (1) – left panel ; (2) – middle panel; (3) – right panel.
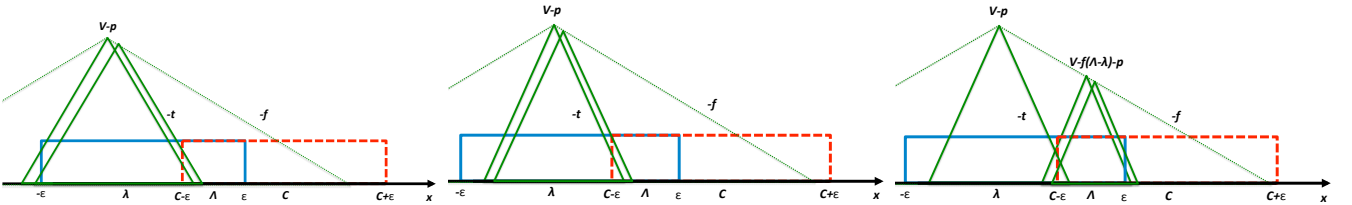
can be analyzed in a similar manner. Also, setting (b) follows from combining Case (3) for setting (a) and setting (b).

Case (1). In this case $\lambda - \frac{V-p}{\epsilon} < -\epsilon$ (represented in left panel of Figure C-2). Without any pivoting, the final product's location is λ, and with incremental pivoting the final product's location is $\Lambda > \lambda$. Then, the expected revenues are respectively

$$\pi_\lambda = \left(\frac{V-p}{t} + \lambda - (-\epsilon)\right)\cdot\tilde{r} + \left(\frac{V-p}{t} + \lambda - (C-\epsilon)\right)\cdot(1-\tilde{r});$$

$$\pi_\Lambda = \left(\frac{V-p-f(\Lambda-\lambda)}{t} + \Lambda - (-\epsilon)\right)\cdot\tilde{r} + \left(\frac{V-p-f(\Lambda-\lambda)}{t} + \Lambda - (C-\epsilon)\right)\cdot(1-\tilde{r}).$$

Author: *Analysis of the Lean Startup*
Article submitted to *Marketing Science*; manuscript no. MKSC-18-0058.R4

51

Then, we have

$$\pi_\Lambda - \pi_\lambda = \left(-\frac{f(\Lambda - \lambda)}{t} + \Lambda - \lambda\right) \cdot \tilde{r} + \left(-\frac{f(\Lambda - \lambda)}{t} + \Lambda - \lambda\right) \cdot (1 - \tilde{r})$$

$$= \left(1 - \frac{f}{t}\right)(\Lambda - \lambda)\tilde{r} + \left(1 - \frac{f}{t}\right)(\Lambda - \lambda)(1 - \tilde{r}) = \left(1 - \frac{f}{t}\right)(\Lambda - \lambda) > 0.$$

Hence, in this case, the pivot should happen until it borders Case (2), i.e., until $\Lambda^* - \frac{V - p - f(\Lambda^* - \lambda)}{t} = -\epsilon$.

Case (2). This case is represented by the middle panel of Figure C-2. Without any pivoting, the final product's location is λ, and with incremental pivoting the final product's location is $\Lambda > \lambda$. Then, the expected revenues are respectively

$$\pi_\lambda = \left(\frac{V - p}{t} + \lambda - \left(-\frac{V - p}{t} + \lambda\right)\right) \cdot \tilde{r} + \left(\frac{V - p}{t} + \lambda - (C - \epsilon)\right) \cdot (1 - \tilde{r});$$

$$\pi_\Lambda = \left(\frac{V - p - f(\Lambda - \lambda)}{t} + \Lambda - \left(-\frac{V - p - f(\Lambda - \lambda)}{t} + \Lambda\right)\right) \cdot \tilde{r}$$

$$+ \left(\frac{V - p - f(\Lambda - \lambda)}{t} + \Lambda - (C - \epsilon)\right) \cdot (1 - \tilde{r}).$$

Then, we have

$$\pi_\Lambda - \pi_\lambda = -\frac{2f(\Lambda - \lambda)}{t} \cdot \tilde{r} + \left(-\frac{f(\Lambda - \lambda)}{t} + \Lambda - \lambda\right) \cdot (1 - \tilde{r})$$

$$= -\frac{2f}{t}(\Lambda - \lambda)\tilde{r} + \left(1 - \frac{f}{t}\right)(\Lambda - \lambda)(1 - \tilde{r}) = \left(-\frac{f}{t}2\tilde{r} + \left(1 - \frac{f}{t}\right)(1 - \tilde{r})\right)(\Lambda - \lambda).$$

Hence, in this case, pivoting is beneficial if and only if

$$\frac{f}{t}2\tilde{r} + \left(1 - \frac{f}{t}\right)(1 - \tilde{r}) > 0 \iff \frac{f}{t} < \frac{1 - \tilde{r}}{1 + \tilde{r}}.$$

Thus, in this case, it is optimal not to pivot if $\frac{f}{t} > \frac{1 - \tilde{r}}{1 + \tilde{r}}$.

If $\frac{f}{t} < \frac{1 - \tilde{r}}{1 + \tilde{r}}$, then pivoting should happen until it borders Case (3), i.e., until $\Lambda^* + \frac{V - p - f(\Lambda^* - \lambda)}{t} = \epsilon$.

Case (3). This case is represented by the right panel of Figure C-2. Suppose that Λ_0 represents the position that first enters this case. We examine whether incrementally pivoting to $\Lambda_1 > \Lambda_0$ is beneficial or not by examining the derivative.

$$\pi_{\Lambda_0} = \left(C - \epsilon - \frac{V - p - f(\Lambda_0 - \lambda)}{t} + \Lambda_0\right) \cdot \tilde{r} + [\epsilon - (C - \epsilon)] + \left(\frac{V - p - f(\Lambda_0 - \lambda)}{t} + \Lambda_0\right) \cdot (1 - \tilde{r});$$

$$\pi_{\Lambda_1} = \left(C - \epsilon - \frac{V - p - f(\Lambda_1 - \lambda)}{t} + \Lambda_1\right) \cdot \tilde{r} + [\epsilon - (C - \epsilon)] + \left(\frac{V - p - f(\Lambda_1 - \lambda)}{t} + \Lambda_1\right) \cdot (1 - \tilde{r}).$$

Then, we have

$$\pi_{\Lambda_1} - \pi_{\Lambda_0} = \left(-\frac{f\Lambda_1}{t} - \Lambda_1 - \left(-\frac{f\Lambda_0}{t} - \Lambda_0\right)\right) \cdot \tilde{r} + \left(-\frac{f\Lambda_1}{t} + \Lambda_1 - \left(-\frac{f\Lambda_0}{t} + \Lambda_0\right)\right) \cdot (1 - \tilde{r})$$

$$= -\left(1 + \frac{f}{t}\right)(\Lambda_1 - \Lambda_0)\tilde{r} + \left(1 - \frac{f}{t}\right)(\Lambda_1 - \Lambda_0)(1 - \tilde{r})$$

$$= \left[-\left(1 + \frac{f}{t}\right)\tilde{r} + \left(1 - \frac{f}{t}\right)(1 - \tilde{r})\right](\Lambda_1 - \Lambda_0) = \left[1 - 2\tilde{r} - \frac{f}{t}\right](\Lambda_1 - \Lambda_0).$$

Hence, in this case, pivoting is beneficial if and only if

$$1 - 2\tilde{r} - \frac{f}{t} > 0 \iff \frac{f}{t} < 1 - 2\tilde{r} \iff \tilde{r} < \frac{1}{2}\left(1 - \frac{f}{t}\right).$$

Hence, pivoting should continue until it leaves this region, i.e., $\Lambda^* - \frac{V - p - f(\Lambda^* - \lambda)}{t} = C - \epsilon$.

(ii) If $f > t$, the coverage area does not shift as a result of pivoting. As such, in this case, pivoting is never optimal. $\square$

Proof of Lemma 3.   Let $s_i \equiv s_i(\lambda, v, \rho)$ and $\tilde{r} \equiv \tilde{r}(\lambda, v, \rho | r)$. Taking the expectation of (5), we have

$$
\begin{aligned}
\max_{(\lambda, v, \rho)} \mathbb{E}_{s_i} \pi^*(\tilde{r}) &= \max_{(\lambda, v, \rho)} \mathbb{E}_{s_i} p\left[\frac{V-p}{\epsilon t} - \left(\frac{V-p}{\epsilon t} - \left(1 - \frac{C}{2\epsilon} \right) \right) \min\{\tilde{r}, 1 - \tilde{r}\} \right] \\
&= \max_{(\lambda, v, \rho)} p\left[\frac{V-p}{\epsilon t} - \left(\frac{V-p}{\epsilon t} - \left(1 - \frac{C}{2\epsilon} \right) \right) \mathbb{E}_{s_i} \min\{\tilde{r}, 1 - \tilde{r}\} \right] \\
&= p\frac{V-p}{\epsilon t} - p\left(\frac{V-p}{\epsilon t} - \left(1 - \frac{C}{2\epsilon} \right) \right) \min_{(\lambda, v, \rho)} \mathbb{E}_{s_i} \min\{\tilde{r}, 1 - \tilde{r}\}. \quad \square
\end{aligned}
$$

Proof of Proposition A.1 (Proposition 2 is a special case).   We solve the problem of finding test product that minimizes $\mathbb{E}\min\{\tilde{r}, 1 - \tilde{r}\}$, whose expressions for the 25 regions of $(\lambda, v - \rho)$ are derived in Appendix B. Observe that within each region, the expression is *linear* in λ and $(v - \rho)$, i.e., it is a plane. This implies that the optimal test product that maximizes learning $(\lambda^*, (v - \rho)^*)$ occurs in the boundaries between the regions. To find out which boundaries, we examine the first derivatives with respect to λ and $(v - \rho)$ of each region.

(1A), (2A): We have $\mathbb{E}\min\{\tilde{r}, 1 - \tilde{r}\} = \left(1 - \xi\frac{C}{2\epsilon} \right) \min\{r, 1 - r\}$, so $\frac{\partial}{\partial(v-\rho)} = \frac{\partial}{\partial\lambda} = 0$.

(1B), (2B): We have

$$
\mathbb{E}\min\{\tilde{r}, 1 - \tilde{r}\} = \xi\left(1 - \frac{C}{2\epsilon} \right) \min\{r, 1 - r\} + (1 - \xi) \min\left\{ r, (1 - r)\left(1 + \frac{v-\rho}{2\epsilon t} - \frac{1}{2} - \frac{C-\lambda}{2\epsilon} \right) \right\}.
$$

Let $1 - A \equiv 1$, and $1 - B \equiv \left(1 + \frac{v-\rho}{2\epsilon t} - \frac{1}{2} - \frac{C-\lambda}{2\epsilon} \right)$.

  (a) If $r \leq \frac{1-B}{1+1-B}$, we have $\mathbb{E}\min\{\tilde{r}, 1 - \tilde{r}\} = \xi\left(1 - \frac{C}{2\epsilon} \right) \min\{r, 1 - r\} + (1 - \xi)r$, so $\frac{\partial}{\partial(v-\rho)} = \frac{\partial}{\partial\lambda} = 0$.

  (b) If $r > \frac{1-B}{1+1-B}$, $\mathbb{E}\min\{\tilde{r}, 1 - \tilde{r}\} = \xi\left(1 - \frac{C}{2\epsilon} \right) \min\{r, 1 - r\} + (1 - \xi)(1 - r)\left(1 + \frac{v-\rho}{2\epsilon t} - \frac{1}{2} - \frac{C-\lambda}{2\epsilon} \right)$, so $\frac{\partial}{\partial(v-\rho)} = \frac{(1-\xi)(1-r)}{2\epsilon t} > 0$, and $\frac{\partial}{\partial\lambda} = \frac{(1-\xi)(1-r)}{2\epsilon} > 0$.

(1C), (2C): We have

$$
\begin{aligned}
\mathbb{E}\min\{\tilde{r}, 1 - \tilde{r}\} = &\ \xi\left[\min\left\{ r\left(-\frac{v-\rho}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon} \right), (1-r)\frac{C}{2\epsilon} \right\} + \left(\frac{v-\rho}{2\epsilon t} - \frac{C-\epsilon-t}{2\epsilon} \right) \min\{r, 1 - r\}, \right] \\
&+ (1 - \xi)\left[\min\left\{ r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon} \right), (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(C-\lambda+\epsilon)}{2\epsilon} \right) \right\} + \right. \\
&\left. \min\left\{ r\left[1 - \left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(-\lambda+\epsilon)}{2\epsilon} \right) \right], (1-r)\left[1 - \left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(C-\lambda+\epsilon)}{2\epsilon} \right) \right] \right\} \right].
\end{aligned}
$$

Let $A \equiv r\left(-\frac{v-\rho}{2\epsilon t} + \frac{-\lambda+\epsilon}{2\epsilon} \right)$, $B \equiv \left(-\frac{v-\rho}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon} \right)$. We have $A < B$, and thus we have three cases. We have four ranges of priors.

  (a) If $r \leq \frac{1-B}{1-A+1-B}$, we have $\frac{\partial}{\partial(v-\rho)} = \frac{\partial}{\partial\lambda} = 0$.

  (b) If $\frac{1-B}{1-A+1-B} < r \leq \frac{1}{2}$, then we have $\frac{\partial}{\partial(v-\rho)} > 0$ and $\frac{\partial}{\partial\lambda} > 0$.

  (c) If $\frac{1}{2} < r \leq \frac{B}{A+B}$, then we have $\frac{\partial}{\partial(v-\rho)} < 0$, and $\frac{\partial}{\partial\lambda} < 0$.

  (d) If $r > \frac{B}{A+B}$, we have $\frac{\partial}{\partial(v-\rho)} = \frac{\partial}{\partial\lambda} = 0$.

(1D): We have $\mathbb{E}\min\{\tilde{r}, 1 - \tilde{r}\} = \min\left\{ r\left(-\frac{v-\rho}{2\epsilon t} + \frac{-\lambda+\epsilon}{2\epsilon} \right), (1 - r) \right\}$. Let $A \equiv \left(-\frac{v-\rho}{2\epsilon t} + \frac{-\lambda+\epsilon}{2\epsilon} \right)$, and $B \equiv 1$, so $A < B$.

  (a) If $r \leq \frac{1}{A+1}$, we have $\frac{\partial}{\partial(v-\rho)} < 0$, and $\frac{\partial}{\partial\lambda} < 0$.

  (b) If $r > \frac{1}{A+1}$, then $\frac{\partial}{\partial(v-\rho)} = \frac{\partial}{\partial\lambda} = 0$.

(1E), (5E): We have $\mathbb{E}\min\{\tilde{r}, 1 - \tilde{r}\} = \min\{r, 1 - r\}$. Thus, $\frac{\partial}{\partial(v-\rho)} = \frac{\partial}{\partial\lambda} = 0$.

Author: *Analysis of the Lean Startup*
Article submitted to *Marketing Science*; manuscript no. MKSC-18-0058.R4

53

(2E), (3E), (4E): We have $\mathbb{E}\min\{\tilde{r},\, 1-\tilde{r}\} = \min\left\{r\left(1-\frac{v-\rho}{\epsilon t}\right),\ (1-r)\right\}$. Let $A \equiv \left(1-\frac{v-\rho}{\epsilon t}\right)$, and $B \equiv 1$, so $A < B$.

(a) If $r \leq \frac{1}{A+1}$, we have $\frac{\partial}{\partial(v-\rho)} < 0$, and $\frac{\partial}{\partial\lambda} = 0$.

(b) If $r > \frac{1}{A+1}$, we have $\frac{\partial}{\partial(v-\rho)} = \frac{\partial}{\partial\lambda} = 0$.

(3A): We have

$$\mathbb{E}\min\{\tilde{r},\, 1-\tilde{r}\} = \xi\left[\left(-\frac{\lambda}{\epsilon}\right)\min\{r,1-r\} + \left(\frac{-C+2\epsilon+2\lambda}{2\epsilon}\right)\min\{2r,1-r\}\right] + (1-\xi)\min\{r,1-r\}.$$

So we have,

(a) if $r < \frac{1}{3}$, then $\frac{\partial}{\partial(v-\rho)} = 0$, and $\frac{\partial}{\partial\lambda} > 0$;

(b) if $\frac{1}{3} < r \leq \frac{1}{2}$, then $\frac{\partial}{\partial(v-\rho)} = 0$; and $\frac{\partial}{\partial\lambda} \geq 0$.

(c) if $r > \frac{1}{2}$, then $\frac{\partial}{\partial(v-\rho)} = \frac{\partial}{\partial\lambda} = 0$.

(3B): We have,

$$\mathbb{E}\min\{\tilde{r},\, 1-\tilde{r}\} = \xi\left[\left(-\frac{\lambda}{\epsilon}\right)\min\{r,1-r\} + \left(\frac{-C+2\epsilon+2\lambda}{2\epsilon}\right)\min\{2r,1-r\}\right],$$
$$+(1-\xi)\min\left\{r, 1-\left[(1-r)\left(-\frac{(v-\rho)}{2\epsilon t}+\frac{C-\lambda+\epsilon}{2\epsilon}\right)\right]-r\right\}.$$

Let $1-A \equiv 1$, and $1-B \equiv \left[1-\left(-\frac{v-\rho}{2\epsilon t}+\frac{1}{2}+\frac{C-\lambda}{2\epsilon}\right)\right]$. Observe, $1-A > 1-B$.

Analyzing different regions of r, we have that (a) if $r \leq \frac{1}{2}$ then $\frac{\partial}{\partial(v-\rho)} \geq 0$ and $\frac{\partial}{\partial\lambda} \geq 0$, (b) if $r > \frac{1}{2}$ then $\frac{\partial}{\partial(v-\rho)} > 0$ and $\frac{\partial}{\partial\lambda} > 0$.

(3C): We have

$$\mathbb{E}\min\{\tilde{r},\, 1-\tilde{r}\} = \xi\left[\min\left\{r\left(-\frac{v-\rho}{2\epsilon t}+\frac{(-\lambda+\epsilon)}{2\epsilon}\right), (1-r)\left(-\frac{v-\rho}{2\epsilon t}+\frac{C-\lambda+\epsilon}{2\epsilon}\right)\right\}\right.$$
$$+\left(\frac{v-\rho}{2\epsilon t}-\frac{\lambda+\epsilon}{2\epsilon}\right)\min\{r,1-r\} + \left(\frac{-C+2\epsilon+2\lambda}{2\epsilon}\right)\min\{2r,1-r\}\right]$$
$$+(1-\xi)\left[\min\left\{r\left(-\frac{(v-\rho)}{2\epsilon t}+\frac{(-\lambda+\epsilon)}{2\epsilon}\right), (1-r)\left(-\frac{(v-\rho)}{2\epsilon t}+\frac{(C-\lambda+\epsilon)}{2\epsilon}\right)\right\}+\right.$$
$$\left.\left.\min\left\{r\left[1-\left(-\frac{(v-\rho)}{2\epsilon t}+\frac{(-\lambda+\epsilon)}{2\epsilon}\right)\right], (1-r)\left[1-\left(-\frac{(v-\rho)}{2\epsilon t}+\frac{(C-\lambda+\epsilon)}{2\epsilon}\right)\right]\right\}\right]\right].$$

Examining the different ranges of r, we have (a) if $r \leq \frac{1}{2}$ then $\frac{\partial}{\partial(v-\rho)} \geq 0$ and $\frac{\partial}{\partial\lambda} \geq 0$; (b) if $r > \frac{1}{2}$ then $\frac{\partial}{\partial(v-\rho)} < 0$ and $\frac{\partial}{\partial\lambda} < 0$.

(3D): We have, $\mathbb{E}\min\{\tilde{r},\, 1-\tilde{r}\}$

$$= \xi\left[\min\left\{r\left(1-\frac{v-\rho}{\epsilon t}\right), (1-r)\left(-\frac{(v-\rho)}{2\epsilon t}+\frac{C-\lambda+\epsilon}{2\epsilon}\right)\right\} + \left(\frac{v-\rho}{2\epsilon t}-\frac{(C-\epsilon-\lambda)}{2\epsilon}\right)\min\{2r,1-r\}\right]$$
$$+(1-\xi)\left[\min\left\{r\left(1-\frac{(v-\rho)}{\epsilon t}\right), (1-r)\left(-\frac{(v-\rho)}{2\epsilon t}+\frac{C-\lambda+\epsilon}{2\epsilon}\right)\right\}+\right.$$
$$\left.\min\left\{r\left[1-\left(1-\frac{(v-\rho)}{\epsilon t}\right)\right], (1-r)\left[1-\left(-\frac{(v-\rho)}{2\epsilon t}+\frac{C-\lambda+\epsilon}{2\epsilon}\right)\right]\right\}\right].$$

Examining the different ranges of r, we have (a) if $r > \frac{1}{2}$ then $\frac{\partial}{\partial(v-\rho)} < 0$ and $\frac{\partial}{\partial\lambda} > 0$; (b) if $r \leq \frac{1}{2}$, $\frac{\partial}{\partial\lambda} > 0$ and $\frac{\partial}{\partial(v-\rho)} = \xi\frac{v-\rho}{2\epsilon t}(1-3r) + (1-\xi)\frac{v-\rho}{2\epsilon t}(1-2r) = \frac{v-\rho}{2\epsilon t}(1+(2-5\xi)r)$. Therefore, if $r \leq \frac{1}{2}$, $\frac{\partial}{\partial(v-\rho)} \geq 0$ if $\xi \leq \frac{2}{5}$ and $\frac{\partial}{\partial(v-\rho)} < 0$ if $\xi > \frac{2}{5}$.

(4A): We have $\mathbb{E}\min\{\tilde{r},\, 1-\tilde{r}\} = \xi\left[\left(\frac{\lambda}{\epsilon}\right)\min\{r,1-r\} + \left(\frac{-C+2\epsilon}{2\epsilon}\right)\min\{2r,1-r\}\right] + (1-\xi)\min\{r,1-r\}$. So we

have, $\frac{\partial}{\partial(v-\rho)} = 0$ and $\frac{\partial}{\partial\lambda} > 0$.

(4B): We have, $\mathbb{E}\min\{\tilde{r}, 1-\tilde{r}\} =$

$$\xi\left[\left(\frac{\lambda}{\epsilon}\right)\min\{r, 1-r\} + \left(\frac{-C+2\epsilon}{2\epsilon}\right)\min\{2r, 1-r\}\right] + (1-\xi)\min\left\{r, 1-(1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon}\right) - r\right\}.$$

Let $1-A \equiv\equiv 1$, and $1-B \equiv \left[1-\left(-\frac{v-\rho}{2\epsilon t} + \frac{1}{2} + \frac{C-\lambda}{2\epsilon}\right)\right]$. Observe, $1-A > 1-B$. Thus, if $r \leq \frac{1-B}{1-A+1-B}$, $\frac{\partial}{\partial(v-\rho)} = 0$ and $\frac{\partial}{\partial\lambda} > 0$; if $r > \frac{1-B}{1-A+1-B}$, $\frac{\partial}{\partial(v-\rho)} > 0$ and $\frac{\partial}{\partial\lambda} > 0$.

(4C): We have,

$$\mathbb{E}\min\{\tilde{r}, 1-\tilde{r}\} = \xi\left[\min\left\{r\left(-\frac{v-\rho}{2\epsilon t} + \frac{\lambda+\epsilon}{2\epsilon}\right), (1-r)\left(-\frac{v-\rho}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon}\right)\right\}\right.$$
$$+ \left(\frac{v-\rho}{2\epsilon t} - \frac{(-\lambda+\epsilon)}{2\epsilon}\right)\min\{r, 1-r\} + \left(\frac{-C+2\epsilon}{2\epsilon}\right)\min\{2r, 1-r\}\right]$$
$$+ (1-\xi)\left[\min\left\{r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(\lambda+\epsilon)}{2\epsilon}\right), (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(C-\lambda+\epsilon)}{2\epsilon}\right)\right\} +\right.$$
$$\left.\min\left\{r\left[1-\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(\lambda+\epsilon)}{2\epsilon}\right)\right], (1-r)\left[1-\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(C-\lambda+\epsilon)}{2\epsilon}\right)\right]\right\}\right].$$

Analyzing all the cases, we have (i) if $r \leq \frac{1}{2}$ then $\frac{\partial}{\partial(v-\rho)} > 0$ and $\frac{\partial}{\partial\lambda} > 0$; (ii) if $r > \frac{1}{2}$ then $\frac{\partial}{\partial(v-\rho)} < 0$ and $\frac{\partial}{\partial\lambda} > 0$.

(4D): We have, $\mathbb{E}\min\{\tilde{r}, 1-\tilde{r}\} =$

$$\xi\left[\min\left\{r\left(1-\frac{v-\rho}{\epsilon t}\right), (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon}\right)\right\} + \left(\frac{v-\rho}{2\epsilon t} - \frac{(C-\epsilon-\lambda)}{2\epsilon}\right)\min\{2r, 1-r\}\right],$$
$$+ (1-\xi)\left[\min\left\{r\left(1-\frac{(v-\rho)}{\epsilon t}\right), (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon}\right)\right\} +\right.$$
$$\left.\min\left\{r\left[1-\left(1-\frac{(v-\rho)}{\epsilon t}\right)\right], (1-r)\left[1-\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon}\right)\right]\right\}\right].$$

Examining the different ranges of priors r, we have (a) if $r > \frac{1}{2}$, then $\frac{\partial}{\partial(v-\rho)} < 0$ and $\frac{\partial}{\partial\lambda} > 0$; (b) if $r \leq \frac{1}{2}$, $\frac{\partial}{\partial\lambda} \geq 0$, and (b.1) $\frac{\partial}{\partial(v-\rho)} \geq 0$ if $\xi \leq \frac{2}{5}$ and (b.2) $\frac{\partial}{\partial(v-\rho)} < 0$ if $\xi > \frac{2}{5}$.

(5A): We have, $\mathbb{E}\min\{\tilde{r}, 1-\tilde{r}\} =$

$$\xi\left[\left(\frac{\lambda}{\epsilon}\right)\min\{r, 1-r\} + \left(\frac{C-2\lambda}{2\epsilon}\right)\min\{2r, 1-r\} + \left(\frac{\lambda-C+\epsilon}{\epsilon}\right)\min\{r, 1-r\}\right] + (1-\xi)\min\{r, 1-r\}.$$

Hence, we have $\frac{\partial}{\partial(v-\rho)} = 0$ and $\frac{\partial}{\partial\lambda} > 0$.

(5B): We have

$$\mathbb{E}\min\{\tilde{r}, 1-\tilde{r}\} = \xi\left[\left(\frac{\lambda}{\epsilon}\right)\min\{r, 1-r\} + \left(\frac{C-2\lambda}{2\epsilon}\right)\min\{2r, 1-r\} + \left(\frac{\lambda-C+\epsilon}{\epsilon}\right)\min\{r, 1-r\}\right],$$
$$+ (1-\xi)\min\left\{r, 1-(1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon}\right) - r\right\}.$$

For all ranges of r, we have $\frac{\partial}{\partial(v-\rho)} \geq 0$ and $\frac{\partial}{\partial\lambda} \geq 0$.

(5C): We have, $\mathbb{E}\min\{\tilde{r}, 1-\tilde{r}\} =$

$$\xi\left[\min\left\{r\left(-\frac{v-\rho}{2\epsilon t} + \frac{\lambda+\epsilon}{2\epsilon}\right), (1-r)\left(-\frac{v-\rho}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon}\right)\right\}\right.$$
$$+ \left(\frac{v-\rho}{2\epsilon t} - \frac{(-\lambda+\epsilon)}{2\epsilon}\right)\min\{r, 1-r\} + \left(\frac{C-2\lambda}{2\epsilon}\right)\min\{2r, 1-r\} + \left(\frac{\lambda-C+\epsilon}{\epsilon}\right)\min\{r, 1-r\}\right]$$
$$+ (1-\xi)\left[\min\left\{r\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(\lambda+\epsilon)}{2\epsilon}\right), (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(C-\lambda+\epsilon)}{2\epsilon}\right)\right\} +\right.$$
$$\left.\min\left\{r\left[1-\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(\lambda+\epsilon)}{2\epsilon}\right)\right], (1-r)\left[1-\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{(C-\lambda+\epsilon)}{2\epsilon}\right)\right]\right\}\right].$$

Author: *Analysis of the Lean Startup*
Article submitted to *Marketing Science*; manuscript no. MKSC-18-0058.R4

55

We have, (a) if $r \leq \frac{1}{2}$, then $\frac{\partial}{\partial(v-\rho)} > 0$ and $\frac{\partial}{\partial \lambda} > 0$; (b) if $r > \frac{1}{2}$, then $\frac{\partial}{\partial(v-\rho)} < 0$ and $\frac{\partial}{\partial \lambda} > 0$.

(5D): We have,

$$
\begin{aligned}
\mathbb{E}\min\{\tilde{r}, 1-\tilde{r}\} = \xi &\left[\min\left\{r\left(1 - \frac{v-\rho}{\epsilon t}\right), (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon}\right)\right\}\right. \\
&+ \left(\frac{v-\rho}{2\epsilon t} - \frac{(\lambda - C + \epsilon)}{2\epsilon}\right)\min\{2r, 1-r\} + \left(\frac{\lambda-C+\epsilon}{\epsilon}\right)\min\{r, 1-r\}\Bigg] \\
&+ (1-\xi)\left[\min\left\{r\left(1 - \frac{(v-\rho)}{\epsilon t}\right), (1-r)\left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon}\right)\right\} + \right. \\
&\left.\min\left\{r\left[1 - \left(1 - \frac{(v-\rho)}{\epsilon t}\right)\right], (1-r)\left[1 - \left(-\frac{(v-\rho)}{2\epsilon t} + \frac{C-\lambda+\epsilon}{2\epsilon}\right)\right]\right\}\right].
\end{aligned}
$$

Taking the derivative, we have if $r > \frac{1}{2}$, then $\frac{\partial}{\partial(v-\rho)} < 0$ and $\frac{\partial}{\partial \lambda} > 0$; (ii) if $r \leq \frac{1}{2}$, then $\frac{\partial}{\partial \lambda} > 0$ and (ii-a) $\frac{\partial}{\partial(v-\rho)} \geq 0$ if $\xi \leq \frac{2}{5}$ and $\frac{\partial}{\partial(v-\rho)} < 0$ if $\xi > \frac{2}{5}$.

Combining the expressions for the slopes, we find that if $r > 0.5$, it lies in the boundary between region (1B) $\cup$ (2B) $\cup$ (3B) and region (1C) $\cup$ (2C) $\cup$ (3C); if $r < 0.5$, it lies in the boundary between region (1C) $\cup$ (2C) and region (1D) $\cup$ (2D). If $r = 0.5$, the regions (1C) $\cup$ (2C) $\cup$ (3C) correspond to the optimal test product. $\square$

Proof of Corollary 1. The properties follow directly from the expressions from Proposition 2. $\square$

Proof of Corollary A.1 (Corollary 2 is a special case). The expressions are derived in a straightforward manner by substituting the necessary values to the expressions of $\mathbb{E}_{s_i}\min\{\tilde{r}, 1-\tilde{r}\}$ in Appendix B. $\square$

Proof of Proposition 4. (i) Observe that the set of optimal test product's identified in Proposition 2 whose design Λ lie outside the interval $[0, C]$ must pivot longer distance (lower quality V^* and hence demand) than the optimal test product located in $\Lambda = 0$ or $\Lambda = C$. Hence it is optimal to select $\Lambda = 0$ or $\Lambda = C$.

Moreover, if pivoting is costlier for higher quality test product, it is optimal to minimize v while keeping $v - \rho$ constant, which is achieved when $\rho = 0$.

(ii) If $f > t$, it is suboptimal to do pivoting. Thus, the test product design becomes the same as the second stage problem, resulting in $\lambda^* = \Lambda^*$ in Proposition 1. $\square$

Proof of Lemma 4. Taking the first order condition of $\pi^*(\tilde{r})$ with respect to p, we have

$$
\begin{aligned}
\frac{\partial \pi^*(\tilde{r})}{\partial p} = 0 &\iff \frac{\partial}{\partial p}p\left[\frac{V-p}{\epsilon t} - \left(\frac{V-p}{\epsilon t} - \left(1 - \frac{C}{2\epsilon}\right)\right)\min\{\tilde{r}, 1-\tilde{r}\}\right] = 0 \\
&\iff p\left(-\frac{1}{\epsilon t} + \frac{\min\{\tilde{r}, 1-\tilde{r}\}}{\epsilon t}\right) + \frac{V-p}{\epsilon t} - \left(\frac{V-p}{\epsilon t} - \left(1 - \frac{C}{2\epsilon}\right)\right)\min\{\tilde{r}, 1-\tilde{r}\} = 0 \\
&\iff \frac{V}{\epsilon t}(1 - \min\{\tilde{r}, 1-\tilde{r}\}) + \left(1 - \frac{C}{2\epsilon}\right)\min\{\tilde{r}, 1-\tilde{r}\} = \frac{2p}{\epsilon t}(1 - \min\{\tilde{r}, 1-\tilde{r}\}) \\
&\iff p^* = \frac{V}{2} + \frac{\min\{\tilde{r}, 1-\tilde{r}\}}{\max\{\tilde{r}, 1-\tilde{r}\}}\frac{\epsilon t}{2}\left(1 - \frac{C}{2\epsilon}\right). \quad \square
\end{aligned}
$$

Proof of Proposition 5. Substituting the expression for p^* in Lemma 4,

$$
\frac{V-p^*}{\epsilon t} = \frac{V}{2\epsilon t} - \frac{\min\{\tilde{r}, 1-\tilde{r}\}}{\max\{\tilde{r}, 1-\tilde{r}\}}\frac{1}{2}\left(1 - \frac{C}{2\epsilon}\right),
$$

$$
\frac{V-p^*}{\epsilon t}\max\{\tilde{r}, 1-\tilde{r}\} = \frac{V}{2\epsilon t}\max\{\tilde{r}, 1-\tilde{r}\} - \min\{\tilde{r}, 1-\tilde{r}\}\frac{1}{2}\left(1 - \frac{C}{2\epsilon}\right).
$$

56

Author: *Analysis of the Lean Startup*
Article submitted to *Marketing Science*; manuscript no. MKSC-18-0058.R4

Substituting the expression for p^* and $\frac{V-p^*}{\epsilon t}\max\{\tilde{r},1-\tilde{r}\}$, we have

$$
\begin{aligned}
\pi^*(\tilde{r}) &= p\left[\frac{V-p}{\epsilon t} - \left(\frac{V-p}{\epsilon t} - \left(1-\frac{C}{2\epsilon}\right)\right)\min\{\tilde{r},1-\tilde{r}\}\right] = p\left[\frac{V-p}{\epsilon t}\max\{\tilde{r},1-\tilde{r}\} + \left(1-\frac{C}{2\epsilon}\right)\min\{\tilde{r},1-\tilde{r}\}\right] \\
&= \left[\frac{V}{2} + \frac{\min\{\tilde{r},1-\tilde{r}\}}{\max\{\tilde{r},1-\tilde{r}\}}\frac{\epsilon t}{2}\left(1-\frac{C}{2\epsilon}\right)\right]\cdot\left[\frac{V}{2\epsilon t}\max\{\tilde{r},1-\tilde{r}\} + \min\{\tilde{r},1-\tilde{r}\}\frac{1}{2}\left(1-\frac{C}{2\epsilon}\right)\right] \\
&= \frac{1}{4}\left[V + \frac{\min\{\tilde{r},1-\tilde{r}\}}{\max\{\tilde{r},1-\tilde{r}\}}\epsilon t\left(1-\frac{C}{2\epsilon}\right)\right]\cdot\left[\frac{V}{\epsilon t}\max\{\tilde{r},1-\tilde{r}\} + \min\{\tilde{r},1-\tilde{r}\}\left(1-\frac{C}{2\epsilon}\right)\right] \\
&= \frac{\epsilon t}{4}\left[\frac{V}{\epsilon t} + \frac{\min\{\tilde{r},1-\tilde{r}\}}{\max\{\tilde{r},1-\tilde{r}\}}\left(1-\frac{C}{2\epsilon}\right)\right]\cdot\left[\frac{V}{\epsilon t} - \left[\frac{V}{\epsilon t} - \left(1-\frac{C}{2\epsilon}\right)\right]\min\{\tilde{r},1-\tilde{r}\}\right]
\end{aligned}
$$

Thus, taking the expectation with respect to $\tilde{r}$, we have the expression for $\mathbb{E}_{s_i}\pi^*(\tilde{r},\lambda,v)$. $\quad\square$

LEMMA C.1. *Let $v \equiv v - \rho$ for notational simplicity, and let us denote the probability that the test product results in sale when $\lambda = W$ and $\lambda \neq W$ respectively as*

$$
q_1(v) = \min\left(1,\frac{v}{\epsilon t}\right), \quad q_2(v) = \min\left(1,\frac{v}{2\epsilon t} - \frac{C-\epsilon}{2\epsilon}\right). \tag{C-1}
$$

Then $v^* \in \{\epsilon t, (C-\epsilon)t\}$ maximizes $\delta q_1(v) - (1-\delta)q_2(v)$ for any $\delta \in [0,1]$. Moreover, if $C < 2\epsilon$, then $q_2(\epsilon t) < 1 - q_1((C-\epsilon)t)$. In other words, $\frac{q_2(\epsilon t)}{1-q_1((C-\epsilon)t)+q_2(\epsilon t)} < 0.5$.

Proof of Lemma C.1. We have,

$$
\delta q_1(v) - (1-\delta)q_2(v) = \min\left(\delta,\frac{\delta}{\epsilon t}v\right) - \min\left(1-\delta,\frac{1-\delta}{2\epsilon t}v - \frac{(1-\delta)(C-\epsilon)}{2\epsilon}\right).
$$

Thus, the piecewise linear function can achieve the maximum difference in the point where either $q_1(v)$ or $q_2(v)$ hit the inflection points $v = \epsilon t$ and $v = (C-\epsilon)t$ respectively. Moreover, since $q_2(\epsilon t) = 1 - \frac{C}{2\epsilon}$ and $q_1((C-\epsilon)t) = \frac{C-\epsilon}{\epsilon}$,

$$
q_2(\epsilon t) < 1 - q_1((C-\epsilon)t) \Leftrightarrow 1 - \frac{C}{2\epsilon} < 1 - \frac{C-\epsilon}{\epsilon} \Leftrightarrow \frac{C-\epsilon}{\epsilon} < \frac{C}{2\epsilon} \Leftrightarrow \epsilon > C/2. \quad\square
$$

Proof of Proposition 5 and Corollary 3. To simplify notations, let

$$
D_1 \triangleq \int_{-\frac{V-p}{t}+\Lambda}^{\frac{V-p}{t}+\Lambda} h(x|W=\Lambda)dx, \qquad D_2 \triangleq \int_{-\frac{V-p}{t}+\Lambda}^{\frac{V-p}{t}+\Lambda} h(x|W\neq\Lambda)dx,
$$

so that D_1 and D_2 are demand from launching the ideal product and non-ideal product, respectively. Then, the optimal final stage profit given posterior $\tilde{r}$ is

$$
\pi^*(\tilde{r}) = \max\{\tilde{r}D_1 + (1-\tilde{r})D_2, (1-\tilde{r})D_1 + \tilde{r}D_2\}.
$$

We now examine the launch of test products. To simplify notation, let $v \equiv v - \rho$. Let n denote the period to go. We will prove by induction (a) the expression for $(\lambda_n^*,v_n^*) \,\forall n$, (b) $\pi_n^*(\lambda^*,v^*) = A_n + B_n|0.5 - q_2(\epsilon t)^n|$, $A_n, B_n > 0 \,\forall n$, and (c) the expression for $\beta_n(0.5)$.

(Base Case). Suppose $n = 1$. The expected profit is

$$
\begin{aligned}
\pi(\lambda=0,v|r) &= [q_1(v)r + q_2(v)(1-r)]\cdot\max\left\{\frac{q_1(v)rD_1 + q_2(v)(1-r)D_2}{q_1(v)r + q_2(v)(1-r)}, \frac{q_2(v)(1-r)D_1 + q_1(v)rD_2}{q_1(v)r + q_2(v)(1-r)}\right\} \\
&\quad + [(1-q_1(v))r + (1-q_2(v))(1-r)]\cdot\max\left\{\frac{(1-q_2(v))(1-r)D_1 + (1-q_1(v))rD_2}{(1-q_1(v))r + (1-q_2(v))(1-r)}, \right. \\
&\quad\qquad\left. \frac{(1-q_1(v))rD_1 + (1-q_2(v))(1-r)D_2}{(1-q_1(v))r + (1-q_2(v))(1-r)}\right\} \\
&= \max\{q_1(v)rD_1 + q_2(v)(1-r)D_2, q_2(v)(1-r)D_1 + q_1(v)rD_2\} \tag{C-2} \\
&\quad + \max\{(1-q_2(v))(1-r)D_1 + (1-q_1(v))rD_2, (1-q_1(v))rD_1 + (1-q_2(v))(1-r)D_2\}.
\end{aligned}
$$

Author: *Analysis of the Lean Startup*
Article submitted to *Marketing Science*; manuscript no. MKSC-18-0058.R4

57

If the prior belief $r \geq 0.5$, staying course is optimal in the event of a sale of test product since $q_1(v) \geq q_2(v)$ for all v. Thus,

$$
\begin{aligned}
\pi(\lambda = 0, v | r \geq 0.5) &= q_1(v) r D_1 + q_2(v)(1-r) D_2 \\
&\quad + \max\{(1 - q_2(v))(1-r) D_1 + (1 - q_1(v)) r D_2, (1 - q_1(v)) r D_1 + (1 - q_2(v))(1-r) D_2\} \\
&= \max\{[q_1(v) r + (1 - q_2(v))(1-r)] D_1 + [q_2(v)(1-r) + (1 - q_1(v)) r] D_2, r D_1 + (1-r) D_2\} \\
&= r D_1 + (1-r) D_2 + \max\{0, [(1 - 2r) + \{r q_1(v) - (1-r) q_2(v)\}](D_1 - D_2)\}.
\end{aligned}
$$

Similarly, if $r \leq 0.5$, pivoting is optimal in the event of a no sale of test product since $q_1(v) \geq q_2(v)$ for all v. Thus,

$$
\begin{aligned}
\pi(\lambda = 0, v | r \leq 0.5) &= (1 - q_2(v))(1-r) D_1 + (1 - q_1(v)) r D_2 \\
&\quad + \max\{q_1(v) r D_1 + q_2(v)(1-r) D_2, q_2(v)(1-r) D_1 + q_1(v) r D_2\} \\
&= \max\{[(1 - q_2(v))(1-r) + q_1(v) r] D_1 + [(1 - q_1(v)) r + q_2(v)(1-r)] D_2, (1-r) D_1 + r D_2\} \\
&= (1-r) D_1 + r D_2 + \max\{0, \{q_1(v) r - (1-r) q_2(v)\}(D_1 - D_2)\}.
\end{aligned}
$$

Thus, we have

$$
\pi(\lambda = 0, v | r) = \begin{cases} r D_1 + (1-r) D_2 + \max\{0, [(1 - 2r) + \{r q_1(v) - (1-r) q_2(v)\}](D_1 - D_2)\}, & r \geq 0.5, \\ (1-r) D_1 + r D_2 + \max\{0, \{q_1(v) r - (1-r) q_2(v)\}(D_1 - D_2)\}, & r \leq 0.5. \end{cases}
$$

Similarly, we have

$$
\pi(\lambda = C, v | r) = \begin{cases} r D_1 + (1-r) D_2 + \max\{0, \{(1-r) q_1(v) - r q_2(v)\}(D_1 - D_2)\}, & r \geq 0.5, \\ (1-r) D_1 + r D_2 + \max\{0, [(2r - 1) + \{(1-r) q_1(v) - r q_2(v)\}](D_1 - D_2)\}, & r \leq 0.5. \end{cases}
$$

By Lemma C.1, we need to only consider $v \in \{\epsilon t, (C - \epsilon) t\}$ when evaluating for optimal v^*. If $v^* = \epsilon t$, then $q_1(v^*) = 1$ and $q_2(v^*) < 1$; if $v^* = (C - \epsilon) t$, then $q_1(v^*) > 0$ and $q_2(v^*) = 0$. Identifying the optimal development level of the test product v^* require comparing the weighted difference $\rho q_1(v) - (1 - \rho) q_2(v)$, $\rho \in [0, 1]$.

$$
\rho - (1 - \rho) q_2(\epsilon t) \gtrless_{v^* = (C - \epsilon) t}^{v^* = \epsilon t} \rho q_1((C - \epsilon) t) \iff \rho(1 - q_1((C - \epsilon) t)) \gtrless_{v^* = (C - \epsilon) t}^{v^* = \epsilon t} (1 - \rho) q_2(\epsilon t)
$$

$$
\iff \rho \gtrless_{v^* = (C - \epsilon) t}^{v^* = \epsilon t} \frac{q_2(\epsilon t)}{1 - q_1((C - \epsilon) t) + q_2(\epsilon t)}. \tag{C-3}
$$

For the case $r \geq 0.5$, we have, $\pi(\lambda = 0, v | r) = r D_1 + (1-r) D_2 + \max\{0, [(1 - 2r) + \{r q_1(v) - (1-r) q_2(v)\}](D_1 - D_2)\}$, which involves maximizing $r q_1(v) - (1-r) q_2(v)$. Since, $r > 0.5 > \frac{q_2(\epsilon t)}{1 - q_1((C - \epsilon) t) + q_2(\epsilon t)}$ (Lemma C.1), $v^* = \epsilon t$, i.e., $\pi(\lambda = 0, v = \epsilon t | r) > \pi(\lambda = 0, v = (C - \epsilon) t | r)$.

We next show that $\pi(\lambda = 0, v = \epsilon t | r) \geq \max\{\pi(\lambda = C, v = \epsilon t | r), \pi(\lambda = C, v = (C - \epsilon) t | r)\}$. We have, $\pi(\lambda = 0, v = \epsilon t | r) > \pi(\lambda = C, v = \epsilon t | r) \iff 1 - r - (1-r) q_2(\epsilon t) > (1-r) - r q_2(\epsilon t) \iff r > 0.5$; and $\pi(\lambda = 0, v = \epsilon t | r) > \pi(\lambda = C, v = (C - \epsilon) t | r) \iff 1 - r - (1-r) q_2(\epsilon t) > (1-r) q_1((C - \epsilon) t) \iff 1 - q_1((C - \epsilon) t) > q_2(\epsilon t)$. Thus, if $r > 0.5$, it is optimal to launch the test product with $\lambda = 0$, and at development level $v = \epsilon t$.

Similarly, one can show that it is optimal to launch the test product with $\lambda = C$ at development level $v = \epsilon t$ when $r < 0.5$. Thus, $(\lambda^*, v^*) = (0, \epsilon t)$ if $r \geq 0.5$ and $(\lambda^*, v^*) = (C, \epsilon t)$ if $r < 0.5$.

58

Author: *Analysis of the Lean Startup*
Article submitted to *Marketing Science*; manuscript no. MKSC-18-0058.R4

Thus, combining the expression for $v^* = \epsilon t$, we have

$$\pi(\lambda^*(v), v = \epsilon t | r) = \begin{cases} (rD_1 + (1-r)D_2)p + [(1-2r) + r - (1-r)q_2(\epsilon t)](D_1 - D_2)p, \ r \geq 0.5, \\ ((1-r)D_1 + rD_2)p + [(2r-1) + (1-r) - rq_2(\epsilon t)](D_1 - D_2)p, \ r \leq 0.5. \end{cases}$$

$$= \left[D_1 p - \frac{q_2(\epsilon t)}{2}(D_1 - D_2)p \right] + q_2(\epsilon t)(D_1 - D_2)p \left| r - \frac{1}{2} \right| \qquad \text{(C-4)}$$

$$\equiv A_0 + B_0 |r - 0.5|. \qquad \text{(C-5)}$$

Moreover, observe that

$$\beta_0(v = \epsilon t | r) = \begin{cases} (r + \beta(v = \epsilon t | r))D_1 p + (1 - r - \beta(v = \epsilon t | r))D_2 p, \ r \geq 0.5, \\ (1 - r + \beta(v = \epsilon t | r))D_1 + (r - \beta(v = \epsilon t | r))D_2 p, \ r \leq 0.5. \end{cases}$$

$$= \begin{cases} (1-r)(1 - q_2(\epsilon t)), \ r \geq 0.5, \\ r(1 - q_2(\epsilon t)), \qquad r \leq 0.5. \end{cases}$$

(Induction Step) Suppose claims (a), (b), and (c) are true for $n-1$ periods to go. Then, after launching a test product and observing sales, the expected profit $\pi(\lambda, v|r)$ for the choice of test product with n periods to go is

$$\pi_n(\lambda = 0, v|r) = [q_1(v)r + q_2(v)(1-r)]\pi^*_{n-1}\left(\frac{q_1(v)r}{q_1(v)r + q_2(v)(1-r)} \right)$$

$$+ [(1-q_1(v))r + (1-q_2(v))(1-r)]\pi^*_{n-1}\left(\frac{(1-q_1(v))r}{(1-q_1(v))r + (1-q_2(v))(1-r)} \right),$$

$$\pi_n(\lambda = C, v|r) = [q_1(v)(1-r) + q_2(v)r]\pi^*_{n-1}\left(\frac{q_1(v)(1-r)}{q_1(v)(1-r) + q_2(v)r} \right)$$

$$+ [(1-q_1(v))(1-r) + (1-q_2(v))r]\pi^*_{n-1}\left(\frac{(1-q_1(v))(1-r)}{(1-q_1(v))(1-r) + (1-q_2(v))r} \right).$$

Examining the first expression, we have

$$\pi(\lambda = 0, v|r) = [q_1(v)r + q_2(v)(1-r)]\left\{ A + B \left| \frac{q_1(v)rq_2(\epsilon t)^{n-1}}{q_1(v)r + q_2(v)(1-r)} - \frac{1}{2} \right| \right\}$$

$$+ [(1-q_1(v))r + (1-q_2(v))(1-r)]\left\{ A + B \left| \frac{(1-q_1(v))rq_2(\epsilon t)^{n-1}}{(1-q_1(v))r + (1-q_2(v))(1-r)} \right| \right\}$$

$$= A + B \left\{ [q_1(v)r + q_2(v)(1-r)] \left| \frac{q_1(v)rq_2(\epsilon t)^{n-1}}{q_1(v)r + q_2(v)(1-r)} - \frac{1}{2} \right| \right.$$

$$+ \left. [(1-q_1(v))r + (1-q_2(v))(1-r)] \left| \frac{(1-q_1(v))rq_2(\epsilon t)^{n-1}}{(1-q_1(v))r + (1-q_2(v))(1-r)} - \frac{1}{2} \right| \right\}.$$

Observe that $\frac{(1-q_1(v))r}{(1-q_1(v))r + (1-q_2(v))(1-r)} < r < \frac{q_1(v)r}{q_1(v)r + q_2(v)(1-r)}$, so we have three cases:

(i) If $\frac{(1-q_1(v))rq_2(\epsilon t)^{n-1}}{(1-q_1(v))r + (1-q_2(v))(1-r)} < \frac{q_1(v)rq_2(\epsilon t)^{n-1}}{q_1(v)r + q_2(v)(1-r)} < \frac{1}{2}$ (equivalently, $r < \frac{q_2(v)}{q_1(v)(2q_2(\epsilon t)^{n-1}-1) + q_2(v)}$), then

$$\pi(\lambda = 0, v|r) = A + B \left\{ -[q_1(v)r + q_2(v)(1-r)] \left(\frac{q_1(v)rq_2(\epsilon t)^{n-1}}{q_1(v)r + q_2(v)(1-r)} - \frac{1}{2} \right) \right.$$

$$- \left. [(1-q_1(v))r + (1-q_2(v))(1-r)] \left(\frac{(1-q_1(v))rq_2(\epsilon t)^{n-1}}{(1-q_1(v))r + (1-q_2(v))(1-r)} - \frac{1}{2} \right) \right\}$$

$$= A + B \left\{ q_1(v)rq_2(\epsilon t)^{n-1} - \frac{q_1(v)r + q_2(v)(1-r)}{2} + (1-q_1(v))rq_2(\epsilon t)^{n-1} - \frac{(1-q_1(v))r + (1-q_2(v))(1-r)}{2} \right\}$$

$$= A + B \left\{ \frac{1}{2} - rq_2(\epsilon t)^{n-1} \right\}.$$

Author: *Analysis of the Lean Startup*
Article submitted to *Marketing Science*; manuscript no. MKSC-18-0058.R4

59

(ii)   If $\frac{(1-q_1(v))rq_2(\epsilon t)^{n-1}}{(1-q_1(v))r+(1-q_2(v))(1-r)} < \frac{1}{2} < \frac{q_1(v)rq_2(\epsilon t)^{n-1}}{q_1(v)r+q_2(v)(1-r)}$ (equivalently, $r \in$

$\left[\frac{q_2(v)}{q_1(v)(2q_2(\epsilon t)^{n-1}-1)+q_2(v)}, \frac{1-q_2(v)}{(1-q_1(v))(2q_2(\epsilon t)^{n-1}-1)+1-q_2(v)}\right]$), then

$$
\begin{aligned}
\pi(\lambda=0,v|r) &= A + B\left\{[q_1(v)r+q_2(v)(1-r)]\left(\frac{q_1(v)rq_2(\epsilon t)^{n-1}}{q_1(v)r+q_2(v)(1-r)} - \frac{1}{2}\right)\right. \\
&\quad \left. - [(1-q_1(v))r+(1-q_2(v))(1-r)]\left(\frac{(1-q_1(v))rq_2(\epsilon t)^{n-1}}{(1-q_1(v))r+(1-q_2(v))(1-r)} - \frac{1}{2}\right)\right\} \\
&= A + B\left\{q_1(v)rq_2(\epsilon t)^{n-1} - \frac{q_1(v)r+q_2(v)(1-r)}{2} - (1-q_1(v))rq_2(\epsilon t)^{n-1} + \frac{(1-q_1(v))r+(1-q_2(v))(1-r)}{2}\right\} \\
&= A + B\left\{q_1(v)rq_2(\epsilon t)^{n-1} - q_2(v)(1-r) + \frac{1}{2} - r\right\}.
\end{aligned}
$$

(iii) If $\frac{1}{2} < \frac{(1-q_1(v))rq_2(\epsilon t)^{n-1}}{(1-q_1(v))r+(1-q_2(v))(1-r)} < \frac{q_1(v)rq_2(\epsilon t)^{n-1}}{q_1(v)r+q_2(v)(1-r)}$ (equivalently, $r \geq \frac{1-q_2(v)}{(1-q_1(v))(2q_2(\epsilon t)^{n-1}-1)+1-q_2(v)}$), then

$$
\begin{aligned}
\pi(\lambda=0,v|r) &= A + B\left\{[q_1(v)r+q_2(v)(1-r)]\left(\frac{q_1(v)rq_2(\epsilon t)^{n-1}}{q_1(v)r+q_2(v)(1-r)} - \frac{1}{2}\right)\right. \\
&\quad \left. + [(1-q_1(v))r+(1-q_2(v))(1-r)]\left(\frac{(1-q_1(v))rq_2(\epsilon t)^{n-1}}{(1-q_1(v))r+(1-q_2(v))(1-r)} - \frac{1}{2}\right)\right\} \\
&= A + B\left\{q_1(v)rq_2(\epsilon t)^{n-1} - \frac{q_1(v)r+q_2(v)(1-r)}{2} + (1-q_1(v))rq_2(\epsilon t)^{n-1} - \frac{(1-q_1(v))r+(1-q_2(v))(1-r)}{2}\right\} \\
&= A + B\left\{rq_2(\epsilon t)^{n-1} - \frac{1}{2}\right\}.
\end{aligned}
$$

Thus, after algebraic simplifications, we have

$$
\pi^*(\lambda=0,v|r) = \begin{cases} A+B\left\{\frac{1}{2} - rq_2(\epsilon t)^{n-1}\right\}, & r < \frac{q_2(v)}{q_1(v)(2q_2(\epsilon t)^{n-1}-1)+q_2(v)}, \\ A+B\left\{q_1(v)rq_2(\epsilon t)^{n-1} - q_2(v)(1-r) + \frac{1}{2} - r\right\}, & r \in \left[\frac{q_2(v)}{q_1(v)+q_2(v)}, \frac{1-q_2(v)}{(1-q_1(v))(2q_2(\epsilon t)^{n-1}-1)+1-q_2(v)}\right], \\ A+B\left\{rq_2(\epsilon t)^{n-1} - \frac{1}{2}\right\}, & r \geq \frac{1-q_2(v)}{(1-q_1(v))(2q_2(\epsilon t)^{n-1}-1)+1-q_2(v)} \end{cases}
$$

Using the same logic, we have

$$
\pi^*(\lambda=C,v|r) = \begin{cases} A+B\left\{\frac{1}{2} - (1-r)q_2(\epsilon t)^{n-1}\right\}, & (1-r) < \frac{q_2(v)}{q_1(v)(2q_2(\epsilon t)^{n-1}-1)+q_2(v)}, \\ A+B\left\{q_1(v)(1-r)q_2(\epsilon t)^{n-1} - q_2(v)r + \frac{1}{2} - (1-r)\right\}, & (1-r) \in \left[\frac{q_2(v)}{q_1(v)(2q_2(\epsilon t)^{n-1}-1)+q_2(v)}, \frac{1-q_2(v)}{(1-q_1(v))(2q_2(\epsilon t)^{n-1}-1)+1-q_2(v)}\right], \\ A+B\left\{(1-r)q_2(\epsilon t)^{n-1} - \frac{1}{2}\right\}, & (1-r) \geq \frac{1-q_2(v)}{(1-q_1(v))(2q_2(\epsilon t)^{n-1}-1)+1-q_2(v)} \end{cases}
$$

Observe that $\pi^*(\lambda=0,v|r)$ and $\pi^*(\lambda=C,v|r)$ are symmetric around $r=0.5$, and regardless of v, it is always optimal to launch $\lambda=0$ if $r \geq 0.5$ and $\lambda=C$ if $r \leq 0.5$. This results in

$$
\pi^*(\lambda=0,v|r \geq 0.5) = \begin{cases} A+B\left\{q_1(v)rq_2(\epsilon t)^{n-1} - q_2(v)(1-r) + \frac{1}{2} - r\right\}, & r \leq \frac{1-q_2(v)}{(1-q_1(v))(2q_2(\epsilon t)^{n-1}-1)+1-q_2(v)}, \\ A+B\left\{rq_2(\epsilon t)^{n-1} - \frac{1}{2}\right\}, & r \geq \frac{1-q_2(v)}{(1-q_1(v))(2q_2(\epsilon t)^{n-1}-1)+1-q_2(v)} \end{cases}
$$

$$
\pi^*(\lambda=C,v|r < 0.5) = \begin{cases} A+B\left\{q_1(v)(1-r)q_2(\epsilon t)^{n-1} - q_2(v)r + \frac{1}{2} - (1-r)\right\}, & (1-r) \leq \frac{1-q_2(v)}{(1-q_1(v))(2q_2(\epsilon t)^{n-1}-1)+1-q_2(v)}, \\ A+B\left\{(1-r)q_2(\epsilon t)^{n-1} - \frac{1}{2}\right\}, & (1-r) \geq \frac{1-q_2(v)}{(1-q_1(v))(2q_2(\epsilon t)^{n-1}-1)+1-q_2(v)} \end{cases}
$$

Since both expression involve optimizing $\rho q_1(v) - (1-\rho)q_2(v)$ with respect to v for $\rho > 0.5$, by (C-3), $v^* = \epsilon t$. Furthermore, if $v^* = \epsilon t$, the $q_1(\epsilon t) = 1$, so the inequalities become $\frac{1-q_2(v)}{(1-q_1(v))(2q_2(\epsilon t)^{n-1}-1)+1-q_2(v)} = 1$. Hence, we have

$$
\pi^*(\lambda^*(v),v^* = \epsilon t|r) = \begin{cases} \left[D_1 p - \frac{q_2(\epsilon t)}{2}(D_1-D_2)p\right] + q_2(\epsilon t)(D_1-D_2)p\left\{\frac{1}{2} - q_2(\epsilon t)^n(1-r)\right\}, & r \geq 0.5, \\ \left[D_1 p - \frac{q_2(\epsilon t)}{2}(D_1-D_2)p\right] + q_2(\epsilon t)(D_1-D_2)p\left\{\frac{1}{2} - q_2(\epsilon t)^n r\right\}, & r \leq 0.5. \end{cases}
$$

$$
= A_n + B_n|q_2(\epsilon t)^n r - 0.5|.
$$

Furthermore, the expression becomes

$$
\pi^*(\lambda^*(v),v^* = \epsilon t|r) = \begin{cases} \left[D_1 p - \frac{q_2(\epsilon t)}{2}(D_1-D_2)p\right] + q_2(\epsilon t)(D_1-D_2)p\left\{\frac{1}{2} - q_2(\epsilon t)(1-r)\right\}, & r \geq 0.5, \\ \left[D_1 p - \frac{q_2(\epsilon t)}{2}(D_1-D_2)p\right] + q_2(\epsilon t)(D_1-D_2)p\left\{\frac{1}{2} - q_2(\epsilon t)r\right\}, & r \leq 0.5. \end{cases}
$$

$$
= \begin{cases} rD_1 p + (1-r)D_2 p + (1-r)(1 - [q_2(\epsilon t)]^n)(D_1-D_2)p, & r \geq 0.5, \\ (1-r)D_1 p + rD_2 p + r(1 - [q_2(\epsilon t)]^n)(D_1-D_2)p, & r \leq 0.5. \end{cases}
$$

where

$$
\beta_n^* = \begin{cases} (1-r)(1 - [q_2(\epsilon t)]^n), & r \geq 0.5, \\ r(1 - [q_2(\epsilon t)]^n), & r \leq 0.5. \end{cases} \qquad \square
$$

60

Author: *Analysis of the Lean Startup*
Article submitted to *Marketing Science*; manuscript no. MKSC-18-0058.R4

Proof of Proposition 6. Suppose the entrepreneur chooses $\Lambda = 0$. Then, if $W = 0$, he will earn D_1. If $W = C$, he will earn D_2. If $W = -C$, he will also earn D_2 due to the fact that there are two overlaps. Thus, the expected profit is

$$\pi(\Lambda = 0) = r_0 D_1 + r_C D_2 + (1 - r_0 - r_C) D_2 = r_0 D_1 + (1 - r_0) D_2.$$

Suppose the entrepreneur chooses $\Lambda = C$. Then, if $W = C$, he will earn D_1. If $W = 0$, he will earn D_2. However, if $W = -C$, he will earn 0 since there is no overlap between C and $-C$. Thus, the expected profit is

$$\pi(\Lambda = C) = r_C D_1 + r_0 D_2.$$

Similarly, suppose the entrepreneur chooses $\Lambda = -C$. Then, if $W = -C$, he will earn D_1. If $W = 0$, he will earn D_2. However, if $W = C$, he will earn 0 since there is no overlap between C and $-C$. Thus, the expected profit is

$$\pi(\Lambda = -C) = (1 - r_0 - r_C) D_1 + r_0 D_2.$$

The optimal $\pi^* = \max\{\pi(\Lambda = 0), \pi(\Lambda = C), \pi(\Lambda = -C)\}$. Moreover, $\Lambda^* = 0$ if $\pi(\Lambda = 0) > \pi(\Lambda = C)$ and $\pi(\Lambda = 0) > \pi(\Lambda = -C)$; $\Lambda^* = C$ if $\pi(\Lambda = C) > \pi(\Lambda = 0)$ and $\pi(\Lambda = C) > \pi(\Lambda = -C)$; and $\Lambda^* = -C$ if $\pi(\Lambda = -C) > \pi(\Lambda = 0)$ and $\pi(\Lambda = -C) > \pi(\Lambda = C)$. Working out the inequalities we have our result. $\square$

Proof of Proposition 7. We let $v \equiv v - \rho$ to simplify notation. The Bayesian posterior beliefs are as follows:

$$\bar{r}_0 \triangleq P(W = 0 | \text{Sale of } (\lambda, v)) = \frac{P(\text{Sale of } (\lambda, v) | W = 0) \cdot r_0}{P(\text{Sale of } (\lambda, v))},$$

$$\bar{r}_C \triangleq P(W = C | \text{Sale of } (\lambda, v)) = \frac{P(\text{Sale of } (\lambda, v) | W = C) \cdot r_C}{P(\text{Sale of } (\lambda, v))},$$

$$\underline{r}_0 \triangleq P(W = 0 | \text{No sale of } (\lambda, v)) = \frac{(1 - P(\text{No sale of } (\lambda, v) | W = 0)) \cdot r_0}{1 - P(\text{No sale of } (\lambda, v))},$$

$$\underline{r}_C \triangleq P(W = C | \text{No sale of } (\lambda, v)) = \frac{(1 - P(\text{No Sale of } (\lambda, v) | W = C)) \cdot r_C}{1 - P(\text{No sale of } (\lambda, v))},$$

where

$$P(\text{Sale of } (\lambda, v) | W = 0) = \int_{-v/t + \lambda}^{v/t + \lambda} h(x | W = 0) dx,$$

$$P(\text{Sale of } (\lambda, v) | W = C) = \int_{-v/t + \lambda}^{v/t + \lambda} h(x | W = C) dx,$$

$$P(\text{Sale of } (\lambda, v)) = \sum_{w \in \{-C, 0, C\}} r_w P(\text{Sale of } (\lambda, v) | W = w).$$

To simplify notations, we let

$$q_1(v) \triangleq \int_{-\frac{v}{t} + \lambda}^{\frac{v}{t} + \lambda} h(x | W = \lambda) dx, \qquad q_2(v) \triangleq \int_{-\frac{v}{t} + \lambda}^{\frac{v}{t} + \lambda} h(x | |W - \lambda| = C) dx,$$

so that $q_1(v)$ and $q_2(v)$ are probability of sales that the test product will resul in sales when launching the ideal design and adjacent-to-ideal design, respectively. (If $\lambda = C$ and $W = -C$, then the probability is zero.) The $q_1(v)$ and $q_2(v)$ is identical to those in (C-1). We will suppress $q_1 \equiv q_1(v)$ and $q_2 \equiv q_2(v)$.

Author: *Analysis of the Lean Startup*
Article submitted to *Marketing Science*; manuscript no. MKSC-18-0058.R4

61

Suppose $\lambda = 0$.

$$
\begin{aligned}
\mathbb{E}\pi(v) &= P(sale)\pi(\bar{r_0}, r_C^-) + (1 - P(sale))\pi(\underline{r_0}, \underline{r_C}) \\
&= P(sale)p\max\left\{\frac{q_1 r_0}{P(sale)}D_1 + (1 - \frac{q_1 r_0}{P(sale)})D_2,\ \frac{q_2 r_C}{P(sale)}D_1 + \frac{q_1 r_0}{P(sale)})D_2,\ (1 - \frac{q_1 r_0 + q_2 r_C}{P(sale)}D_1 + \frac{q_1 r_0}{P(sale)})D_2\right\} \\
&\quad + P(nosale)p\max\left\{\frac{(1-q_1)r_0}{P(nosale)}D_1 + (1 - \frac{(1-q_1)r_0}{P(nosale)})D_2,\ \frac{(1-q_2)r_C}{P(nosale)}D_1 + \frac{(1-q_1)r_0}{P(nosale)})D_2,\ (1 - \frac{(1-q_1)r_0 + (1-q_2)r_C}{P(nosale)}D_1 + \frac{(1-q_1)r_0}{P(nosale)})D_2\right\} \\
&= p\max\{q_1 r_0 D_1 + (1 - r_0)q_2 D_2,\ q_2 r_C D_1 + r_0 q_1 D_2,\ (1 - r_0 - r_C)q_2 D_1 + r_0 q_1 D_2\} \\
&\quad + p\max\{(1-q_1)r_0 D_1 + (1-r_0)(1-q_2)D_2,\ (1-q_2)r_C D_1 + r_0(1-q_1)D_2,\ (1 - r_0 - r_C)(1-q_2)D_1 + r_0(1-q_1)D_2\}
\end{aligned}
$$

Suppose $\lambda = C$.

$$
\begin{aligned}
\mathbb{E}\pi(v) &= P(sale)\pi(\bar{r_0}, r_C^-) + (1 - P(sale))\pi(\underline{r_0}, \underline{r_C}) \\
&= P(sale)p\max\left\{\frac{q_2 r_0}{P(sale)}D_1 + (1 - \frac{q_2 r_0}{P(sale)})D_2,\ \frac{q_1 r_C}{P(sale)}D_1 + \frac{q_2 r_0}{P(sale)})D_2,\ (1 - \frac{q_2 r_0 + q_1 r_C}{P(sale)}D_1 + \frac{q_2 r_0}{P(sale)})D_2\right\} \\
&\quad + P(nosale)p\max\left\{\frac{(1-q_2)r_0}{P(nosale)}D_1 + (1 - \frac{(1-q_1)r_0}{P(nosale)})D_2,\ \frac{(1-q_1)r_C}{P(nosale)}D_1 + \frac{(1-q_2)r_0}{P(nosale)})D_2,\ (1 - \frac{(1-q_2)r_0 + (1-q_1)r_C}{P(nosale)}D_1 + \frac{(1-q_2)r_0}{P(nosale)})D_2\right\} \\
&= p\max\{q_2 r_0 D_1 + r_C q_1 D_2,\ q_1 r_C D_1 + r_0 q_2 D_2,\ 0 + r_0 q_2 D_2\} \\
&\quad + p\max\{(1-q_2)r_0 D_1 + [(1-r_0) - r_C q_1]D_2,\ (1-q_1)r_C D_1 + r_0(1-q_2)D_2,\ (1 - r_0 - r_C)D_1 + r_0(1-q_2)D_2\}.
\end{aligned}
$$

Suppose $\lambda = -C$.

$$
\begin{aligned}
\mathbb{E}\pi(v) &= P(sale)\pi(\bar{r_0}, r_C^-) + (1 - P(sale))\pi(\underline{r_0}, \underline{r_C}) \\
&= P(sale)p\max\left\{\frac{q_2 r_0}{P(sale)}D_1 + (1 - \frac{q_2 r_0}{P(sale)})D_2,\ 0 + \frac{q_2 r_0}{P(sale)})D_2,\ (1 - \frac{q_2 r_0}{P(sale)})D_1 + \frac{q_2 r_0}{P(sale)})D_2\right\} \\
&\quad + P(nosale)p\max\left\{\frac{(1-q_2)r_0}{P(nosale)}D_1 + (1 - \frac{(1-q_1)r_0}{P(nosale)})D_2,\ \frac{(r_C)}{P(nosale)}D_1 + \frac{(1-q_2)r_0}{P(nosale)})D_2,\ (1 - \frac{(1-q_2)r_0 + r_C}{P(nosale)}D_1 + \frac{(1-q_2)r_0}{P(nosale)})D_2\right\} \\
&= p\max\{q_2 r_0 D_1 + (1 - r_0 - r_C)q_1 D_2,\ 0 + r_0 q_2 D_2,\ q_1(1 - r_0 - r_C)D_1 + r_0 q_2 D_2\} \\
&\quad + p\max\{(1-q_2)r_0 D_1 + [1 - r_0 - (1 - r_0 - r_C)q_1]D_2,\ r_C D_1 + r_0(1-q_2)D_2,\ (1 - r_0 - r_C)(1-q_1)D_1 + r_0(1-q_2)D_2\}.
\end{aligned}
$$

By Lemma C.1, it suffices to check $v = \epsilon t$ and $v = (C - \epsilon)t$. If $v = \epsilon t$, then $q_1 = 1$ and $q_2 < 1$. If $v = (C - \epsilon)t$, then $q_1 < 1$ and $q_2 = 0$. Comparing the values, we find that $v^* = \epsilon t$, thus for simplicity, we will focus on analyzing this case without writing down the expression for $v = (C - \epsilon)t$.

(i) Suppose $r_0 = 0$. Then,

$$\pi(\lambda = 0) = p\max\{q_2 D_2,\ q_2 r_C D_1,\ (1 - r_C)q_2 D_1\} + p\max\{(1-q_2)D_2,\ (1-q_2)r_C D_1,\ (1 - r_C)(1-q_2)D_1\},$$

$$\pi(\lambda = C) = p\max\{r_C q_1 D_2,\ q_1 r_C D_1,\ 0\} + p\max\{(1 - r_C q_1)D_2,\ (1-q_1)r_C D_1,\ (1 - r_C)D_1\},$$

$$\pi(\lambda = -C) = p\max\{(1 - r_C)q_1 D_2,\ 0,\ q_1(1 - r_C)D_1\} + p\max\{(1 - (1 - r_C)q_1)D_2,\ r_C D_1,\ (1 - r_C)(1-q_1)D_1\}.$$

When $v = \epsilon t$,

$$\pi(\lambda = 0) = p\max\{r_C D_1, (1 - r_C)D_1\},$$
$$\pi(\lambda = C) = pD_1,$$
$$\pi(\lambda = -C) = pD_1.$$

Hence, $\lambda^* = \pm C$, $v^* = \epsilon t$.

(ii) Suppose $r_C = 0$ (or $1 - r_0 - r_C = 0$. For simplicity, we will omit this case). Then,

$$
\begin{aligned}
\pi(\lambda = 0) &= p\max\{q_1 r_0 D_1 + (1 - r_0)q_2 D_2,\ r_0 q_1 D_2,\ (1 - r_0)q_2 D_1 + r_0 q_1 D_2\} \\
&\quad + p\max\{(1-q_1)r_0 D_1 + (1-q_2)(1-r_0)D_2,\ (1-q_1)r_0 D_2,\ (1-q_2)(1-r_0)D_1 + r_0(1-q_1)D_2\},
\end{aligned}
$$

$$
\begin{aligned}
\pi(\lambda = C) &= p\max\{q_2 r_0 D_1,\ q_2 r_0 D_2,\ q_2 r_0 D_2\} \\
&\quad + p\max\{(1-q_2)r_0 D_1 + (1-r_0)D_2,\ (1-q_2)r_0 D_2,\ (1-r_0)D_1 + (1-q_2)r_0 D_2\},
\end{aligned}
$$

$$
\begin{aligned}
\pi(\lambda = -C) &= p\max\{q_2 r_0 D_1 + (1-r_0)q_1 D_2,\ r_0 q_2 D_2,\ q_1(1-r_0)D_1 + r_0 q_2 D_2\} \\
&\quad + p\max\{(1-q_2)r_0 D_1 + (1-r_0)(1-q_1)D_2,\ (1-q_2)r_0 D_2,\ (1-r_0)(1-q_1)D_1 + (1-q_2)r_0 D_2\}.
\end{aligned}
$$

62

Author: *Analysis of the Lean Startup*
Article submitted to *Marketing Science*; manuscript no. MKSC-18-0058.R4

When $v = \epsilon t$,

$$\pi(\lambda = 0) = p[r_0 D_1 + (1 - r_0)[q_2 D_1 + (1 - q_2)D1],$$

$$\pi(\lambda = C) = p[q_2 r_0 D_1 + (1 - q_2)(1 - r_0)D_1],$$

$$\pi(\lambda = -C) = p[(1 - r_0)D_1 + r_0[q_2 D_1 + (1 - q2)D1].$$

Hence, $\lambda^* = C$ is ruled out. and $\lambda^* = 0$ if $r_0 > 0.5$ and $\lambda^* = -C$ otherwise.

(iii) Finally, suppose $r_0 = 1/3$. Let $r_C = 1/3$.

$$\pi(\lambda = 0) = p \max\{\frac{1}{3}q_1 D_1 + \frac{2}{3}q_2 D_2, \ \frac{1}{3}q_2 + \frac{1}{3}q_1 D_2, \ \frac{1}{3}q_2 D_1 + \frac{1}{3}q_1 D_2\}$$
$$+ p \max\{\frac{1}{3}(1 - q_1)D_1 + \frac{2}{3}(1 - q_2)D_2, \ \frac{1}{3}(1 - q_2)D_1 + \frac{1}{3}(1 - q_1)D_2, \ \frac{1}{3}(1 - q_2)D_1 + \frac{1}{3}(1 - q_1)D_2\},$$

$$\pi(\lambda = C) = p \max\{\frac{1}{3}q_2 D_1 + \frac{1}{3}q_1 D_2, \ \frac{1}{3}q_1 D_1 + \frac{1}{3}q_2 D_2, \ \frac{1}{3}q_2 D_2\}$$
$$+ p \max\{\frac{1}{3}(1 - q_2)D_1 + (\frac{2}{3} - \frac{1}{3}q_1)D_2, \ \frac{1}{3}(1 - q_1)D_1 + \frac{1}{3}(1 - q_2)D_2, \ \frac{1}{3}D_1 + \frac{1}{3}(1 - q_2)D_2\},$$

$$\pi(\lambda = -C) = p \max\{\frac{1}{3}q_2 D_1 + \frac{1}{3}q_1 D_2, \ \frac{1}{3}q_2 D_2, \ \frac{1}{3}q_1 D_1 + \frac{1}{3}q_2 D_2\}$$
$$+ p \max\{\frac{1}{3}(1 - q_2)D_1 + (\frac{2}{3} - \frac{1}{3}q_1)D_2, \ \frac{1}{3}D_1 + \frac{1}{3}(1 - q_2)D_2, \ \frac{1}{3}(1 - q_1)D_1 + \frac{1}{3}(1 - q_2)D_2\}.$$

Observe that $\pi(\lambda = C) = \pi(\lambda = -C)$. Also, when $v^* = \epsilon t$, then $q_1 = 1$ and we have

$$\pi(\lambda = 0) = p(\frac{2}{3}D_1 + \frac{2}{3}q_2 D_2 - \frac{1}{3}q_2 D_1),$$

$$\pi(\lambda = C) = \pi(\lambda = -C) = p(\frac{2}{3}D_1 + \frac{1}{3}D_2).$$

Since $q_2 < 0.5$, $\lambda = 0$ is ruled out as optimal. Thus, if $r_C > 1/3$, then $\lambda^* = C$, otherwise $\lambda^* = -C$. $\quad\square$