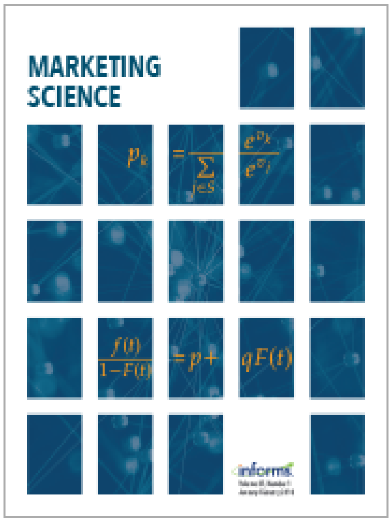


Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Time Inconsistency and Product Design: A Strategic Analysis of Feature Creep

Sanjay Jain

informs
http://pubsonline.informs.org/journal/mksc

MARKETING SCIENCE
Vol. 38, No. 5, September–October 2019, pp. 835–851
ISSN 0732-2399 (print), ISSN 1526-548X (online)

Time Inconsistency and Product Design: A Strategic Analysis of Feature Creep

Sanjay Jain[a]

[a] Naveen Jindal School of Management, University of Texas at Dallas, Richardson, Texas 75080

Contact: sanjay.jain@utdallas.edu, http://orcid.org/0000-0001-6348-9526 (SJ)

Received: October 27, 2017
Revised: September 15, 2018; February 17, 2019
Accepted: March 4, 2019
Published Online in Articles in Advance: August 28, 2019

https://doi.org/10.1287/mksc.2019.1170

Abstract. Many new products have several features that are not used by most consumers. One explanation for consumers buying products with features that they do not use is that consumers value features and ease of use differently at the point of purchase and at the time of usage. This behavior can be explained by consumer biases such as hyperbolic discounting. Some researchers claim that such consumer biases encourage firms to offer too many features at the cost of making products less user friendly. This paper examines how consumer biases such as hyperbolic discounting and naive expectations affect pricing, product design, firm profits, and welfare. We build a game theoretic model in which consumers need to invest in learning in order to use the additional features of a product. In our model, consumers use quasi-hyperbolic discounting, and some consumers have naive beliefs about their future learning behavior. Contrary to intuition, we show that hyperbolic discounting encourages firms to focus more on ease of learning rather than on investing in additional features. We also find that consumer biases do not always hurt consumers but can sometimes lead to improvement in consumer welfare. From a public policy perspective, our results suggest that educating consumers to reduce the number of overly optimistic consumers could be welfare reducing. We also show that firms can improve learning by making features complementary or by having a consistent design to encourage sequential learning of features. Our analysis also reveals that one approach to make consumers learn existing features may be to add more complementary features. We also investigate whether allowing consumers to choose between simple and feature-rich products would enhance welfare. We find that providing consumers with more choice can lead to reduced consumer welfare.

History: Ganesh Iyer served as the senior editor and Juanjuan Zhang served as associate editor for this article.
Supplemental Material: The online appendix is available at https://doi.org/10.1287/mksc.2019.1170.

Keywords: product design • pricing • behavioral economics • hyperbolic discounting • game theory

1. Introduction

Many new products such as software, car navigation systems, and cell phones include several features that are not used by most consumers. For example, a typical smartphone allows consumers to take photos, check emails, track appointments, play music, and perform hundreds of other functions. Several new refrigerators are internet connected and can send reminders about grocery purchases. Software, such as Microsoft Office, has hundreds of features, and the number of features keeps on increasing with new versions. This phenomenon is often referred to as *feature creep* in the literature (e.g., Elliott 2007). When a product has several features, it requires consumers to invest significant effort in learning how to use the product. For example, there are courses offered for consumers to learn how to use Microsoft Office, and manuals can be several hundred pages long. Consequently, most consumers end up using only the basic features. Researchers have suggested that these additional features decrease the usability of products and lead to frustration and lower customer satisfaction. For example, Thompson et al. (2005) suggest that firms should be aware that overloading products with features can hurt their profits in the long run. The literature on product design also warns against the tendency of designers to make feature-rich products and cautions against feature creep (e.g., Thompson et al. 2005, Rust et al. 2006, Elliott 2007). Despite these warnings, new products continue to add more features.

There are several reasons why a firm may offer feature-rich products. Feature-rich products can enable better price discrimination when consumers have varying preferences for different features. For example, Bakos and Brynjolfsson (1999) find that bundling of information goods can increase a monopolist's profits. Feature-rich products can also be beneficial when consumers' future preferences for features are uncertain. Zhang and Padmanabhan (2011) examine feature overload when features have option value and consumers are ex ante uncertain whether they will use multiple features. In their framework, firms trade off between usability and option value. They also show that feature-rich products can improve profits by achieving better price discrimination.

Although these streams of research provide potential explanations for adding many features, they are inconsistent with some observed consumer behaviors. For example, Thompson et. al. (2005) find that consumers' valuations for feature-rich products differ at the time of purchase and at the time of usage. Meyer et al. (2008) show that consumers incorrectly project their usage of additional features and consequently value a durable good more at the time of purchase than at the time of usage. Research has also shown that consumers may prefer low-feature products when they estimate their usage before purchasing (e.g., Hamilton and Thompson 2007, Goodman and Irmak 2013). These inconsistencies can be explained by the presence of consumer biases.

Meyer et. al (2008) propose and experimentally validate that hyperbolic discounting by consumers and incorrect expectations about learning behavior can explain why consumers prefer to buy feature-rich products. They also explain how such biases can result in consumers buying a product under the incorrect assumption that they will use many features. Researchers have suggested that firms may be tempted to take advantage of these biases by offering feature-rich products. They argue that this is bad for the firms in the long run because it will decrease consumer satisfaction (e.g., Rust et al. 2006). Some researchers have suggested that firms are forcing consumers to buy feature-rich products because these are the only available options, and therefore, there is an opportunity for simple products in the market (e.g., Darlin 2008, Pogue 2008, Hern 2015, Ludunke 2017). These results suggest that reducing bias such as overestimation of usage of additional features can potentially improve welfare and reduce feature creep.

Whereas prior consumer behavior research has examined the impact of consumer biases on consumer choice, it has not explored the strategic implications of these biases for firms. Firms can react to consumer biases by strategically adjusting prices, product features, or product usability. Intuition would suggest that biases that lead consumers to overestimate their chances of using the new features would result in feature creep. Similarly, intuition suggests that consumer biases would hurt welfare and that providing consumers with simple products would improve welfare. However, it is not clear that these intuitions will survive once we allow firms to strategically adjust pricing and product design to account for consumer biases. The purpose of this paper is to analyze these issues. Such an analysis is useful from both managerial and public policy perspectives. From a managerial perspective, the analysis can inform firms about optimal pricing and product design decisions. From a public policy perspective, such an analysis would help public policymakers assess the importance of educating consumers about making better decisions in terms of improving decision making by eliminating biases.

We develop a model in which a firm sells a feature-rich product. The base features can be used without any learning, but the additional features require learning by consumers. We assume that consumers use hyperbolic discounting. Hyperbolic discounting captures the phenomenon that consumers value additional features differently at the time of purchase and at the time of usage. In particular, at the time of usage, some consumers find the effort required to learn to use the product not to be worthwhile. Consequently, these consumers do not use the additional features. Consistent with prior research, we allow some consumers to rationally anticipate their future actions, whereas others are overly optimistic about their future learning behavior (e.g., O' Donoghue and Rabin 1999). In this context, we examine cases where some consumers buy feature-rich products under the incorrect assumption that they will use the additional features.

Our analysis reveals several interesting results. First, we show that, contrary to intuition, hyperbolic discounting does not lead to overinvestment in feature-rich product but rather encourages firms to focus more on making products easier to learn. In other words, if firms focus on increasing features rather than usability, this is not due to hyperbolic discounting but despite it. However, the presence of consumers with overly optimistic expectations can encourage a focus on features relative to learning costs. Furthermore, we find that hyperbolic discounting and the presence of naive consumers can force firms to lower prices, sometimes increasing overall consumer welfare. This result therefore clarifies that public policy initiatives educating consumers about their biases may have unintended consequences such as increased prices and reduced consumer and social welfare. Thus, unlike prior research, where awareness of one's bias can improve welfare, our results show the opposite (see, e.g., DellaVigna and Malmendier 2004, Heidhues and Kőszegi 2010). We also explore what firms can do to encourage learning of features. Our results show that firms can increase learning of existing features by adding more complementary features. These features could be more advanced features that need to be learned sequentially. Alternatively, firms can use common architecture and consistent design, which would also encourage learning. Thus, our results show that a solution to feature creep is not always to reduce the number of features but instead could be to strategically add new features that can bolster learning. Finally, we examine whether offering consumers a choice of both basic and feature-rich products will improve welfare. Our results show that when consumers can choose between both the basic and feature-rich products, consumer welfare can decline.

This paper adds to the growing literature in marketing and economics that has modeled the strategic impact of hyperbolic discounting (e.g., Laibson 1997, DellaVigna and Malmendier 2004, Gilpatric 2009, Jain 2012). This paper also adds to the literature on hyperbolic discounting in the context of durable goods. There is empirical evidence to suggest that individuals have discount rates for durables such as air conditioners, water heaters, freezers, and refrigerators that exceed market interest rates and are consistent with hyperbolic discounting (Hausman 1979, Gately 1980, Ruderman et al. 1987). However, hyperbolic discounting in a durable goods context has not been incorporated into formal models (for an exception, see Nocke and Peitz (2003)). Our paper incorporates hyperbolic discounting in a durable goods framework. This paper is more broadly related to the growing literature in marketing that has included psychological and sociological phenomena in formal models (see, e.g., Amaldoss and Jain 2005, Cui et al. 2007, Syam, Krishnamurthy, and Hess 2008, Lim 2010, Rao and Schaefer 2013). The remainder of this paper is organized as follows. In Section 2, we discuss the main model. In Section 3, we discuss several extensions of our model. Section 4 concludes the paper.

2. Model

We first consider the case of a monopolist offering a new product.[1] In Section 3.1, we consider competition. The sequential analysis allows us to understand the differential impacts of competition. The firm produces a product that has basic features that are easy to use and do not require any learning. In addition, the product also offers several extra features that require learning. For example, consumers could upgrade from a feature phone to a smartphone, which makes it easier to text and also offers augmented features that could require effort such as downloading apps and learning to use them. There is little learning required to use the basic features of a smartphone, such as calling and texting. However, in order to use advanced features, a consumer may need to invest time in learning those features. Similarly, basic calculations in spreadsheet programs or editing in word processors require minimal learning. However, these programs come with a variety of additional features that can require significant learning costs for the consumers.[2] We denote these learning costs by L. The product is durable, and the consumer uses the product for N periods. We assume that all consumers place a value u on the basic features. The parameter u is the net present value of the usage utility over the N periods. The sequence of decisions is as follows. In period 0, the firm decides on the price and product design. In period 1, the consumer decides whether to purchase the product. In period 2, the consumer decides whether to invest in learning the additional features. The consumer receives benefits from the features (base and additional, if learned) from period 3 to period $N+2$.

We assume that consumers have heterogeneous preferences for the additional features. In particular, we assume that an additional feature has net present valuation $\phi(\theta) = w - \theta$. The parameter θ is distributed according to a continuous concave distribution function $F(\cdot)$ with range $(0, 1)$.[3] Several distributions, such as the uniform distribution, left triangular distribution, and various families of the beta distribution, satisfy this assumption. The heterogeneity over the additional features is reasonable. For example, pivot tables may be of high value to a financial analyst or a student but of little value to casual users of Microsoft Excel. In addition to heterogeneity for the value of additional features, consumers could have heterogeneous learning costs. We analyze this in Section 3.5. It is also important to note that our framework with learning costs can also be interpreted in terms of the usability of new features. We discuss this later.

Prior research has shown that consumers rate the product features and the cost of learning differently at the point of purchase and the point of usage (see, e.g., Thompson et al. 2005, Meyer et al. 2008). In order to model this time inconsistency, we use the hyperbolic discounting framework, which has often been used to model self-control problems. Meyer et al. (2008) argue that because the benefits of a durable good are not realized immediately, consumers discount the future using a hyperbolic discount function. Their experimental results support this hypothesis. Following their approach, we assume that the consumer uses a quasi-hyperbolic discount function.[4] Quasi-hyperbolic discounting has been used extensively in the literature to model self-control issues (see, e.g., Laibson 1997, Gruber and Kőszegi 2001, Machado and Sinha 2007, Jain 2012). It represents the idea that some consumers discount the immediate future more than subsequent periods. We assume that the consumer uses the following function:

$$D(t) = \begin{cases} 1 & \text{if } t = 0, \\ \beta & \text{otherwise.} \end{cases} \tag{1}$$

Note that we are assuming that the exponential discount factor (for both the firm and the consumers) is 1.[5] This simplifies the presentation and is a common assumption in the literature (e.g., Gilpatric 2009).[6] It is easy to see from (1) that hyperbolic discounting would lead to time inconsistency. In other words, a consumer's preferences change depending on the time at which he or she makes the decision.[7] Note that this captures the idea that the weight a consumer places on features at the time of purchase is more than that at the time of usage. This is because w and L are weighted equally at the time of purchase, but at the time of usage, learning costs become more important, and the relative weight of features to learning costs is $\beta < 1$. This change in weighting over time for features versus learning costs is consistent with prior research (Meyer et al. 2008).[8]

It is important to note that this framework also admits an alternate interpretation of learning costs in terms of usability. Assume that the additional features provide delayed benefits of $\phi(\theta)$ but require immediate costs L to operate. As long as the benefits are obtained with delay and L is a constant, the analysis with this alternate interpretation remains the same. In this case, higher L implies that the product is less usable, and therefore, L can be interpreted as signifying complexity of usage.[9] Our framework could also allow for the possibility that w is uncertain and that consumers must incur learning costs L to determine whether the additional features are worth using (Meyer et al. 2008). In this alternative interpretation, w is the expected maximum benefit that a consumer could receive from the additional features. As long as there is a time lag between the time at which the consumer incurs the learning costs and the benefits that he or she gets from the additional features, the analysis remains the same.

Given that the preferences are time inconsistent, consumers need to predict whether they will invest in learning after purchase. There is substantial research that shows that some consumers overestimate their discount parameter (e.g., O' Donoghue and Rabin 1999). In order to model naive expectations, we assume that an η segment of consumers has rational expectations, whereas $(1-\eta)$ consumers have naive expectations.[10] We call the consumers with rational expectations *segment 1 consumers*. Also, we label the consumers with naive expectations as *segment 2 consumers*. We assume that consumers with naive expectations believe that their future selves have discount parameter $\hat{\beta} = 1$. Alternatively, we could assume that $\hat{\beta} > \beta$. In Section 3.6, we discuss an alternate formulation in which all consumers have expectations $\hat{\beta} \geq \beta$.

The consumers who have rational expectations expect to use the product iff $\beta(w-\theta) - L > 0$. Because we want to consider the case where at least some consumers would learn, we assume that $\beta w > L$. Define the marginal consumer who will use the product by θ_1. We have

$$\theta_1 = \frac{\beta w - L}{\beta}. \tag{2}$$

The firm could potentially offer the consumers pricing contract based on time of payment and also could offer a menu of contracts. However, because actual usage of features is difficult to observe, we assume that the firm cannot offer monetary rewards conditional on usage.[11] In the online appendix, we show that given these constraints, the optimal contract is for the firm to offer free financing such that the consumer signs the contract in period 1 but pays price p in period 2.[12]

If $\theta < \theta_1$, then this consumer will buy only if

$$U_2(\theta) = \beta(u + w - \theta - L - p) > 0, \tag{3}$$

which implies that the consumer will purchase if $\theta < \theta_2$, where

$$\theta_2 = u + w - L - p. \tag{4}$$

Thus, segment 1 consumers will buy only if $\theta \leq \min(\theta_1, \theta_2)$. Now consider the consumers with naive expectations. They will expect to learn as long as $\phi(\theta) - L > 0$. Let θ_3 be such that $\phi(\theta_3) - L = 0$. Therefore,

$$\theta_3 = w - L. \tag{5}$$

Because $\theta_3 > \theta_1$, some naive consumers in (θ_1, θ_3) will incorrectly believe that they will learn to use the features. If $\theta_1 < \theta_2$, then consumers with naive expectations in the set $(\theta_1, \min(\theta_2, \theta_3))$ will purchase the product and not use the additional features. Note that $\theta_2 < \theta_3$ if $p^* > u$. If this were not true, then all consumers would buy even if they did not anticipate using the additional features. For the firm to sell the feature-rich product, $p^* > u$ must hold because otherwise the firm could make more profits by selling the base product to everyone.

Using (2) and (4), it follows that $\theta_1 < \theta_2$ when

$$p \leq u + \frac{(1-\beta)L}{\beta} = \hat{p}. \tag{6}$$

Therefore, the demand function is given by

$$D(p) = \begin{cases} \eta F(\theta_1) + (1-\eta)F(\theta_2) & p < \hat{p}, \\ F(\theta_2) & \text{otherwise.} \end{cases} \tag{7}$$

To keep the analysis simple, we assume that the marginal cost for the product is zero. Therefore, the corresponding profit function is

$$\Pi(p) = \begin{cases} [\eta F(\theta_1) + (1-\eta)F(\theta_2)]p & p < \hat{p}, \quad (8a) \\ F(\theta_2)p & \text{otherwise.} \quad (8b) \end{cases}$$

Note that the profit function implies that for $p \geq \hat{p}$, the presence of naifs does not play a role in profits or consumer and social welfare analysis. Let us denote the price that maximizes (8a) by p_1* and the price that maximizes (8b) by p_2*. With this setup, we can derive the firm's optimal decisions.

We first derive the optimal prices. Note that the relevant profit function depends on whether the rational expectation constraint binds. In particular, the firm's profit function is different when $p \leq \hat{p}$ from the case when the firm charges a price higher than $\hat{p}$. It turns out that for some range of parameters, the optimal price is $\hat{p}$, at which the rationality constraint just binds, that is $\theta_1\theta_2$" is better as you suggested.

When consumers have low values of β (i.e., high levels of self-control problems), the firm charges a low price, and incorrect expectations of some consumers lead them to buy when they would not otherwise. In this case, the firm charges a price p_1*. However, for large values of β, consumer naiveté does not play a role. We therefore have the following.

Lemma 1. *The optimal prices are*[13]

$$p^* = \begin{cases} p_1^* & \text{if } \beta \leq \beta_1, \\ \hat{p} & \text{if } \beta \in (\beta_1, \beta_2), \\ p_2^* & \text{if } \beta \geq \beta_2. \end{cases} \tag{9}$$

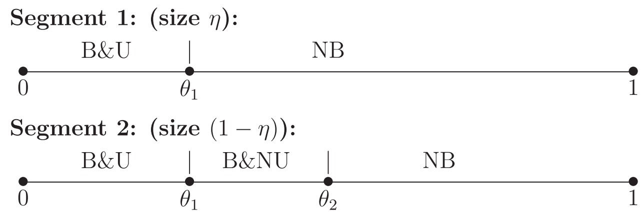
We could also state the lemma in terms of critical values of L or w. In particular, the case where some consumers buy under incorrect expectations will occur when learning costs are relatively high. For low values of learning costs, all consumers who buy will learn to use the features. Note that if $p^* \geq \hat{p}$, then all consumers who buy the product will learn to use the additional features. However, we are interested in situations in which some consumers incorrectly anticipate using the product while failing to do so. Therefore, we will primarily focus on the case when $\beta < \beta_1$.[14] Figure 1 depicts this situation. The profit function in this case is

$$\Pi = [\eta F(\theta_1) + (1-\eta)F(\theta_2)]p. \tag{10}$$

Proposition 1. *A firm charges a higher price as the hyperbolic discount parameter (β) and the proportion of sophisticated consumers (η) increase. However, the firm's profits increase with β but decrease with η.*

This proposition shows that a reduction of the self-control problem (i.e., an increase in β) and an increase in the proportion of sophisticated consumers lead to higher prices for consumers. The other part of the proposition shows that increases in the hyperbolic discounting parameter and the proportion of sophisticated consumers have differential effects on profits. In order to understand this result, let us first consider

Figure 1. Monopoly: Demand Pattern



Note. NB, Do not buy; B&U, buy and use additional features; B&NU, buy and do not use additional features.

the impact of β. Recall that an increase in β implies that consumer bias in terms of hyperbolic discounting is reduced. As β increases, consumers are more willing to learn to use the additional features. This increases the value of the product and consequently leads to higher prices by the monopolist.

Now let us consider the impact of η on prices. Intuition might suggest that as more consumers correctly anticipate that they will not learn to use the product, prices will decrease. In fact, the opposite is true. In other words, prices increase as more consumers have rational expectations. In order to understand this result, note that the marginal consumer in the segment with rational expectations is given by θ_2, which is independent of price. By contrast, the marginal consumers in the set of consumers who have naive expectations is denoted by θ_3, which is decreasing in price. Therefore, segment 2 consumers are more price sensitive than segment 1 consumers. As η increases, the size of segment 1 consumers increases, and therefore, the overall price sensitivity decreases, leading to higher prices.

The next part of the proposition shows that whereas increases in η and β lead to higher prices, they affect firm profits differently. Given the impact of β on prices, the impact on profits is intuitive. However, somewhat counterintuitively, profits decrease in η even though the firm is able to charge a higher price. To see this, note that although the demand from segment 2 consumers increases as η increases, an increase in η leads to fewer segment 2 consumers buying the product. Using the envelope theorem, the overall impact of η on the firm's profits is given by

$$\frac{\partial \Pi^*}{\partial \eta} = [F(\theta_1) - F(\theta_2^*)]p^* < 0. \tag{11}$$

Note that the result shows that the presence of more sophisticated consumers and consumers with more self-control can lead to higher prices but may not always increase the firm's profits.

A related question is whether a decrease in consumer biases is always beneficial to consumers and society. Most researchers view hyperbolic discounting as a bias that is harmful to consumers. The negative impact of hyperbolic discounting has been established in several other contexts, such as savings and health (e.g., Laibson 1997, DellaVigna and Malmendier 2004). Indeed, there is research that suggests how consumers could reduce their tendency to use hyperbolic discounting. For example, researchers have shown that strategies such as outcome elaboration can be effective in improving self-control, that is, increasing β (Haws et al. 2012). Research has also suggested that self-awareness of one's self-control problems can mitigate the negative impact of hyperbolic discounting (e.g., Heidhues and Kőszegi 2010). In other words, intuition would suggest that an increase in η and β should lead to higher consumer and social welfare. However, this intuition does not take into account the effect of the firm's strategic actions. We now analyze how β and η affect welfare while taking into account the firm's optimal pricing strategy.

In order to analyze the impact of hyperbolic discounting on consumer and social welfare, we need to address the issue of consumer preferences. Because hyperbolic discounting leads to changing preferences over time, the analysis of consumer and social welfare requires one to choose which consumer preferences one should use. The usual approach is to obtain the consumer surplus by looking at the consumer's long-run preferences by setting $\beta = 1$ (see, e.g., Gruber and Kőszegi 2001, O'Donoghue and Rabin 2003, DellaVigna and Malmendier 2004, Gilpatric 2009, Jain 2012).[15] Note, however, that the actual β and η will still impact social and consumer welfare because they affect demand and prices. Social welfare in this case is therefore given by

$$SW = \int_0^{\theta_2} (u + w - L - \theta)dF(\theta) + (1 - \eta) \int_{\theta_2}^{\theta_3^*} udF(\theta). \tag{12}$$

Note that the last term includes the consumers who end up buying the product under incorrect expectations but do not invest in learning. These consumers enjoy only the base features.[16] Consumer welfare is the difference between social welfare and firm profits.

We have the following result.

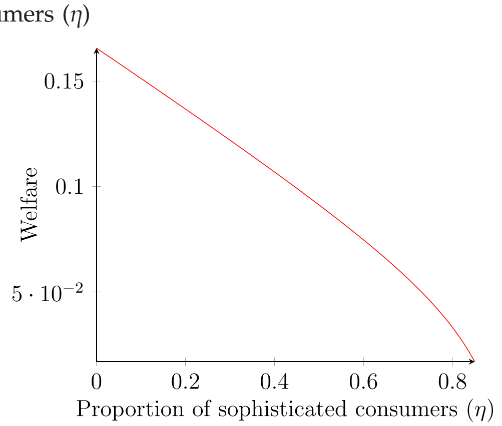
Proposition 2. *Social welfare decreases as the proportion of sophisticated consumers (η) increases but increases with the hyperbolic discount parameter β.*

The first part of this proposition shows that if more consumers become sophisticated, social welfare declines. This goes against the advice that consumers need to be educated to anticipate that they may not learn the additional features. Thus, from a public policy perspective, it is not clear that investment in programs to increase awareness of self-control problems is necessarily good for society. There are two reasons for this. First, an increase in η leads to higher prices, which decrease consumption. Furthermore, an increase in η decreases the set of consumers who buy a product that they use only partially. In particular, some consumers buy the product under the incorrect assumption that they will use the additional features but end up using only the basic features. Note that although this segment of consumers makes a negative surplus (because $u < p^*$), from a welfare perspective, an increase in this segment is still good for society. For example, consumers who upgrade to a smartphone will get some additional features over a basic cell phone and still get additional utility u from the upgrade but overpay in anticipation of use of other features. From a welfare perspective, this switch is still welfare enhancing but decreases consumer welfare.

The impact of η on consumer surplus is ambiguous because η reduces the number of consumers who overpay for a product when they do not learn the additional features. By contrast, an increase in η also increases prices and thus hurts consumers. When L is small (so the segment of consumers who overpay is small), consumer welfare could also decline with η because consumers pay higher prices (see Figure 2).

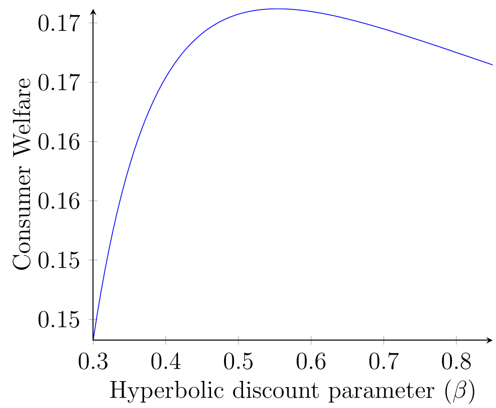
The second part of the proposition shows a dichotomy between η and β. We find that an increase in self-control (i.e., an increase in β) enhances social welfare. This is despite the fact that prices are higher in this case, and therefore, fewer segment 2 consumers buy. However, note that unlike η, an increase in β improves learning and therefore increases the set of consumers in segment 1 who buy and the set of consumers who learn in segment 2. The overall impact is an increase in social welfare. Whereas social welfare always improves with β, consumer surplus could still decline with β because of higher prices as a result of a higher β (see Figure 3). In other words, improved self-control can harm consumer welfare.

Figure 2. (Color online) Monopoly: Consumer Welfare Changes with the Proportion of Sophisticated Consumers (η)



Note. The parameters used for the figure are as follows: $w = 0.75$, $L = 0.3$, $u = 0.8$, $\beta = 0.5$, and $F(\cdot)$ is uniform.

Figure 3. (Color online) Monopoly: Consumer Welfare Changes with the Hyperbolic Discount Parameter β



Note. The parameters used for the figure are as follows: $w = 0.75$, $L = 0.15$, $u = 0.8$, $\eta = 0.3$, and $F(\cdot)$ is uniform.

Another important issue is whether firms have too much incentive to add new features. Suppose the firm can choose to decrease L or increase w. The firm could increase w by adding more features. The firm can reduce L either by changing the design or providing consumers with help in learning, for example, by providing tutorials, webinars etc. Assume that the cost to make a unit change in both parameters is the same and is denoted by $c(\cdot)$, where $c'(\cdot) > 0$ and $c''(\cdot) > 0$. The assumption of equal costs enables us to remove differential costs as an explanation for different incentives to invest in ease of learning or new features. The firm can choose to increase w by ω or decrease L by λ units at a cost of $c(\omega)$ or $c(\lambda)$. Note that we are assuming that w and L are independent. This allows us to study the impact of consumer biases on investments, independent of cost correlations. In Section 3.3, we consider the case when increasing features lead to higher learning costs or lower usability. We continue to assume that the firm still operates in the region in which some consumers buy under the incorrect assumption that they will learn to use the product. In this case, the profit function is given by

$$\Pi = \left[\eta F(\theta_1(w+\omega, L-\lambda)) + (1-\eta)F(\theta_2(w+\omega, L-\lambda))\right] \cdot p - c(\lambda) - c(\omega). \tag{13}$$

If the firm sets prices after making investments in ω and λ, using the envelope theorem, we have

$$\frac{\partial \Pi^*}{\partial \omega} = \eta f(\theta_1)p^* + (1-\eta)f(\theta_2^*)p^* - c'(\omega). \tag{14}$$

Similarly,

$$\frac{\partial \Pi^*}{\partial \lambda} = \frac{\eta f(\theta_1)p^*}{\beta} + (1-\eta)f(\theta_2^*)p^* - c'(\lambda). \tag{15}$$

We have the following result.

Proposition 3.

(a) *The firm invests more in reducing learning costs relative to investing in new features.*

(b) *If $f(\cdot)$ is uniform, then the ratio $\frac{c'(\lambda^*)}{c'(\omega^*)} = \frac{\eta+(1-\eta)\beta}{\beta} = \zeta_m > 1$. This ratio increases as the proportion of sophisticated consumers (η) increases and decreases as the hyperbolic discount parameter β increases.*

The first part of this proposition shows that, contrary to intuition, consumer biases such as hyperbolic discounting do not lead to feature creep. To see this, consider the case where there are no biases; that is, $\beta = 1$ and $\eta = 1$. In this case, $\frac{c'(\lambda^*)}{c'(\omega^*)} = 1 < \zeta_m$. Let us understand the intuition for the results. Note that for uniform $f(\cdot)$, (14) can be rewritten as

$$\frac{\partial \Pi^*}{\partial \omega} = p^* - c'(\omega). \tag{16}$$

Similarly, (15) becomes

$$\frac{\partial \Pi^*}{\partial \lambda} = p^*\left[\frac{\eta}{\beta} + (1-\eta)\right] - c'(\lambda). \tag{17}$$

It is easy to see that $c'(\lambda^*) > c'(\omega^*)$, and therefore, the firm has higher incentives to invest in making products easier to learn. Thus, the claim that the monopolist would invest more in features in the presence of hyperbolic discounting does not hold. It is also useful to compare how the relative investments change with β and η. The second part of the proposition shows that as consumer self-control improves (i.e., β increases), firms invest more in features relative to making products easier to use. Interestingly, as η increases (i.e., more consumers have rational expectations), the relative profitability of investing in w decreases. In other words, hyperbolic discounting and the presence of naifs have opposing impacts on incentives to make products easy to use. Whereas low β tends to favor products that are easy to use, naive expectations about future actions lead to more emphasis on adding new features at the expense of usability. *Therefore, contrary to intuition, firms' focus on improving features at the expense of increasing learning cost is not due to hyperbolic discounting but despite it.*

Taken together, the results of our model clarify the roles of two aspects that we study: hyperbolic discounting, which leads to overweighting benefits at the time of purchase, and the presence of naive consumers. As the tendency of consumers to overly weigh decreases (i.e., as β increases), prices, sales, firm profits, and social welfare increase, but consumer welfare could decrease. By contrast, as more consumers become better at foreseeing their future actions, prices increase, but sales, total profits, and social welfare decrease. Furthermore, our results show that a reduced tendency to overly weigh benefits (i.e., an increase in β) incentivizes firms to focus more on features rather than ease of learning. By contrast, an increase in the presence of sophisticated consumers leads firms to focus on easier-to-learn products.

3. Model Extensions

In the base model, we made several assumptions. In this section, we will relax some of these assumptions in order to examine the robustness of our results. First, we assumed that the firm is a monopolist. In Section 3.1, we consider a duopoly. This enables us to understand how consumer biases affect competition, which, in turn, affects prices and product design. In Section 3.2, we allow consumers to learn over multiple periods and allow for asymmetric benefits from features. Prior research has argued that an increase in features hurts usability, that is, increases L (e.g., Thompson et al. 2005). In our base model, we assumed that these two parameters are independent. In Section 3.3, we explore how the incentives to invest in ease of learning versus features change when we consider the case where w and L are positively correlated. In the base model, we assumed that the firm can only offer the feature-rich product and does not offer a product with only basic features. Although this assumption may be reasonable when there are fixed costs of maintaining separate product lines, it need not always hold. In Section 3.4, we relax this assumption and explore how allowing the firm to sell the base product along with the feature-rich product changes our results. Finally, in Section 3.5, we consider the case where consumers have heterogeneous learning costs. In Section 3.6, we consider the case where all consumers have $(\beta, \hat{\beta})$ preferences with $\beta \le \hat{\beta} \le 1$. Thus, all consumers are overly optimistic.

3.1. Competition

Now we examine the case where there are two firms in the market. Assume that firm 1 is located at 0 in the Hotelling line and firm 2 is located at 1. We assume that the firms are symmetric and that the distribution $f(\cdot)$ is a symmetric log-concave function with range $(0, 1)$. This includes the uniform distribution, several families of the beta distribution, and the truncated normal (Bagnoli and Bergstrom 2005). In order for expectations to have significance, we assume that the consumer in the middle of the Hotelling line does not learn; that is, $\beta w - L - \frac{1}{2} < 0$. As before, we will focus primarily on the case where some consumers buy the feature-rich product even though they end up not using the additional features. This is true as long as β is not too large.

First, consider segment 1. These consumers form rational expectations. Because, by assumption, $\beta w - L - \frac{1}{2} < 0$, the market is not fully covered in this segment. For rational expectations to matter, as before, we need that $p_i < \hat{p} = u + \frac{(1-\beta)}{\beta} L$, where $i \in \{1, 2\}$. In this case, consumers from $(0, \theta_1(p_1))$ will buy from firm 1. Analogously, consumers from $(1 - \theta_1(p_2), 1)$ will buy from firm 2. Finally, segment 2 consumers assume that they will invest in learning the additional features even when some of these consumers will end up not learning to use the additional features. We assume that $\beta(w - L - 1) > 0$ so that all consumers expect to learn to use the additional features. The demand in this case depends on whether the market is fully covered or not. If the market is not covered, then the marginal consumer is defined by θ_2. If the market is covered, then the marginal consumer is indifferent between the two products. Let θ_2^c be the marginal consumer, where

$$\theta_2^c = \frac{1 + p_2 - p_1}{2}. \tag{18}$$

If segment 2 is not fully covered, then the analysis reduces to the monopoly case, and therefore we focus on the case when segment 2 is fully covered. With this setup, the demand is as depicted in Figure 4, and the profit for firm 1 is given by

$$\Pi_1 = \left[\eta F(\theta_1) + (1 - \eta) F(\theta_2^c)\right] p_1. \tag{19}$$

The first-order condition implies that

$$\left[\eta F(\theta_1) + (1 - \eta) F(\theta_2^c)\right] - p_1 \left[\frac{(1 - \eta) f(\theta_1^c)}{2}\right] = 0. \tag{20}$$

Using symmetry, it follows that

$$p_1^* = \frac{\beta\left[2\eta F(\theta_1) + (1 - \eta)\right]}{(1 - \eta) f(\frac{1}{2})}. \tag{21}$$

As before, for the expectations to matter, we need that $\hat{p} > p_1^*$. This happens when β is not too large.

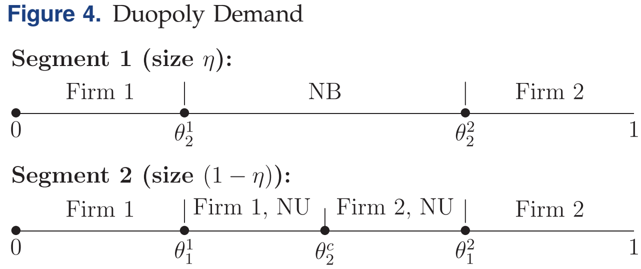
Proposition 4.

(a) *Prices increase in the proportion of sophisticated consumers (η) and the hyperbolic discount parameter β.*

(b) *Firms' profits increase in β. However, profits increase in η iff $\eta > \frac{1-4F(\theta_2)}{(1-2F(\theta_2))}$.*

(c) *Social welfare declines as the proportion of sophisticated consumers η increases and increases as the*

Figure 4. Duopoly Demand



Note. NB, do not buy; NU, do not use.

hyperbolic discount parameter β increases. If $\eta > \frac{1-4F(\theta_2)}{(1-2F(\theta_2))}$, then consumer surplus declines as η increases.

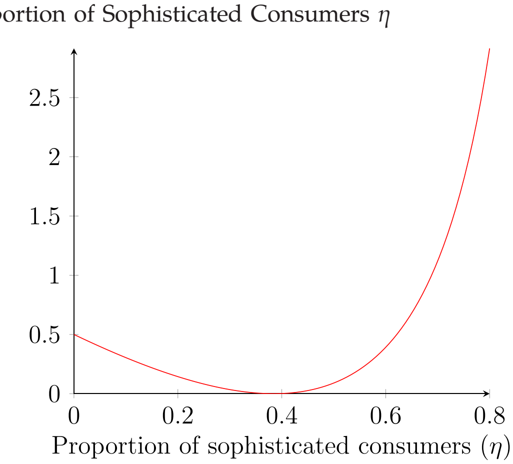
Let us first examine the impact of β on prices and profits. From (21), it follows that an increase in β leads to higher prices. Unlike the monopoly case, this increase in price does not affect the sales in segment 2. Therefore, firm profits and total sales increase as β increases.

Now let us consider the impact of η on the firm's prices and profits. The effect of η on prices is consistent with the monopoly result. Note, however, that in addition to the direct effect of η on prices, there is a strategic effect of an increase in η. In particular, because prices are strategic complements, an increase in η leads to reduced price competition and higher prices. However, we find that, unlike in the monopoly case, profits can increase in η. Recall that under monopoly an increase in η always leads to a reduction in profits. This is because as η increases, sales to consumers with naive expectations decline. In a competitive setting, the impact of change in η on firm 1's profits is

$$\frac{d\Pi_1^*}{d\eta} = \frac{\partial \Pi_1}{\partial \eta} + \frac{\partial \Pi_1}{\partial p_2} \cdot \frac{\partial p_2^*}{\partial \eta}. \tag{22}$$

The first term is the monopoly effect, which is negative. However, the second term refers to the strategic impact. We know that prices increase as η increases. Therefore, the second term is positive. It turns out that if η is large enough, the competitive effect dominates, and firms' profits improve with η.[17] Therefore, unlike the monopoly case, the firm's profits are U-shaped in η (see Figure 5). This condition implies that if most consumers are sophisticated, the firm would benefit if even more consumers become sophisticated. By contrast, if most consumers are naive, then the firm would be hurt if they became sophisticated. The intuition for the proposition is as follows. First, note that optimal prices are convex in η.[18] Therefore, for low values of η, the competitive effect is small, and the monopoly effect dominates. However, as η increases, the competitive effect dominates, and firms' profits increase as η increases.

Figure 5. (Color online) Duopoly Profits Change with the Proportion of Sophisticated Consumers η



Note. The parameters used for the figure are as follows: $w = 2$, $L = 0.7$, $\beta = 0.25$, and $F(\cdot)$ is uniform.

The last result shows that, as in the monopoly case, an increase in η leads to decline in social welfare, whereas an increase β increases welfare. The result also shows that the presence of naifs in the market can increase consumer welfare. This is because as η increases (i.e., as the proportion of naifs decreases), prices increase, which hurts welfare. However, an increase in η also leads to fewer consumers buying products that they use only partially. For large η, the first effect dominates. Note that the result also implies that for some parameters of η, social welfare, consumer welfare, and profits all could decline as η increases.[19] These results again reinforce our earlier result that an increase in the proportion of sophisticated consumers can hurt welfare.

Now consider the impact of changes in L or w on equilibrium profits. Let us consider the case where both firms independently choose the learning costs and features. Let us assume that both firms initially have learning cost L and that additional features are valued at w. Suppose that firm i can invest to increase w to $w + \omega_i$ and reduce L to $L - \lambda_i$. We assume that the cost function is the same for both and is given by $c(\cdot)$, where $c'(\cdot) > 0$ and $c''(\cdot) > 0$. The profit function for firm 1 is given by

$$\begin{aligned}\Pi_1 = {} & \eta F(\theta_1^i(w + \omega_1, L - \lambda_1))p_1 + (1-\eta)F(\theta_2^i(w + \omega_1, \\ & w + \omega_2, L - \lambda_1, L - \lambda_2))p_1 - c(\lambda_1) - c(\omega_1),\end{aligned} \tag{23}$$

where $\theta_1^i = \frac{\beta(w+\omega_i)-L+\lambda_i}{\beta}$ and $\theta_2^i = \frac{1-p_1+p_2+\omega_1+\lambda_1-\omega_2-\lambda_2}{2}$. The firms first choose (ω_i, λ_i) and then set prices. We solve the game backward. We have the following result.

Proposition 5.

(a) *Firms invest more in learning relative to features.*

(b) *If $f(\cdot)$ is uniform, then $\frac{c'(\lambda^*)}{c'(\omega^*)} = \frac{4\eta+\beta(1-\eta)}{(3\eta+1)\beta} = \zeta_c > \zeta_m$. This ratio decreases as the hyperbolic parameter β increases and with the proportion of sophisticated consumers η.*

The first part of this proposition again shows that even under competition, firms invest more in reducing learning costs. The second part of the proposition shows that under competition, firms focus more on ease of learning relative to the case of monopolist. In other words, contrary to intuition, competition does not lead to more focus on features at the expense of usability and in fact leads to more focus on ease of learning. To understand this result, first note that an increase in λ_1 increases valuation for segment 1 more than an increase in ω_1, whereas the effects are same for segment 2. Therefore, an increase in λ_1 is in general better for the firm. However, in a competitive setting, λ_1 and η_1 also change the nature of price competition. If firm 2 reacts to an increase in investments with a more aggressive pricing response, then the benefits of investments are reduced. This strategic effect is absent in the case of monopoly. Because λ_1 improves valuation more in segment 1 relative to an increase in ω_1, overall price competition is less intense with an increase in λ_1, and therefore, the strategic effect favors investment in ease of learning relative to features.[20] Consistent with monopoly, an increase in self-control (i.e., increased β) leads to more features and less focus on ease of learning. The last part of the proposition is also consistent with the base model and shows that improved self-control leads to more features; that is, self-control problems are not the reason for feature creep.

3.2. Infinite-Period Model with Asymmetric Features

Now we consider a more general model in which the consumer can use the product for several periods and also can choose to learn over several periods. In the base model, we assumed that the consumer learns in the period before usage and that this learning allows him or her to use all the features. The advantage of allowing consumers to learn over multiple periods is that it will allow us to more precisely capture procrastination in learning and, more important, show how firms can design products in such a way that this procrastination can be reduced and consumers can learn to use the additional features.

Consider the case where the product has two additional features, and the consumer can learn a single feature or both the features in any period. For simplicity, assume that the learning costs of both features are the same, but the benefits could differ. If both the benefits and learning costs are the same, then the main results of the base model will apply.[21] We consider the case where consumers use the product for infinite periods and can also invest in learning in each of the periods.[22]

We will assume that the two features differ in terms of w. This analysis will help us to study situations in which some consumers invest in learning one feature, whereas others learn both. The consumer receives utility $\nu + \sum_{i\in I} \omega_i(\theta)$ in each period that he or she uses the product after learning to use the additional features, where I is the set of features that the consumer learns. Otherwise, the consumer receives utility ν. In an infinite-period model, we need to have an exponential discount factor $\delta \in (0,1)$. Define $w(\theta) = \frac{\omega_i(\theta)}{1-\delta} = w_i - \theta$ and $u = \frac{\nu}{1-\delta}$. In each period, the consumer decides whether to invest in learning or procrastinates. If the consumer learns, then from period 1's perspective, he or she receives utility $\beta\delta w_i(\theta)$, which represents the discounted utility that he or she will receive from the next period. However, as before, the consumer incurs an immediate cost L for learning, where L is the learning cost for each feature. In order to characterize the equilibrium, we need to consider how consumers form expectations about the behavior of their future selves. The naive consumers always assume that their future selves will learn as long as $\delta w_i(\theta) - L > 0$. Note that the assumption here is that a consumer with low θ values both features more highly. Thus, we are assuming that the feature valuations are correlated. We later discuss how relaxing this assumption will affect our results. Naive consumers will prefer to defer by at least one period if

$$\beta\delta w_i(\theta) - L < \beta\delta[\delta w_i(\theta) - L], \tag{24}$$

which reduces to the condition that

$$\theta > w_i - \frac{L(1-\beta\delta)}{(1-\delta)\beta\delta} = \theta^i_{2n}. \tag{25}$$

Consumers with $\theta > \theta^i_{2n}$ will prefer to delay. Given their wrong expectations, these consumers will never learn feature i. It is useful to compare this with the earlier case where we allowed consumers to learn in only one period. The condition there would be $\theta > w - \frac{L}{\beta\delta}$. Thus, the multiperiod framework leads to less learning by naive consumers because they indefinitely procrastinate.[23] Without loss of generality, assume that $w_1 \le w_2$. This implies that $\theta^1_{2n} < \theta^2_{2n}$. Also, note that in this case, some consumers in the range $(\theta^1_{2n}, \theta^2_{2n})$ will learn only feature 2 if they buy the product.

However, these consumers expect to learn feature i if $\delta(w_i - \theta) - L > 0$. In this case, the critical value of consumers who expect to use feature i is given by

$$\delta(w - \theta^i_{3n}) - L_i > 0. \tag{26}$$

This reduces to the condition

$$\theta^i_{3n} = w_i - \frac{L}{\delta}. \tag{27}$$

Note that $\frac{L(1-\beta\delta)}{(1-\delta)\beta\delta} > \frac{L_1}{\delta}$. Thus, $\theta^i_{3n} > \theta^i_{2n}$, and some consumers who expect to learn feature i will not learn it. If the marginal consumer is such that he or she does not expect to learn both the features, then the firm will not introduce the other feature, and therefore, let us focus on the case where the marginal consumer expects to learn both features. As before, assume that the price is paid in period 2. If the consumer expects to learn both the features, he or she will buy if

$$u + \delta(w_1 + w_2 - 2\theta) - 2L - p > 0. \tag{28}$$

Note that the consumer can use the basic features immediately after purchase but must wait at least one period for learning the additional features. Also, the consumer will expect to learn both features in period 1, although he or she may end up not learning either. The marginal consumer is at θ_3, which is defined by

$$\theta_3 = \frac{w_1 + w_2}{2} - \left(\frac{2L + p - u}{2\delta}\right). \tag{29}$$

If, as before, we assume that the product can be learned by all consumers, absent hyperbolic discounting, the marginal consumer will be at θ_3.

3.2.1. Sophisticated Consumers. Now consider the sophisticated consumers. In order to characterize their behavior, we need to make additional assumptions about their beliefs. We follow O'Donoghue and Rabin (2001) and focus on a perception-perfect strategy. In such a strategy, the consumer optimizes in each period given his or her beliefs. Furthermore, the beliefs are constrained to be dynamically consistent such that a consumer's belief of his or her actions in period τ must be the same for all time periods $t < \tau$ [for a formal definition, see O'Donoghue and Rabin 2001)]. With this it can be shown that if learning feature i is worthwhile (i.e., $\beta\delta w_i(\theta) - L > 0$), then a consumer will delay learning by at most $d^*(\theta)$ periods, where

$$d^*(\theta) = \max_d \{d | \beta\delta(w_i - \theta) - L < \beta\delta^d(\delta(w_i - \theta) - L)\}. \tag{30}$$

There are thus several pure-strategy perception-perfect equilibria where the consumer learns in any period $\{0, 1, 2, \ldots, d^*(\theta)\}$. Furthermore, as long as $\beta\delta w_i(\theta) - L > 0$, learning immediately is always an equilibrium and is also the most efficient strategy from the consumer's long-term perspective. We will use this equilibrium selection criterion. Note, however, that this assumption plays a role only when the consumer's learning constraint does not bind. In such a case, when the consumer expects to learn affects his or her willingness to pay. Because we are focused on situations in which the learning constraint binds, the assumption is not critical but simplifies the presentation.

As long as $\beta\delta w_i(\theta) - L > 0$, this consumer will learn feature i. Because $w_2 \geq w_1$, some consumers will learn only feature 2. Define

$$\theta^i_{2s} = w_i - \frac{L_i}{\beta\delta} < w_i - \frac{L}{\delta}. \tag{31}$$

Note that $\theta^1_{2s} \leq \theta^2_{2s}$ because $w_2 \geq w_1$. In this case, the perception-perfect equilibrium would entail that consumers from $(0, \theta^1_{2s})$ will learn both features, whereas consumers in $(\theta^1_{2s}, \theta^2_{2s})$ will learn only feature 2. It is also interesting to note that $\theta^i_{2s} < \theta^i_{2n}$. In other words, naive consumers are less likely to learn features. This is because naive consumers assume that they will learn both features in the next period. Because these expectations are sometimes incorrect, some naive consumers in an infinite-period model end up perpetually procrastinating. It is possible that the marginal consumer in the sophisticated segment learns only one feature. Nevertheless, as long as the demand in segment 1 is determined by the learning constraint, an increase in η will lead to higher prices, as before. Also, the firm will have higher incentives to reduce learning costs.

3.2.2. Interrelated Features. Now consider the case where the features are interrelated. There are several reasons this could happen. Some features are designed to be intermediate and must be learned in order to learn the more advanced features. For example, before using applications in a cell phone, the consumer needs to learn how to install those apps. Manuals and help are often structured in order to encourage consumers to learn certain features before others. Alternatively, features could be complementary, and learning of one feature could increase the value of learning the other feature. For example, consumers with multiple files in Microsoft Excel can more efficiently use pivot tables if they know the VLOOKUP function. Another reason for features to be interrelated is that product features often share a common architecture, which leads to overall consistency. In this case, learning one feature makes it easier to learn the other. In our framework, this would imply that learning any feature reduces the cost of learning subsequent features. Although we will present the results in terms of benefits, the intuition remains the same when we formulate the problem in terms of learning costs. Finally, firms can make features interrelated by offering rewards, such as achievement awards, free in-app games, etc., for learning some features. We will start the model with the first interpretation, where the consumer either can independently learn both features or must learn them sequentially. Nevertheless, our framework is general enough to accommodate the other aspects above, and we will later discuss what this implies for product design.

Let us assume that the learning constraint binds for segment 1. If feature i needs to be learned before feature j, then we call feature i an *intermediate feature* and feature j an *advanced feature*. We have the following result.

Proposition 6. *If the firm can choose to design products with sequential or independent learning and* $w_2 > w_1 + \left(\frac{1-\beta}{\beta}\right)L$, *then in the sequential learning case, the firm will designate feature* 2 *as the advanced feature and feature* 1 *as the intermediate feature. Furthermore,*

(a) *If the exponential discount factor* δ *is close to* 1, *then the firm's profits weakly increase under the sequential case relative to the case where features can be learned independently.*

(b) *More consumers learn feature* 1 *under the sequential case compared with the independent case. More sophisticated consumers learn both feature* 1 *and feature* 2, *whereas fewer naive consumers learn feature* 2.

The result shows that the firm is better off making the more advanced features (which need to be learned before the intermediate features) of higher value. Furthermore, by making features interrelated, the firm can improve learning of all the features by sophisticated consumers. In order to understand this result, first note that under the simultaneous case, the sophisticated consumer will learn feature i only if $\theta \le w_i - \frac{L}{\beta\delta}$. If $\theta < w_1 - \frac{L}{\beta\delta}$, the decision of the sophisticated consumer is the same for independent and sequential learning because the consumer learns both features immediately. Recall that, by assumption, $w_2 > w_1$. If feature 2 is designated as the advanced feature, then a consumer at $\theta = w_2 - \frac{L}{\beta\delta} \equiv \theta^2_{2s}$ learns only feature 2 when he or she can learn both features independently. However, this changes when the consumer needs to learn feature 1 before feature 2. In particular, consumers at θ^2_{2s} will strictly prefer to learn both features as long as $w_2 > w_1 + \frac{1-\beta}{\beta}L$. To see this, note that these consumers will not learn both the features in period 1 but could consider learning feature 1 in period 1 and feature 2 in period 2. If they do this, then their expected utility is

$$U(\theta^2_{2s}) = \beta\delta(w_1 - \theta^2_{2s}) - L + \beta\delta(\delta(w_2 - \theta^2_{2s}) - L) \quad (32)$$

$$= -\beta\delta(w_2 - w_1) + \delta(1-\beta)L. \quad (33)$$

This is positive as long as $\frac{w_2 - w_1}{L} < \frac{1-\beta}{\beta}$. Therefore, in this case, all the consumers who learn only feature 2 in period 1 now learn both features. Note that if $\beta \to 1$, then there is no benefit to this sequential learning. Therefore, this result is driven by hyperbolic discounting. The firm can take advantage of consumers' hyperbolic discounting parameter to improve learning. To understand this result, note that consumers in period 1 expect that they will learn feature 2 in period 2. If they were to learn feature 1 alone, then their utility would be negative, and they would not do so. However, they receive a positive utility from period 1's perspective from learning in period 2. The utility of learning feature 2 for a consumer at θ^2_{2s} from period 1's perspective is positive, whereas it is zero from period 2's perspective, as can be seen from (32). This is because the period 1 consumer discounts both the learning cost and the benefits similarly, whereas the period 2 consumer weighs the learning cost more because of hyperbolic discounting. This is what makes it attractive for the period 1 consumer to learn feature 1, which he or she would not have done otherwise.[24]

Now consider the firm's demand. In the independent case, if the marginal consumer expects to use only feature 2, under the sequential case, he or she expects to use both. However, sequential learning will delay his or her learning of feature 2 by one period. As long as δ is close to 1, the willingness to pay will increase. This will lead to (weakly) increasing demand and profits.[25] Details are in the online appendix.

Now consider learning by naive consumers. These consumers expect to always learn in the next period. When these consumers need to learn features sequentially, they either learn both the features immediately or defer learning for both. In the latter case, naive consumers do not learn both the features. Consider the consumer who learns feature 1 immediately under independent learning. This consumer is located at $\theta < \theta^1_{2n}$, whereas the consumer who learns feature 2 immediately is located at $\theta < \theta^2_{2n}$. Recall that $\theta^2_{2n} > \theta^1_{2n}$. With sequential learning, consumers with $\theta \in (\theta^1_{2n}, \theta^2_{2n})$ cannot learn feature 2 before learning feature 1. Therefore, they have to choose whether to learn both features immediately or delay. They will learn immediately if

$$\beta\delta(w_1 - \theta + w_2 - \theta) - 2L \ge \beta\delta[\delta(w_1 - \theta + w_2 - \theta) - 2L], \quad (34)$$

which reduces to the condition that

$$\theta \le \frac{w_1 + w_2}{2} - \frac{(1-\beta\delta)L}{\beta\delta(1-\delta)} = \theta^a_{2n}. \quad (35)$$

It is easy to see that $\theta^1_{2n} < \theta^a_{2n} < \theta^2_{2n}$. Thus, more consumers who were learning only feature 2 now learn feature 1. However, fewer consumers who were learning feature 2 in the independent case now learn. As it turns out, the consumers in $(\theta^a_{2n}, \theta^2_{2n})$ indefinitely postpone learning. Overall, we see that although more naive consumers learn feature 1 in the sequential case, it comes at the cost of learning feature 2 for others.

It is also interesting to examine how the choice of features can affect the incentives for firms to invest in ease of learning versus features. The basic idea that firms will have more incentives to invest in learning remains valid even in this formulation. However, the firms' incentives are weakly higher for improving feature 2 rather than feature 1 in the independent feature case. When products are interrelated, because consumers learn both features, the firm will have the same incentives to invest in both features.[26]

Now let us examine the implications of our analysis for the alternate interpretations of interrelated features. First, consider the case where features are complementary. This could be due to higher benefits obtained from learning subsequent features or to reduced learning costs as a result of a common architecture or consistent design. To model this, we could consider a case where the consumer receives w for each feature when the features are independent and $w = w_1 < w_2$ when features are independent and feature 1 is learned before feature 2.[27] Clearly, more consumers will learn in this case, and profits and welfare will unambiguously increase. *Interestingly, this suggests that one way that firms can increase learning of features is to add complementary features.* This is in contrast to the common advice that feature creep can be tackled only by reducing features. From a design perspective, use of common architecture, consistent design, and complementary features can lead to improved learning, profits, and welfare. The result also suggests that manuals and help should be designed to encourage sequential learning. Similar arguments also show that firms can improve profits and welfare when they offer (noncash) rewards for consumers to learn both features. Recall that our analysis assumed perfect correlation between valuation of the two products. To see how the results would differ if we relaxed this assumption, let us consider the extreme case where they are perfectly negatively correlated. Thus, feature 1 has valuation $w_1 - \theta$ for a consumer at θ and $w_2 - 1 + \theta$. Then the total valuation for both features is $w_1 + w_2$ and is a constant. In this case, similar to the results in the bundling literature, we find that the firm can benefit if consumers learn the features sequentially rather than independently.

3.3. Learning Costs and w Are Correlated

It is often argued that adding additional features can lead to increased learning costs and reduced usability (Thompson et al. 2005). In our base model, we did not account for this possibility. Now we relax this assumption. Assume that learning costs are an increasing function of w. In particular, assume that $L = \psi(w)$, where $\psi' > 0$ and $\psi''(\cdot) > 0$. In other words, costs are an increasing and convex function of additional features. As before, let us assume that we are in a situation where some consumers buy even when they do not learn to use the additional features. In this case, the monopolist's profit function is given by

$$\Pi = \left[\eta F(\theta_1) + (1-\eta)F(\theta_2)\right]p. \tag{36}$$

In order to focus on the trade-off between additional features and usability, we set $c(w) = 0$. The firm will nevertheless choose an interior w as long as learning costs are sufficiently convex in w. Using the envelope theorem, it follows that

$$\frac{\partial \Pi^*}{\partial w} = p^*\left[\eta f(\theta_1^*)\left(1 - \frac{\psi'(w)}{\beta}\right) + (1-\eta)(1-\psi'(w))f(\theta_2^*)\right]. \tag{37}$$

From this, we see that if $f(\cdot)$ is uniform, then at the optimal w^*, we have

$$\psi'(w^*) = \frac{\beta}{\left[\beta + \eta(1-\beta)\right]} < 1. \tag{38}$$

Define $\hat{w}$ as $\psi'(\hat{w}) = 1$. Thus, the marginal utility of increasing w beyond $\hat{w}$ is negative. Absent self-control problems, the firm will choose $w^* = \psi^{-1}(1)$. However, in the presence of self-control problems, the firm chooses fewer features in order to make the product easier to learn. This is consistent with the result in the base model.

3.4. The Firm Can Offer Both the Base Product and a Feature-Rich Product

Now consider the case where the firm can offer not only the feature-rich product but also the base product. Some have argued that the practice of offering only feature-rich products is harmful to consumers and that consumers would be better off if they could buy simpler products (e.g., Hern 2015). In reality, this may not always be feasible because there may be fixed costs associated with carrying a larger product line. If we assume that these costs are low, then the firm can sell both products. Denote the price of the base product by p_b and that of the feature-rich product (with two features) by p_t. This can also be viewed as an unbundling strategy, where the firm offers consumers an upgrade option from the base product at price $p_t - p_b$. First, it is easy to see that the firm would charge a price of u for the base product.[28] If the firm continued to charge the optimal price, as in the single-product case, and charges u for the base product, no consumer would switch, but all nonbuyers would buy. The profits for the firm therefore would strictly increase. Of course, optimal prices would change, and the firm's profits would be even higher. To determine the optimal prices, we continue to assume that the firm operates in the region where some consumers buy the feature-rich product under overly optimistic expectations. The relevant profit function when the firm charges p_t in the two-product case is now

$$\Pi = (p_t - u)\left[\eta F(\theta_1) + (1-\eta)F(\theta_2^t)\right] + u, \tag{39}$$

where θ_2^t is the marginal consumer in the naive segment who buys the advanced product rather than the basic product. Let us denote the optimal price in the single-product case by p_s^* and that in the two-product case by p_t^*. In this case, it is immediate from (39) that $p_t^* > p_s^*$.[29] We can also establish that p_t^* increases in η and β. This is consistent with our prior results. Now let us examine whether introduction of the base product could change the incentives for investing in new features or ease of learning and also how the relative incentives change. We have the following result.

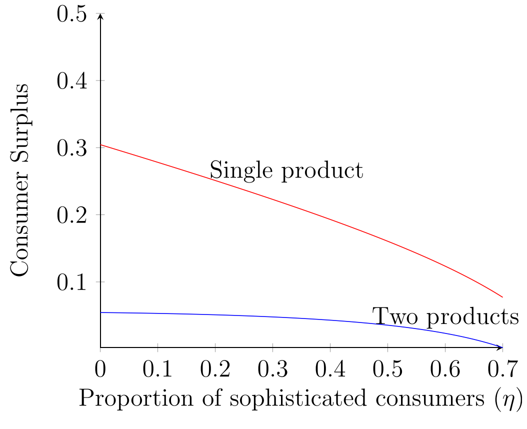
Proposition 7. *If $f(\cdot)$ is uniform, then we have the following:*

(a) *Profits and social welfare increase, but consumer surplus declines with the option of buying the base product.*

(b) *The firm benefits more by making products easier to learn rather than increasing product features. Firms invest less in ease of learning and new product features when they can sell both the base and the feature-rich product.*

It would generally be expected that offering more choices would make consumers better off. However, interestingly, this is not true, and for uniform $f(\cdot)$, consumer surplus is lower when the simpler product is available (see Figure 6). Because prices are higher when firm offers two products, it is easy to see that the consumers who continue to purchase the feature-rich product are worse off. However, there are some consumers who also switch down because of this increase in price. For uniform $f(\cdot)$, the first effect dominates, and consumer surplus declines. The social welfare, however, unambiguously improves as long as $\theta_1 < \theta_2$. There is, however, a caveat to this result. First, it assumes that it is costless to offer an additional product line. If there is a fixed cost associated with offering the base product, then social welfare could also be lower under the two-product case. In particular, because the firm has higher incentives to introduce the base product, for some values of fixed introduction costs, the firm may want to offer two versions, but it makes both society and consumers worse off. Another reason why social welfare could be negatively impacted by the introduction of two products is that offering the base product reduces innovation incentives, which therefore could reduce social surplus.

Figure 6. (Color online) Consumer Surplus: Two Products vs. a Single Product



Note. The parameters used for the figure are as follows: $w = 1$, $L = 0.4$, $u = 1$, $\beta = 0.6$, and $F(\cdot)$ is uniform.

The second part of the proposition shows that the condition under which investing in ease of learning is more beneficial relative to investing in features is the same as in the base case, where the firm offers only a single product. However, the ability of the firm to sell the base product along with the feature-rich product makes it less attractive for the firm to invest in making the feature-rich product easier to learn or adding more features. The reason is that marginal value from catering to the consumers buying the feature-rich product is lower. When the firm offers only a single product, an increase in w or a decrease in L brings in additional consumers whose marginal value is given by the price. However, when the firm sells the base product, all consumers buy, and therefore, a consumer who switches from buying the base product to buying the feature-rich product brings in a marginal revenue of $p_t^* - u$. For a uniform distribution, $p_t^* = p_s^* + \frac{u}{2}$. Therefore, $p_t^* - u < p_s^*$. In other words, the additional revenue from converting consumers is less than the revenue obtained in the single-product case. As a result, the firm's incentives to innovate are reduced when it can offer both the base and feature-rich product.

3.5. Consumers Have Heterogeneous Learning Costs

In the base model, we assumed that all consumers have the same learning costs. It is plausible that learning costs vary across consumers. For example, learning costs could depend on familiarity with the product category. Previous research has shown that experts have lower learning costs than nonexperts (e.g., Ziefle 2002). Learning costs could also differ because of age, education, and income. For example, young people often find it easier (and even enjoyable) to learn new features in technology products. Similarly, some consumers are tech savvy and can quickly and effortlessly learn new features. By contrast, it is more difficult for some consumers to learn new features. This could be due to lack of expertise, ability, or time. For example, high-income consumers may find it more costly to invest in learning new features. In order to model this, we assume that the learning costs for one section of consumers are zero. We label this segment of consumers *experts*. We assume the size of this segment of consumers to be α. The remaining $(1-\alpha)$ segment of consumers incurs a learning cost $L > 0$. The main results of this paper continue to hold in this alternate formulation. However, α does decrease the incentives for the firm to invest in learning. This is intuitive because learning costs are irrelevant for the experts.

3.6. Consumers Have Imperfect Expectations; That Is, $\hat{\beta} \in (\beta, 1)$

In the base model, we assumed that there are two segments: one with naive expectations and one where consumers have correct expectations. In our modeling of naive consumers, we assumed that these consumers believe that their future selves will have no self-control issues and $\hat{\beta} = 1$. An alternate approach is to use the $(\beta, \hat{\beta})$ formulation, where all consumers have incorrect expectations, denoted by $1 \geq \hat{\beta} \geq \beta$. In this case, the marginal consumer who will learn will be at $\hat{\theta}_1 = w - \frac{L}{\hat{\beta}} > w - \frac{L}{\beta} = \theta_1$. Therefore, the segment of consumers in $(\theta_1, \hat{\theta}_1)$ will incorrectly perceive that they will learn to use the additional features. Note that in such a formulation, $\hat{\beta}$ will play a role only when the price is sufficiently low that the learning constraint binds. This requires that $\theta_2 = u + w - L - p \geq \hat{\theta}_1$. However in any optimal solution, $\theta_2 > \hat{\theta}_1$ is not possible because the firm can then increase its price with no loss in sales. Therefore, the only case where $\hat{\beta}$ will matter in such a formulation is when $\hat{\theta}_1 = \theta_2$. This happens when u is sufficiently large.[30] The price in this scenario is given by $\hat{p} = u + \frac{L(1-\hat{\beta})}{\hat{\beta}}$. This price is decreasing in $\hat{\beta}$. In other words, as consumers' naiveté increases, prices go down. Also, it is easy to see that $\hat{\theta}_1$ is increasing in $\hat{\beta}$. Therefore, as consumers become more naive, social welfare goes up. Also, as before, the firm's incentives to invest in learning are higher in this alternate formulation. Thus, many of the key results continue to hold in this case.

4. Conclusion

The purpose of this paper is to examine how hyperbolic discounting and naive expectations affect a firm's pricing, profits, and product design; consumer surplus; and social welfare. Our results provide insights into the following questions:

1. *Does hyperbolic discounting and the presence of naive consumers lead to feature creep?* We find that hyperbolic discounting would encourage firms to make easier-to-learn products. Therefore, our result shows that feature creep is not due to the presence of consumer biases such as hyperbolic discounting. However, the presence of naifs can increase incentives to add more features at the cost of making products harder to learn.

2. *Would educating consumers about their biases improve welfare?* Our results show that improvement in self-control increases social welfare, but as more consumers become aware of their biases, social welfare declines. Furthermore, an increase in β and the number of sophisticated consumers can lead to a reduction in consumer welfare. Therefore, from a public policy perspective, investing in such education programs could be counterproductive.

3. *How does competition affect a firm's incentives to make products easier to use?* We find that competition can increase the incentives for firms to make products easier to learn.

4. *Can a choice of base and feature-rich products improve surplus?* Our results show that unbundling complex features and selling them as upgrades can improve profits for the monopolist but hurts consumer welfare.

5. *How can firms design products to improve learning and profits?* Our results suggest that if firms can encourage consumers to learn features sequentially, then more consumers could learn to use features, and profits could improve. In such cases, the firm will need to make the advanced feature relatively more valuable or easier to learn. This can be done via consistency in design, common architecture, incorporating complementarity in features, offering rewards to learn features in a sequence, or writing manuals to encourage sequential learning. Our results also show that sometimes addition of a new feature can increase learning of existing features as long as the learning is sequential. Thus, in contrast to the recommendations from the prior literature, this paper shows that the problems arising out of feature creep could be solved by adding more features.

In our analysis, we assumed that the firm cannot observe the actual usage of advanced features. With new technologies, it may be possible to monitor usage. This raises the possibility that the firm could charge based only on usage. In this case, the firm charges u for buying the base product and a separate price if additional features are used. It is easy to see that for a small u, the firm will not use this strategy. Furthermore, if the cost of carrying two product lines is low, then this strategy is dominated by the strategy of offering the base product and an upgrade. However, when u and the cost of introduction of two product lines are large, then unbundling such prices could be profitable. Another possibility is for firms to provide monetary awards to customers if they learn to use additional features. This would encourage learning, potentially increasing sales and welfare. It is easy to show that with the ability to provide monetary awards based on usage of advanced features, the firm can always make more profits and yet ensure that all consumers learn the additional features (see the online appendix for a proof). However, partial usage is still possible when the firm optimally sets prices and rewards.

Another issue that we did not consider is the availability of other commitment devices that could encourage learning. Consider the case where the firm can costlessly develop a commitment device that consumers can purchase, and consumers who purchase the commitment device will learn to use the additional features. Because hyperbolic discounting reduces profits in our framework, this would suggest that the firm could benefit from offering such a device. In the online appendix, we show that this is indeed the case, and the firm would in fact sell such commitment devices at zero prices. This would enhance profits and lead to higher prices. Nevertheless, social welfare improves.

In this paper, we focused on hyperbolic discounting and naive expectations and did not examine the impact of other biases such as projection bias on product design. The impacts of these biases on product design would be an interesting area for future research. Future research can also empirically test the implications of our theoretical model.

Acknowledgments
The author thanks Wilfred Amaldoss, Ashutosh Prasad, Upender Subramanian, seminar participants at Carnegie Mellon University, two anonymous reviewers, the associate editor, and the editor for their comments. The usual disclaimer applies.

Endnotes

[1] Alternatively, we could assume that the firm is offering an upgrade over an existing technology that is competitively supplied.

[2] Alternatively, we could assume that there are two types of features: some require minimal learning, and others require higher costs of learning. As long as features of the first type have a low cost of learning such that everyone learns, our results are the same.

[3] The assumption of concavity is not critical for several of our results. However, concavity does ensure that the optimal price is unique.

[4] Although we use the hyperbolic function as Meyer et al. (2008) do, our approach is a little different. In their framework, consumers must invest in learning to determine whether the new features are helpful and would increase the usage utility from the product. We assume that learning is required in order to use the additional features. Nevertheless, our framework can be easily adapted to consider this situation, as we discuss later.

[5] Because the exponential discount factor for the firm is 1, it is indifferent between payments at different points in time. We use a tie-breaking rule that in such cases the firm weakly prefers earlier to later payments.

[6] We relax this assumption in Section 3.2, where we also allow consumers to learn over multiple periods.

[7] To see this, consider a consumer with discounting as in (1). For this consumer, the discount rate between now and the next period is $\frac{1-\beta}{\beta} > 0$, whereas the discount rate between time period $t-1$ and t is given by 0.

[8] Although we use the quasi-hyperbolic discount function, this assumption is not critical. What we require is that consumers have time-inconsistent preferences and that some consumers are unable to foresee this. Alternate approaches with the same two key features would also lead to similar results. We use quasi-hyperbolic discounting because it is a parsimonious model with a single parameter. It also has been widely used in extant literature. Furthermore, hyperbolic discounting has wide empirical support.

[9] It is also possible that increased features could make usage of the base product more difficult. If the additional costs are L_0 and $\beta u - L_0 > 0$, then again our main results will hold.

[10] O'Donoghue and Rabin (1999) label the consumers with rational expectations as *sophisticates* and the consumers with incorrect expectations as *naifs*.

[11] In Section 3.2, we do consider the possibility of nonmonetary rewards.

[12] Consistent with prior literature, we assume that the firm cannot charge negative prices. Because, the firm has an exponential discount factor of 1, the firm could potentially offer an exploitative contract in which it pays the consumer p and expects a future payment of $\frac{p}{\beta}$. One way to rule this out is to add an additional stage in which the consumer decides whether to enter the store (see the online appendix). However, such exploitation is still possible when the firm can offer a menu of contracts and there are naive consumers who enter under the false assumption that they will choose the nonexploitive contract.

[13] Proofs of this and all other propositions are in the online appendix.

[14] If $p^* > \hat{p}$, then hyperbolic discounting has no impact, and the model reduces to one in which the consumer has no bias.

[15] Harris and Laibson (2002) justify this by asking the discount factor one would use to advise consumers. This approach is also consistent with the idea that hyperbolic discounting is a bias and not does represent the consumer's "true" preferences (Akerlof 1991).

[16] If we view L in terms of usability, we need to modify this equation slightly and replace u with $u - L_0$, as discussed in endnote 9. The main insights continue to hold.

[17] Note that the condition is always satisfied if $F(\theta_2) \in (\frac{1}{4}, \frac{1}{2})$.

[18] This can be seen immediately from (21).

[19] To see this, note that if $\eta = \frac{1-4F(\theta_2)}{(1-2F(\theta_2))} \equiv \eta_1$, then $\frac{d\Pi_1^*}{d\eta} = 0$, but at this point, $\frac{dSW}{d\eta} < 0$, and therefore, for some $\eta < \eta_1$, firm profits, consumer surplus, and social welfare decrease as η increases.

[20] The effect of price competition when we allow firms to coordinate investment levels is even stronger. In particular, when firms coordinate their investments, an investment in ease of learning reduces price competition even more relative to an investment in features.

[21] However, fewer consumers will learn when we consider a multiperiod model because there is a tendency to procrastinate. Nevertheless, the essential insights of the base model will hold when the costs and benefits of the two features are the same.

[22] Alternatively, we could allow consumers to use the product for a finite N. However, the analysis becomes more messy, but the essential insights are similar. We could also allow consumers to use the product for only two periods. This analysis leads to very similar results but could be considered limiting in that there could be strong end-effect bias. Therefore, we consider an infinite-period version.

[23] In a finite version of the version with N periods, these consumers will learn in period $N-1$.

[24] The idea is similar to the notion that assigning goals can improve performance [for an example in the context of sales force compensation design, see Jain (2012)].

[25] Demand may not change if the marginal consumer in the independent case is at θ_{2s}^2, although these consumers are still better off.

[26] Welfare comparison between the two cases is ambiguous because overall usage of features is ambiguous. Also, the prices could be higher or lower in the interrelated case compared with the independent case, depending on the set of parameters.

[27] Because what matters is that the second feature is more valuable or easier to learn, it is irrelevant which feature (i.e., feature 1 or feature 2) the consumer learns first.

[28] Suppose not. If the firm charges a higher price, then no one buys the base product, whereas if it offers a lower price, then each consumer who buys the base product makes a positive surplus (from the perspective of period in which he or she makes the purchase). The firm can then raise prices and make more profits.

[29] Formal proof is in the online appendix.

[30] The precise condition for this to hold is that $u > \frac{F(\hat{\theta}_1)}{f(\hat{\theta}_1)} - \frac{L(1-\hat{\beta})}{\hat{\beta}}$.

References

Akerlof G (1991) Procrastination and obedience. *Amer. Econom. Rev.* 81(2):1–19.

Amaldoss W, Jain S (2005) Pricing of conspicuous goods: A competitive analysis of social effects. *J. Marketing Res.* 42(1):30–42.

Bagnoli M, Bergstrom T (2005) Log-concave probability and its applications. *Econom. Theory* 26(2):445–469.

Bakos J, Brynjolfsson E (1999) Bundling information goods: Pricing, profits and efficiency. *Management Sci.* 45(12):1613–1630.

Cui T, Raju J, Zhang Z (2007) Fairness and channel coordination. *Management Sci.* 3(8):1303–1314.

Darlin D (2008) The year of the simpler gadget. *New York Times* (December 21), BU3.

DellaVigna S, Malmendier U (2004) Contract design and self-control: Theory and evidence. *Quart. J. Econom.* 129(2):353–402.

Elliott B (2007) Anything is possible: Managing feature creep in and innovation rich environment. *IEEE Internat. Engrg. Management Conf.* (Institute of Electrical and Electronics Engineers, Piscataway, NJ), 304–307.

Gately D (1980) Individual discount rates and the purchase and utilization of energy-using durables: Comment. *Bell J. Econom.* 11(1):373–374.

Gilpatric S (2009) Search for related content slippage in rebate programs and present-biased preferences. *Marketing Sci.* 28(2):229–238.

Goodman J, Irmak C (2013) Having versus consuming: Failure to estimate usage frequency makes consumers prefer multifeature products. *J. Marketing Res.* 50(1):44–54.

Gruber J, Kőszegi B (2001) Is addiction "rational"? Theory and evidence. *Quart. J. Econom.* 116(4):1261–1303.

Hamilton R, Thompson D (2007) Is there a substitute for direct experience? Comparing consumers' preferences after direct and indirect product experiences. *J. Consumer Res.* 34(4):546–555.

Harris C, Laibson D (2002) Hyperbolic discounting and consumption. Dewatripont M, Hansen LP, Turnovsky S, eds. *Advances in Economics and Econometrics: Theory and Applications,* Eighth World Congress, vol. 1 (Cambridge University Press, Cambridge, UK), 258–298.

Hausman J (1979) Individual discount rates and the purchase and utilization of energy-using durables. *Bell J. Econom.* 10(1):33–54.

Haws KL, Bearden WO, Nenkov GYJ (2012) Consumer spending self-control effectiveness and outcome elaboration prompts. *J. Acad. Marketing Sci.* 40(5):695–710.

Heidhues P, Kőszegi B (2010) Exploiting naïvete about self-control in the credit market. *Amer. Econom. Rev.* 100(5):2279–2303.

Hern A (2015) From coffee-makers to cameras: Six gadgets to simplify your life. *Guardian* (January 6), https://www.theguardian.com/technology/2015/jan/06/from-coffee-makers-to-cameras-six-gadgets-to-simplify-your-life.

Jain S (2012) Self-control and incentives: An analysis of multi-period quota plans. *Marketing Sci.* 31(5):855–869.

Laibson D (1997) Golden eggs and hyperbolic discounting. *Quart. J. Econom.* 112(2):443–477.

Lim N (2010) Social loss aversion and optimal contest design. *J. Marketing. Res.* 47(4):777–787.

Ludunke B (2017) Lessons learned from the failure of Ubuntu Touch. *Network World* (April 25), https://www.networkworld.com/article/3192311/lessons-learned-from-the-failure-of-ubuntu-touch.html.

Machado F, Sinha R (2007) Smoking cessation: A model of planned vs. actual behavior for time-inconsistent consumers. *Marketing Sci.* 26(6):834–850.

Meyer R, Zhao S, Han J (2008) Biases in valuation vs. usage of innovative product features. *Marketing Sci.* 27(6):1083–1096.

Nocke V, Peitz M (2003) Hyperbolic discounting and secondary markets. *Games Econom. Behav.* 44(1):77–97.

O'Donoghue T, Rabin M (1999) Doing it now or doing it later. *Amer. Econom. Rev.* 89(1):103–124.

O'Donoghue T, Rabin M (2001) Choice and procrastination. *Quart. J. Econom.* 116(1):121–160.

O'Donoghue T, Rabin M (2003) Self-awareness and self-control. Loewenstein G, Read D, Baumeister R, eds. *Time and Decision: Economic Perspectives on Intertemporal Choice* (Russell Sage Foundation, New York), 217–243.

Pogue D (2008) It's so simple, it's beautiful. *Internat. Herald Tribune* (March 20), 17.

Rao R, Schaefer R (2013) Conspicuous consumption and dynamic pricing. *Marketing Sci.* 32(5):786–804.

Ruderman H, Levine M, McMahon J (1987) The behavior of the market for energy efficiency in residential appliances including heating and cooling equipment. *Energy J.* 8(1):101–124.

Rust R, Thompson D, Hamilton R (2006) Defeating feature fatigue. *Harvard Bus. Rev.* 84(2):40–46.

Syam N, Krishnamurthy P, Hess J (2008) Models of miswanting and regret of customized products. *Marketing Sci.* 27(3):379–397.

Thompson DV, Hamilton RW, Rust RT (2005) Feature fatigue: When product capabilities become too much of a good thing. *J. Marketing Res.* 42(4):431–442.

Zhang K, Padmanabhan V (2011) Feature overload. Working Paper 2011/66/MKT, INSEAD, Fontainebleau, France.

Ziefle M (2002) The influence of user expertise and phone complexity on performance, ease of use and learnability of different mobile phones. *Behav. Inform. Tech.* 21(5):303–311.