

Service Quality Variability and Termination Behavior

S. Siriam

Stephen M. Ross School of Business
University of Michigan

Pradeep K. Chintagunta

University of Chicago

Puneet Manchanda

Stephen M. Ross School of Business
University of Michigan

Ross School of Business Working Paper

Working Paper No. 1224

January 2014

This work cannot be used without the author's permission.

This paper can be downloaded without charge from the
Social Sciences Research Network Electronic Paper Collection:
<http://ssrn.com/abstract=2267549>

Service Quality Variability and Termination Behavior

S Sriram*

Ross School of Business
University of Michigan

Pradeep Chintagunta
Booth School of Business
University of Chicago

Puneet Manchanda
Ross School of Business
University of Michigan

January 2014

*We thank seminar participants at Bocconi, Colorado State, Columbia, Delaware, Georgia Tech, Georgetown, Houston, Hebrew University, Illinois Institute of Technology, Kellogg, Johns Hopkins, LBS, Madison, NUS, Pittsburgh, Purdue, Stanford, Syracuse, Tilburg, Toronto, Univ of Illinois UC, University of Texas at Dallas, University of Utah, University of Washington, and Washington University, St. Louis. The first author acknowledges the financial support from 3M Corporation. The second author thanks the Kilts Center for Marketing at the Booth School of Business, University of Chicago for financial support. We would like to thank Sungjoon Nam for help with cleaning the data and an anonymous company for providing us the data. The authors are solely responsible for all errors and omissions.

Service Quality Variability and Termination Behavior

Abstract

While researchers have documented a positive relationship between the average quality of a service and customer retention, the effect of variability on customer retention has been viewed more ambiguously in the literature. We investigate the roles of the level and variability in quality in the context of a new video on demand service in driving customer retention. We find that while high average quality helps in retaining customers, high variability leads to higher termination rates. Apart from these main effects, we empirically document the presence of an interaction effect between average service quality and its variability on termination rates; customers who experience low variability are more responsive to mean quality compared to those experiencing high variability. As an extreme outcome, at the lower end of the quality spectrum, customers experiencing high variability have a higher retention rate than those experiencing low variability; this is contrary to what the main effect of variability would imply. We postulate a mechanism involving risk aversion and learning, which can induce this interaction effect and test this against several alternative explanations. Our results reinforce the notion that high service quality is associated with lower termination rates. Moreover, our estimates suggest that households exhibit risk aversion, implying that, on average, variability increases termination. Based on the model and estimation results, we document that in the context of new services where customers are likely to learn about their quality, households that experience low variability in service are likely to be more responsive to the quality level. This differential responsiveness results in an interaction effect between service quality level and its variability. In terms of managerial implications, we show that while increasing the average quality might be effective in retaining customers at low quality levels, lowering variability is likely to be more appropriate at high quality levels.

1 Introduction

When customers of a service experience low level of quality, on average, they are more likely to terminate the service compared to those experiencing higher average quality. This can be due to a variety of reasons, including disillusionment and/or inability to use the service (e.g., cellphone usage for a traveling sales person depends on the quality of the service in the places traveled). The importance of maintaining a high level of service quality in order to increase customer retention is echoed in the academic marketing literature (see e.g., Bolton, Lemon, and Bramlett (2006)).

Researchers in marketing have also investigated the effects of service quality variability on customer retention (see Kannan and Proenca (2010) for a review of this literature). Unlike the recommendation about the average service quality, the effect of variability on consumer retention has been viewed more ambiguously in the literature. The dominant view is that customers are likely to penalize a service for high variability, possibly because of “risk” or variability aversion, resulting in lower adoption (Meyer (1981)) and retention (Rust, Inman, Jia, and Zahorik (1999)). On the other hand, researchers have also argued that variability can increase retention if it induces a significant number of positive customer experiences (e.g., Bolton, Lemon, and Bramlett (2006)).

In this paper, we investigate the roles of the mean and variability in service quality in driving *customer retention* in the context of a new video on demand (VOD) service. We provide model-free evidence that the effect of mean quality on customer retention would depend on the variability of the service and vice versa. We postulate a mechanism involving risk aversion and learning (since the service is new) that can rationalize these data patterns. First, consistent with the literature, if customers are intrinsically risk averse, they would dislike variability. Second, in the context of a new service, customers tend to be uncertain about the average quality and learn about it over time (e.g., Mehta, Rajeev, and Srinivasan (2004), Narayanan, Manchanda, and Chintagunta (2005)). In the presence of uncertainty, high variability can deter learning about the true quality. As a result, customers experiencing high variability are less likely to update their beliefs about the true quality of the service. Consequently, termination rates would decrease with quality at a slower rate when variability is high. As an extreme outcome, households that have a higher prior belief about the quality of the service vis-à-vis what they actually experience are less likely to terminate in the presence of high variability. Thus, the key tradeoff is the following – high variability reduces the expected utility via risk aversion and thereby encourages termination. On the other hand, high variability can deter learning about

the true quality of the service. This, in turn, can lead to higher retention amongst customers with high prior quality beliefs who receive low quality. The tradeoff between risk aversion and learning deterrence is also alluded to in the experimental psychology literature (e.g., Pleskac (2008)). Our model-free data patterns are consistent with such a tradeoff.

The mechanism for the interaction effect that we propose is similar in spirit to Sun (2012), albeit with some key notable differences. In her paper, Sun (2012) argues that at the higher end of the average quality spectrum, high variability can lower *adoption*, possibly due to risk aversion. On the other hand, at lower average levels, high variability can lead to higher *adoption* by informing customers that it is a niche product/service. Meyer (1981) demonstrates such an interaction effect between average quality and variability, i.e., purchase intent increases more steeply with mean quality when variability is low. Since *retention* deals with consumers' own experiences of service quality *post-adoption* rather than the information contained in others' experiences *prior to adoption*, the "niche appeal" argument advanced by Sun (2012) for the interaction effect does not apply in our case.

We calibrate the model incorporating the roles of risk aversion and learning and demonstrate the differential effects of variability on customer retention in the context of a new video on demand (VOD) service. We exploit the notion that the cross-sectional difference in the average quality across households and the temporal quality variation within a household are determined exogenously. Given that we observe the quality experienced by each household over time and have information on the corresponding variability, we are able to identify the effects of risk aversion. Our results reinforce the notion that high service quality is associated with lower termination rates. Moreover, our estimates suggest that households exhibit risk aversion, suggesting that high-variability might increase termination rates, on average. Based on the model and estimation results, we document that in the context of new services where customers are likely to learn about the true quality over time, households that experience low variability in service are likely to be more responsive to the quality level. It is this differential responsiveness that results in the interaction effect between service quality level and its variability. Furthermore, we document conditions under which this interaction effect is likely to result in lower termination rates amongst customers experiencing high variability. Importantly, our model estimates can replicate the interaction effect of mean and variability in signal quality on termination as observed in our data; whereas high variability is associated with higher termination rates at high mean quality levels, it lowers termination at lower quality levels.

There are several managerially relevant implications of our analysis and findings. First, consistent with

the large body of research in the services marketing literature, our results show that for long-standing (i.e., not new) customers, a high level of average quality and a low level of variance drive increased retention. However, the implications are nuanced for new customers. Here we find that while increasing the average quality might be effective in retaining customers at low quality levels, lowering variability is likely to be more appropriate at high quality levels. This has implications for how firms (in industries beyond the VOD business under consideration) with limited resources might choose to allocate them. An example of this is the cellular phone service industry where firms face tradeoffs between offering a reliable service (more towers for coverage) and improving average quality by adopting new technologies (such as 4G LTE).¹

We also provide several extensions and robustness checks for our model and estimation results. First, the presence of learning may lead to households making intertemporal tradeoffs, i.e., be willing to continue with the service in order to learn about its quality (see e.g., Erdem and Keane (1996), Hitsch (2006)). We provide an extension of our model to account for such forward-looking behavior of households. Second, while the Bayesian model we specify for household learning assumes that the true quality and the quality “signals” for households come from a normal distribution, the signal quality data lie in a finite range. We therefore modify the learning model to account for this feature of the data to assess the sensitivity of our results. Third, we discuss an extension of our model that allows for household uncertainty regarding the variance of signal quality. In addition to these model extensions, we investigate several alternate model specifications. In particular, we look at a model that allows for the effects of extreme positive and negative signals received and one that incorporates the effects of the “gain” and “loss” associated with the most recent signal received by the household.

The rest of the paper is organized as follows. We first discuss the data used in our analysis and describe some key patterns. We then present the model, discuss its implications, and review the estimation approach. Next, we present our empirical results. We then provide results from model extensions and alternative specifications. Finally, we provide some concluding comments.

¹We elaborate on this later in the paper.

2 Data

The data used in this study come from a VOD service which was test marketed for over a year in three US cities: Jacksonville, FL, Salt Lake City, UT and Spokane, WA. Our data span the first 13 months of this test market from October 2003 to October 2004.² The service allowed subscribing households to rent movies from a database of 100 titles that reside in a set-top box with a built-in antenna. During each month, the movie database was to be updated by up to a maximum of 40 new titles via a terrestrial signal from a local broadcasting tower to the set-top box. The one-time activation fee was \$29.99, the monthly subscription fee for the service was \$7.99 and the average per-movie rental fee was \$1.99. Not paying the subscription fee in a given month resulted in termination of the service. Reactivating required re-payment of the activation fee. The data contain household level information on subscription and termination each month. The data used in our analysis consists of 3246 households who adopted the service of which 695 (i.e., 21.4%) terminated it during the period of our analysis. None of the terminating households re-adopted the service. Of the households in our sample, 44.1% were from Salt Lake City, 43.3% were from Jacksonville and the remaining 12.6% were from Spokane.³

A unique characteristic of the service is that while 40 movies can potentially be updated in its database every month, the actual outcome (the number of movies updated) depends on the signal quality that the household receives during that period.⁴ The movies are updated on a weekly basis with 10 being the maximum number of movies that can be downloaded in a given week. As new movies get downloaded, an equal number of the oldest movies (based on the time stamp at the time of downloading) get deleted from the set top box. If the household experiences low service quality in a given week, it would not receive all the scheduled movies; the movies that don't get downloaded in that week will never be available to the household.

The set top box records the signal quality received by each household during each period on a scale of 0-2.4, and our data contain this information.⁵ The data were transmitted from the set top box to the firm via

²The service was subsequently launched in a few other US cities.

³Nam, Manchanda, and Chintagunta (2010) use the same data to study the effect of word of mouth on customer acquisition. In contrast, the question in our paper pertains to how a customer's own experience with the service affects their retention conditional on having joined the service.

⁴Signal quality only affects the number of movies that are updated and not their picture or sound quality. In other words, if 30 movies are updated in a given month, all 30 movies will be of the expected quality.

⁵Conversations with the firm that provided these data revealed that a linear approximation of the signal quality on the 0-2.4 scale to the number of potential movies updated each month (i.e., 0-40) is reasonable.

the telephone line. A survey of 8% of subscribers showed that the number of movies received (a) was the most important factor affecting satisfaction with the service; and (b) explained over 50% of the variation in satisfaction across subscribers. This implies that signal quality is an important determinant of the overall quality of the service. Furthermore, the service quality literature, (e.g., Parasuraman, Zeithaml, and Berry (1988)) has discussed the differences between objective quality of a service and customer perception. The strong correlation between the satisfaction scores and the signal quality as recorded in set-top box suggest that this objective quality measure is a reasonable proxy for the perception of service quality. Therefore, for the remainder of the paper, we use the terms “quality,” “service quality” and “signal quality” interchangeably.

Given that the signals are transmitted terrestrially, the average signal quality varies across households depending on where they are located. To illustrate this point, we present the distribution of the average signal quality received by households in Figure 1. The figure suggests that there is considerable heterogeneity in signal quality across households, with some not getting any new movies during their tenure (i.e., average signal quality of 0). The average quality received by households was 1.415, while the median was 1.492. Furthermore, within a household, the signal quality could vary over time depending on factors such as the weather, changes in the orientation of the antenna atop the set-top box, etc. - the average variance experienced by households was 0.447. However, there is significant heterogeneity amongst households in the extent to which they experience variability in signal quality. In Figure 2, we provide the distribution of within-household variability in signal quality. Based on these descriptive analyses, we conclude that there is significant heterogeneity across households in average signal quality and its variability over time.

Several aspects of the data make it well suited for our research. First, given that the set top box records the signal quality received by households during each period, we are able to observe an important component of service quality over time. This helps us understand how the mean service quality and its variability are related to customer tenure. Second, the cross-sectional and temporal variation in signal quality is induced by factors such as household’s location and weather conditions, respectively. As a result, the household cannot systematically alter the quality it receives. Further, while the firm can make investments to improve the average quality across all households, it does not do so during the period of our data. Even if the firm makes such an investment, it cannot influence the level of quality received by a particular household in the market. Together, these factors allow us to treat signal quality as being exogenous to the subscriber. This feature of the data allows us to measure the effects of quality without having to control for endogeneity or

selection issues. Finally, the data contain information from the time of service activation for all households. This enables us to investigate how customers' beliefs about the service evolve over time as they obtain more information.

2.1 Data Patterns

In this section, we provide model-free evidence on how the mean and variance of quality influence retention rates. First, we discuss the marginal effects of mean and variability on termination and then consider the joint effects. Next, we document evidence that households learn about the quality of the service over time. We also present evidence that the rate of learning is related to the variance in signal quality experienced by these households.

2.1.1 Marginal and Joint Effects of Average Service Quality and Variability

It has been documented in the literature (see Bolton, Lemon, and Bramlett (2006)) that service termination (continuation) rates decrease (increase) with quality. In our data, the correlation between the binary outcome of termination (1) / continuation (0) for each household and that household's average signal quality over its tenure with the service is -0.29 and is statistically significant at the 5% level. In other words, households that, on average, receive higher signal quality, terminate less than those receiving lower quality. As an alternative illustration of this relationship, we find that the average termination rates among households experiencing higher than median quality is lower than those experiencing low quality (11.4% vs. 31.4%, significant at the 5% level), determined based on a median split.

Next, we investigate the association between termination rates for households and the extent of variability in signal quality experienced by them. Once again, this is a cross-sectional association in which we correlate the binary household-level outcomes of termination/continuation with households' signal quality variances over their durations of service. Our data indicate a positive correlation of 0.09, which is statistically significantly different from 0 at the 5% level of significance. This shows that households experiencing higher variability in their service are more likely to terminate than those with low variability. In addition, we compare the average termination rates for households experiencing above versus below median signal

quality variance. These results show that households with high signal quality variability tend to terminate more than those with low variability (18.5% vs. 24.3%, significant at the 5% level). This relationship is consistent with aversion to variance or more generally, risk aversion.

To understand whether the effect of variability in service quality differs across average quality levels, we compare the termination rates at high and low levels of both mean and variance in signal quality. The results in Table 1 suggest an interaction effect of mean and variance - high variability is associated with higher termination rates at high mean quality levels, it lowers termination at lower quality levels. A potential explanation for this pattern stems from the quality of the service being bounded from above and below. As a result, both low- and high-variability customers at the lower end of the quality spectrum are likely to have similar realizations of low quality outcomes. On the other hand, those with higher variability have more upside potential. Similarly, at higher quality levels, high-variability households have a higher probability of experiencing low quality outcomes than their low variability counterparts. Hence, high-variability households are likely to receive more (fewer) movies than their low-variability counterparts at lower (higher) average quality levels. If this is true, the differences in average quality across households experiencing low- versus high-variability can explain the outcome in Table 1. To evaluate the validity of this explanation, we need to compare the termination rates *after* controlling for the average quality experienced. In Table 2, we present the termination rates among low- and high-variability households on a fine grid of average quality. Comparing across the high and low variability levels at high average quality levels, we find that high variability leads to higher termination rates than low variability. However, comparing across variability levels at low average quality levels, we find the opposite, i.e., higher variability leads to lower termination rates at low average quality levels. In other words, high variability can increase termination rates at high quality levels but helps in retaining customers who experience low average quality.⁶ Thus, the mechanism by which higher variability lowers termination requires further investigation. Furthermore, note that households experiencing high variability stay with the service longer prior to termination compared to low-variability households. As we discuss subsequently, this pattern is suggestive of high variability deterring the ability of households to learn about the quality of the service that they receive.

To establish the above interaction effect more formally, we provide the results from a log-logistic hazard

⁶Note that the results in Table 2 just rule out the (obvious) explanation that the differential termination rate is due to differences in average quality. As the downside at the lower end of the quality spectrum is limited, households experiencing high variability are likely to receive higher signal quality, on average. However, these results do not suggest that the bounded nature of signal quality is not an important aspect to consider in our model formulation. We formally account for the bounded nature of the data generating process in a later section.

model in Table 3. The dependent variable in this analysis is the time to termination for each household and accounts for right censoring, i.e., some households might not have terminated the service by the end of the horizon. We include the average signal quality received, quality variance, and interaction between mean and variance as the key covariates in this analysis. Further, we control for heterogeneity by including household demographics as well as dummy variables for the month in which the household activates the service. The results reveal that all 3 covariates have statistically significant effects on time to termination. The results once again point to the presence of a significant interaction effect. Note that due to the presence of the interaction effect, we need to be careful in interpreting the coefficients. The negative main effect for variability implies that when the mean quality is low, variability lowers the termination hazard. On the other hand, when the mean quality exceeds 1.47 (i.e., $0.28/0.19$), variability increases the termination hazard. This is consistent with the data patterns reported in Tables 1 and 2.

2.1.2 Evidence of Learning

While the above analysis considers the cross-sectional variation across households in our sample, it does not reflect how termination behavior varies based on the history of realized signal quality. If households are uncertain about the quality they receive, on average, the temporal variation in signal quality within a household would imply that a single realization is not sufficient to fully resolve this uncertainty. Rather, each realization provides a noisy signal of the true average quality being received by the household. Under these circumstances, households could learn over time about the true quality of the service they receive. In Figure 3, we plot the proportion of subscribers who terminate the service after receiving “ n ” periods of high quality signals or “ n ” periods of low quality signals for various values of “ n .” Here a “high” or “low” quality signal is defined as one that falls above or below (respectively) of the median signal quality. The median is computed across signals received by all households over the data period. The figure shows higher termination rates amongst households receiving bad quality signals, with the effect becoming more pronounced with the number of periods for which the quality is low. At the same time, termination rates decline slightly but do not appear to change much when households receive many periods of good signals. Together, these patterns seem to indicate that households learn about the quality they receive from the service over time.

Next, we look at how the patterns in Figure 3 vary by households subject to low versus high variability.

These are depicted in Figure 4. This figure suggests that households that experience low variability and receive low quality signals over several time periods tend to terminate in higher proportions than households receiving high quality signals (i.e., there is limited overlap in their 95% confidence bands). On the other hand, there is no significant difference in termination rates based on the history of receiving low or high quality signals amongst households experiencing high variability (i.e., there is significant overlap in their 95% confidence bands). Together, these results suggest that the rate of learning is related to the variability in signal quality experienced by households. We conjecture that while low variability enables households to discern the true quality of the service, high variability renders such an inference difficult. In a Bayesian learning sense, this is likely to arise because households that experience high variability are likely to place a greater weight on the prior belief rather than on the signal.⁷ Note that this inference is consistent with our earlier observation in Table 2 that households experiencing high variability stay with the service longer prior to termination compared to low-variability households.

In summary, our data patterns suggest that (a) termination decreases with average quality, (b) termination increases with variability, (c) there is an interaction effect between the average quality and variability, (d) customers learn about quality over time, and (e) the learning is hindered for households that experience high variability. While (a) - (c) underscore the importance of both mean and variance, (d) points to uncertainty among households in the average quality they receive, and (e) probably suggests that the rate of learning is related to the variability that consumers experience. In the next section, we propose a model involving risk aversion and learning that reflects these factors. Subsequently, we will discuss alternative mechanisms that could rationalize our observed data patterns and compare them with our proposed mechanisms.

3 Model

3.1 Overview

We characterize the indirect utility that household h derives from subscribing to the service at time t as

⁷A more parsimonious account of the data that does not require Bayesian learning, is that consumers use the information on the trajectory or slope of the most recent quality signals. While not reported here, we compared such models with the formal Bayesian learning model described later. The Bayesian learning model yielded superior model fit, even after accounting for the additional parameters that needed to be estimated. We thank Dan Bartels and an anonymous reviewer for suggesting this alternative explanation.

$$\tilde{V}_{1ht} = \alpha_h + g(\tilde{Q}_{ht}) + \psi D_h + \tau S_t + \varepsilon_{1ht}, \quad (1)$$

where, α_h is the household's intrinsic preference for the service, \tilde{Q}_{ht} is the household's belief about the quality of the service at time t , $g(\cdot)$ is a function that captures how quality influences the utility from subscribing, D_h are household-specific demographic characteristics that shifts the intrinsic preference for the service, and S_t are month dummies that capture seasonality in preference for the service. The term ε_{1ht} is an idiosyncratic shock experienced by household h at time t .⁸

A household's intrinsic preference for the service, α_h , and their responsiveness to average signal quality, β_h are likely to be functions of the extent to which the household uses the service.⁹ The rationale is that households that rent a lot of movies are likely to derive more consumption value from the service and thus find it more valuable. Similarly, high usage households are also likely to be more sensitive to the quality of the service. We accommodate this feature in our estimation described later.

Since the household makes the subscription decision at the beginning of each period, it is uncertain about the quality of the service it will receive over the course of the month and whether it is worth the subscription fee. As discussed in the data section, there is significant heterogeneity across households in terms of the average quality of the service that they receive. Therefore, it is conceivable that a household that is new to the service is uncertain about the average quality of the service that it would receive. Furthermore, there is temporal variation in quality experienced by a given household over time due to weather conditions. Therefore the household is likely to be uncertain about (a) the average quality of the service it experiences and (b) the actual quality it will experience during period t for which it is making the subscription decision. Hence, the household's decision is based on the expected utility from subscribing, where the expectation is taken over the distribution of uncertainty. Formally, the expected utility that the household derives from subscribing to the service at time t is:

$$\bar{V}_{1ht} = E[\tilde{V}_{1ht}] = \alpha_h + E[g(\tilde{Q}_{ht})] + \psi D_h + \tau S_t + \varepsilon_{1ht}. \quad (2)$$

⁸Technically, the utility is a function of the monthly subscription price. However, since there is no variation in the subscription price in the data, we cannot econometrically identify the price effect. As a result, the price effect is subsumed within the household's intrinsic preference.

⁹We thank the AE for pointing this out.

We assume that at time t , the uncertainty about the household's true quality is distributed $N(Q_{ht}, \sigma_{Q_{ht}}^2)$ and the temporal uncertainty is distributed $N(0, \sigma_h^2)$. Therefore, the total uncertainty perceived by the household about the quality of the service at time t is the sum of the household's perceived uncertainty about the average quality of the service it receives ($\sigma_{Q_{ht}}^2$) and the temporal variation it experiences (σ_h^2), i.e., $\sigma_{ht}^2 = \sigma_{Q_{ht}}^2 + \sigma_h^2$. Over time, the household accumulates experience with the service and updates its beliefs accordingly. In what follows, we assume that households are Bayesian learners. Consequently, the mean converges to the true average quality received by the household and the uncertainty ($\sigma_{Q_{ht}}^2$) converges to 0. However, temporal variation in signal quality experienced (σ_h^2) will persist.

We normalize the observed (by the researcher) component of the per-period utility from not subscribing to the service to 0. Therefore, the total per-period utility from terminating the service, including the idiosyncratic shock is

$$V_{0ht} = \varepsilon_{0ht}. \quad (3)$$

We assume that the household-specific idiosyncratic shocks, ε , follow a type I extreme value distribution. The corresponding probability of subscribing is:

$$Prob_{ht} = \frac{\exp(\bar{V}_{1ht})}{1 + \exp(\bar{V}_{1ht})}. \quad (4)$$

We recognize that since termination is, for the most part, an irreversible process, the household is likely to consider more than the current period utility while making the subscription decision. Later, we consider an extension that accounts for this forward-looking behavior.

3.2 Risk Aversion

The expression for the expected utility will depend on the specification of the $g(\cdot)$ function. If $g(\cdot)$ is specified as linear function of \tilde{Q}_{ht} , the expected utility will also be a linear function of the household's perception about the quality of the service, \bar{Q}_{ht} . Furthermore, the expected utility will be invariant to the uncertainty about service quality. As discussed in the previous section, our data suggest that households are

averse to variability, on average. In order to capture this, we use a quadratic specification for $g(\cdot)$. We use the quadratic specification because researchers have used it in the past to effectively capture risk aversion (see, for example, Erdem and Keane (1996), Ching (2010)). Furthermore, it is computationally easier to estimate a model with quadratic utility as opposed to other functional forms such as constant absolute risk aversion (CARA) (e.,g., Chan and Hamilton (2006), Ching and Ishihara (2010), Crawford and Shum (2005)), especially for the model that also accounts for forward looking behavior (described later).¹⁰ The corresponding expression for the expected utility in Equation 2 is

$$\bar{V}_{1ht} = E[\tilde{V}_{1ht}] = \alpha_h + \beta_h \bar{Q}_{ht} - \gamma \bar{Q}_{ht}^2 - \gamma \sigma_{ht}^2 + \psi D_h + \tau S_t + \varepsilon_{1ht}. \quad (5)$$

Recall that $\sigma_{ht}^2 = \sigma_{Q_{ht}}^2 + \sigma_h^2$ is the total uncertainty perceived by household h regarding the quality of the service it would experience during period t .

3.3 Learning

We assume that households update their beliefs about the true quality of the service in a Bayesian manner. A household has a prior belief about the quality of the service at the time of activation. We assume that this prior belief follows a normal distribution such that $\tilde{Q}_{h0} \sim N(Q_0, \sigma_{Q_{h0}}^2)$. (The prior belief also needs to be high enough that the household activates the service in the first place) During each period t , household h notes the quality of the service it receives (via the number of new movies in its set-top box) and updates its belief accordingly. As is common in the literature (see, for example Coscelli and Shum (2004)), we assume that the quality of the signals that the household receives each period comes from a normal distribution with variance σ_h^2 . In contrast to a majority of empirical studies, however, we as researchers, observe the actual signal quality received by each household in each period.

Given the conjugacy of the prior and signal, the posterior mean belief about the signal quality after t periods of subscribing (i.e., the prior at the beginning of period $t + 1$) based on the information it has accumulated till time t can be written as

¹⁰While not reported here (available from the authors), results from the CARA assumption are very similar to those from the quadratic model when there is no forward-looking behavior.

$$\bar{Q}_{ht} = \bar{Q}_{ht-1} + \frac{\sigma_{Q_{ht}}^2}{\sigma_{Q_{ht}}^2 + \sigma_h^2} [Q_{ht} - \bar{Q}_{ht-1}], \quad (6)$$

where, Q_{ht} is the actual signal quality experienced by the household during period t and the posterior variance at the end of the period t (i.e., prior variance in period $t+1$) is

$$\sigma_{Q_{ht}}^2 = \frac{1}{\frac{1}{\sigma_{Q_{ht-1}}^2} + \frac{1}{\sigma_h^2}} = \frac{1}{\frac{1}{\sigma_0^2} + \frac{\tau_{ht}}{\sigma_h^2}}. \quad (7)$$

In the above expression, τ_{ht} is the number of periods that household h has been with the service as of period t . Since households join the service at different time periods, the number of periods that household h is with the service (τ_h) would not correspond to the number of periods that the service has been active (t). Note that Equation (6) implies that the learning model essentially tracks the idea of updating last period's mean belief with the deviation of the current period's signal from that last period's mean belief. The influence of the deviation term, $[Q_{ht} - \bar{Q}_{ht-1}]$ which is $\frac{\sigma_{Q_{ht}}^2}{\sigma_{Q_{ht}}^2 + \sigma_h^2}$, declines over time as the household converges to its true value.

3.4 Model Implications

The uncertainty experienced by households is likely to influence the expected utility from subscription in three different ways. First, from Equation 5, the expected utility is a decreasing function of the total uncertainty, $\sigma_{ht}^2 = \sigma_{Q_{ht}}^2 + \sigma_h^2$, because of risk aversion. Households that experience high variability in their signals (i.e., large σ_h^2) are more likely to terminate because of risk aversion compared to those with low variability. Second, the temporal variation in signal quality has an additional effect via Bayesian learning. From Equation 7, we can see that the rate at which a household's uncertainty about its average quality reduces as it accumulates information is a decreasing function of σ_h^2 . Since the uncertainty about average quality ($\sigma_{Q_{ht}}^2$) enters the total uncertainty, $\sigma_{ht}^2 = \sigma_{Q_{ht}}^2 + \sigma_h^2$, the presence of high temporal variation in service quality is likely to have an additional adverse effect via risk aversion. Note that the posterior uncertainty about one's own average quality is bounded from above by the prior variance, $\sigma_{Q_{ht}}^2$, i.e., $\sigma_{Q_{ht}}^2 < \sigma_{Q_{h0}}^2, \forall t > 0$. Hence, if the prior variance is relatively small compared to the temporal variation, the second adverse effect

of high variability has a limited impact on termination.

Third, Equation (6) suggests that the rate at which households update their beliefs about the true quality of the service will be a decreasing function of the variability in their signals (i.e., σ_h^2). Therefore, if two households have the same average quality in steady state, the household with lower σ_h^2 will converge to this quality faster than the one with higher σ_h^2 . Thus, households that experience low variability are likely to be more responsive to quality than those with high variability; the weight on Q_{ht} in Equation (6) is larger for households with lower σ_h^2 . This differential effect of quality as a function of variability faced by households is the mechanism by which our model rationalizes the interaction effect on termination.

Implication 1: For new products or services where households are uncertain about the quality they experience, those experiencing low temporal variability in quality are likely to be more responsive (in terms of termination) to the average quality level compared to those experiencing high variability. This leads to an interaction effect between average quality and variability on termination.

Since the households in our data voluntarily signed up for the service, they must have had a high prior expectation about the quality of the service (or a very large positive shock at the time of activating the service although this is less likely for all the subscribers). Moreover, the firm advertised the service based on the maximum number of movies that would be updated each week. Hence, we can reasonably assume that a majority of households have a prior belief that exceeds the true quality they experience.¹¹ Since households with higher values of σ_h^2 are likely to update their beliefs slowly, this learning deterrence would imply that they would stay with the service longer. On the other hand, learning deterrence is likely to adversely affect customer retention amongst households receiving higher signal quality vis-à-vis their prior belief about the service at the time of activation. Consequently, learning deterrence is likely to work in the same way as risk aversion for these households.

Implication 2: The interaction effect between quality and variability will lead to lower termination rates amongst households experiencing high variability if these households have prior beliefs about the quality of the service that are higher than what they actually receive.

When variability increases, so would the penalty from risk aversion. This would lower the utility from continuing with the service. At the same time, learning deterrence would imply that the utility from staying

¹¹We assess the validity of this assumption in a later section.

with the service would increase for households that have higher priors than their true quality. The net effect of increasing variability would depend on the relative magnitude of these two factors. Equation 6 implies that the rate of learning deterrence is dictated by $\frac{\sigma_{Q_{ht}}^2}{\sigma_{Q_{ht}}^2 + \sigma_h^2}$, with lower values leading to slower updating. This function is bounded between 0 and 1 and becomes flatter as σ_h^2 increases. Therefore, if the average variability is very high, the marginal benefit of learning deterrence would be relatively small for both high- and low-variability households. Thus, the risk aversion effect is likely to dominate when the overall service variability is very high. Consequently, the lower termination amongst high variability households at low quality levels would cease to exist if the overall variability is very high. Nevertheless, we would still see an interaction effect in terms of differential responsiveness to quality improvements amongst low- and high-variability households.

Implication 3: The interaction effect will lead to lower termination amongst households experiencing high variability only if the level of variability is not too high; if variability is very high, risk aversion will dominate and high variability will always lead to higher termination even if the interaction effect exists.

Implication 1 shows how the model can rationalize the interaction effect, implication 2 shows how this translates into the observed termination pattern whereas implication 3 highlights the boundary condition associated with the effect.

4 Estimation

4.1 Overview

In a typical learning model, there are four parameters of interest: (i) mean of the prior belief about the unknown entity (Q_0), (ii) prior variance (e.g., $\sigma_{Q_{h0}}^2$), (iii) variance of the signals about the unknown entity (e.g., σ_h^2), and (iv) true value of the unknown entity (Q_h). Researchers typically estimate the true value of the unknown entity when they do not observe the exact signals received by the households.¹² Since we observe the signal strength received by individual households during each period, we do not estimate Q_h . We fix the signal variance, σ_h^2 , based on the actual variance in signal quality received by the household.

¹²In these situations, the information is assumed to come from a distribution with unknown mean and variance. The researcher would then simulate the information signals by making draws from this distribution, with the unknown moments of the distribution being estimated from the data.

Note that this assumption is consistent with the model-free evidence that we presented earlier that the rate at which customers learn about the quality of the service (or the perceived precision of the experienced quality) is related to the variability that they experience; the rate of learning is slower in the presence of high variability.¹³ Thus, the only components of the learning model that we estimate are the heterogeneous prior mean and variance about the signal quality.

In our application, we use city dummies, month dummies, income, number of children, and the number of elderly people in the household as the demographic characteristics that shift the intrinsic preference for the service. The city dummies pick up differences in unobserved characteristics of subscribing households across the three test markets.¹⁴ With little loss in model fit, we discretized the continuous demographic variables by performing a median split with the high values coded as 1 and the low values coded as 0. Such a discretization is especially useful when extending the model to account for forward looking behavior.

As discussed earlier, we allow the parameters α_h (choice intercept), and β_h (responsiveness to signal quality) to vary with the household's intrinsic propensity to use the service. Although our data include information on the number of movies rented by each household over time, we cannot directly introduce the information directly into the model. This is because usage is an endogenous decision made by the household based on (a) the average quality of the service that it receives and (b) its intrinsic propensity to use the service. Rather, we need to isolate the intrinsic propensity of a household to use the service separately from the influence of signal quality and use this information in the model specification.

We accomplish this by using a two-step approach. In the first step, we regress the number of movies rented by each household during any time t , R_{ht} , on a flexible function of the average signal quality received by that household until t . Formally,

$$R_{ht} = \omega_h + H\left(\tilde{Q}_{ht}\right) + \nu_{ht}, \quad (8)$$

where $H\left(\tilde{Q}_{ht}\right)$ is a flexible function of \tilde{Q}_{ht} , the average signal quality experienced by household h until time t and ω_h is the intrinsic propensity of the household to use the service, captured by household-specific fixed

¹³Another way of justifying this assumption is that households have experienced other terrestrial transmission methods e.g., over-the-air television signals that have similar issues with variability. As those signals do not involve downloading however, inferences regarding the average number of movies downloaded is still uncertain although the variability is known. Subsequently, we verify the sensitivity of our key results to this assumption by incorporating uncertainty in this variance. We thank Andrew Ching for suggesting this explanation.

¹⁴Unobserved characteristics might arise due to differences in the firm's advertising efforts across markets. If these differences in marketing efforts are correlated with the average signal quality in these markets, not accounting for the unobserved characteristics is likely to yield biased estimates.

effects. The term u_{ht} is an IID error term. In our empirical application, we specify $H(\tilde{Q}_{ht})$ as a flexible third degree polynomial function of \tilde{Q}_{ht} .^{15 16}

We allow for heterogeneity in the prior variance, choice intercept, and the signal quality effect. We model this unobserved heterogeneity using a latent class specification with consumers belonging to different segments. Specifically, we assume there are R segments of households such that the parameter vector, $\Theta^r = \{Q_{r0}, \sigma_{Q_{r0}}^2, \alpha_r, \beta_r\}$ includes the set of heterogeneous parameters for segment r , $r = 1, 2, \dots, R$. Based on Equation 4, the likelihood of observing the history of household h belonging to segment r can be written as

$$L_{hr} = \prod_{t=1}^{T_h} (Prob_{hrt})^{I_{ht}} (1 - (Prob_{hrt}))^{(1-I_{ht})}, \quad (9)$$

where, I_{ht} is an indicator for whether the household subscribes to the service (1) or terminates (0). The corresponding overall likelihood for all households is:

$$L = \prod_{h=1}^H \left(\sum_{r=1}^R \pi_h^r L_{hr} \right). \quad (10)$$

In the above expression, π_h^r is the probability that household h belongs to segment r , $r = 1, 2, \dots, R$ and L_{hr} is the likelihood for household h conditional on it belonging to segment r . Following the literature on concomitant variable latent class models (see Dayton and McReady 1988; Gupta and Chintagunta 1994), we allow the probability that household h belongs to segment r , π_h^r , to be a function of the household's intrinsic propensity to use the service, $\bar{\omega}_h$. Specifically, we have

$$\pi_h^r = \frac{\exp(\lambda_r + \chi_r \bar{\omega}_h)}{1 + \sum_{r'=1}^{R-1} \exp(\lambda_{r'} + \chi_{r'} \bar{\omega}_h)}, \quad (11)$$

where, $\{\lambda_r, \chi_r\}$ are parameters to be estimated. For identification, we set the parameters of the R^{th} segment

¹⁵Note that \tilde{Q}_{ht} is different from \bar{Q}_{ht} in the Bayesian model; while the former captures the actual average signal quality received by household h until t , the latter captures the household's perception about the average quality of the service at time t .

¹⁶A more general approach would be to formally model usage and subscription decisions as outcomes of utility maximizing behavior by households, with signal quality having a direct bearing on both decisions. Since the two decisions are temporally separated, the subscription decision would be a function of expected usage. Such a formulation will enable us to characterize the effects of both the intrinsic preference for the service and the usage propensity on the subscription decision. However, such a formal treatment we believe, detracts from the main point of the paper, i.e., the effect on termination behavior of the interaction between the mean and variance in service quality.

to zero.¹⁷

4.2 Identification

In most common applications of learning models, researchers estimate the prior mean and the signal variance and fix the prior variance to 1 for identification (see, for example, Narayanan, Manchanda, and Chintagunta (2005)). As discussed earlier, since we use the signal variance information from the data, we can infer the ratio of the signal variance to the prior variance, and consequently, the prior variance based on the rate of learning; a lower ratio is synonymous with faster learning. The challenge is to identify the prior mean separately from the prior variance; a household can hold a high posterior belief after receiving low signals either because it started off with a high prior belief or because it perceives a very low prior uncertainty. There are two arguments for the identification of the prior mean separately from the prior variance. First, unlike in most common applications, we observe the signal quality. This additional information allows us to characterize the evolution of the posterior belief conditional on the prior mean and variance precisely; in scenarios where we do not observe the actual realized signals, we need to estimate the mean of the underlying distribution from which signals are generated and integrate over this distribution to derive the posterior belief. Second, the identification of the prior mean separately from the prior variance is enabled by the functional form of the Bayesian learning model. Specifically, as a household accumulates more information, its posterior belief becomes more precise. As a result, the household becomes progressively less prone to updating its beliefs in response to realized signals. Since the evolution of the posterior variance is governed by a specific functional form (as in Equation 7), this enables us to identify the prior mean separately from the prior variance. Thus, identification is partly based on the functional form of the learning model.¹⁸

In addition to the prior mean and variance, we estimate the following parameters: the intrinsic preference for the service, α_h , effect of perception about signal quality, β_h , the effect of demographic characteristics, ψ , eleven month fixed effects, τ , and more importantly, the risk aversion parameter, γ . As previously noted,

¹⁷

An alternative approach would be to let α_h and β_h be linear functions of ω_h , such that $\alpha_h = \bar{\alpha} + \zeta_\alpha \omega_h$ and $\beta_h = \bar{\beta} + \zeta_\beta \omega_h$. However, such a formulation would add an additional state dimension when extending the model to one with option value. Further, there could be a concern with any residual endogeneity in usage if we directly introduce the fixed effects into the utility function. The results from the myopic version of the model yielded very similar substantive results under the proposed and alternative formulations.

¹⁸We would like to thank an anonymous reviewer for suggesting this identification strategy.

one of the unique aspects of the paper is the ability to estimate the risk-aversion parameter γ . There are two sources of variation that enable us to identify the risk aversion parameter. The first is the cross-sectional difference in variability in service quality experienced by households. From Equation 4, we can see that as the posterior variance $\sigma_{Q_{ht}}^2 \rightarrow 0$ over time via learning, the total uncertainty experienced by household h , $\sigma_{ht}^2 \rightarrow \sigma_h^2$. Since we observe the variance of quality received by individual households, σ_h^2 , differences in variability experienced by households allows us to identify γ . Second, as households accumulate information about the quality of the service over time, their posterior beliefs become more precise. The rate at which this precision increases would depend on the precision of the information, $1/\sigma_h^2$. This temporal and cross-sectional variation in $\sigma_{Q_{ht}}^2$ further strengthens the case for identification of the risk aversion parameter. In contrast, previous research (see, for example, Crawford and Shum (2005)) has had to base identification of the risk aversion parameter only on the temporal variation in posterior variance, $\sigma_{Q_{ht}}^2$, which is not directly observed by the researcher.

5 Results

5.1 Overview of the Results

We began our estimation by identifying the optimal number of segments in the heterogeneity distribution. To this end, we first estimated models with one, two, and three segments. Based on BIC, we identified the two-segment solution as the best. We present the estimates in Table 4. Recall that in our empirical specification, we characterize segment membership probabilities as functions of each household's intrinsic propensity to use the service. Based on the posterior segment membership probabilities, we assign each household to one of the two segments. We present information on the size, composition, and characteristics of the two segments in Table 5.

The results from Table 5 imply that segment 1 is the larger of the two with approximately 80% of the households, while segment 2 accounts for the remaining 20%. Households in segment 1 have a relatively higher intrinsic preference, but are less responsive to signal quality compared to segment 2. Nevertheless, as expected, both segments respond positively to signal quality.

Our estimate of the risk aversion parameter is 1.22 (standard error = 0.427). We restrict the risk-aversion

term to be positive by exponentiating the corresponding parameter. The statistically significant risk aversion parameter corroborates the pattern in the raw data that, on average, high variability in signal quality is associated with higher probability of termination.

The results in Table 4 imply that the probability of belonging to segment 1 decreases with the propensity to use the service (ω_h). This intuition is reflected in the results in Table 5, which suggest that households in segment 2 use the service more than those in segment 1. However, despite their higher propensity to use the service, households in segment 2 terminate the service at a higher rate than those in segment 1. The low value of the intercept for this segment suggests that households in this segment derive a very low intrinsic utility from the service (after accounting for usage) although they derive a high consumption utility (since as we note above, segment 2 households have a higher propensity to use the service).¹⁹ Furthermore, as we discuss below, households in segment 1 experience higher average signal quality compared to those in segment 2 (1.46 for segment 1 vs. 1.19 for segment 2), which can also partly explain the higher termination rates among segment 2 households. However, there is no difference among the segments in terms of the variance in signal quality that they experience.

The results from Table 5 suggest that households in segment 2 have a prior belief that they will receive the maximum signal quality of 2.4 (standard error = 0.072).²⁰ Relative to this prior belief, households in this segment experience average signal quality of 1.19. On the other hand, households in segment 1 have lower prior beliefs at 1.65. In contrast, these households experience average signal quality of 1.46. Therefore, on average, households in both segments have higher prior beliefs than the actual quality that they receive. Recall that per Implication 2, the interaction effect between quality and variability will lead to lower termination rates amongst households experiencing high variability if these households have prior beliefs about the quality of the service that are higher than what they actually receive. Therefore, our results seem to rationalize the observed data patterns.

The estimated variances for the prior belief (i.e., uncertainty in the prior) are fairly consistent (in the range of 0.086-0.09) across the two segments. Note that we restrict the variances to be positive by expo-

¹⁹

In the alternative formulation where we can characterize the intercept as $\alpha_h = \bar{\alpha} + \zeta_\alpha \omega_h$, we can parse out the intrinsic utility from the service ($\bar{\alpha}$) separately from the consumption utility, $\zeta_\alpha \omega_h$. Although not reported here, the main substantive results are very consistent across these two alternative formulations of how the parameters of the utility function are related to the intrinsic propensity of a household to use the service.

²⁰In order to restrict the estimated prior mean to lie in the range of 0-2.4, we use a logit transformation such that $Q_{r0} = 2.4 \left(\frac{\exp(\rho_r)}{1 + \exp(\rho_r)} \right)$, where ρ_r can take any value between $-\infty$ and $+\infty$. The estimates reported in Table 4 correspond to ρ_r , while those reported in Table 5 are Q_{r0} .

entiating the estimated parameters. Hence, the parameters reported in Table 4 are the logarithms of the variance terms. As discussed in the model section, the rate at which households update their prior beliefs about signal quality via learning would depend on the relative magnitudes of the variability in the signals that they receive vis-a-vis the prior variance. The average variances in signal quality experienced by low- and high-variability households (defined based on a median split) were 0.075 and 0.372, respectively. Given that the signal variance for high-variability households is higher than the estimates for the prior variance, these households are likely to be slow in updating their beliefs about signal quality over time.

We illustrate the relative rate of learning among low- and high-variability households implied by these estimates in Figure 5. As in the data, we assume that the low-variability household perceives a variability of 0.075, while the high-variability household perceives a variability of 0.372. To illustrate the differential rate of updating the prior belief, we consider two identical households (low- and high-variability, respectively) that receive a low average quality of 0.1 in Figure 5. These results suggest that there is marked difference in the rate of updating between the low- and high-variability households. For example, whereas the low-variability household in segment 2, starting with a prior of 2.4, would update its belief about the average service quality to less than 1 within three periods, it would take the high-variability household nine periods to do so. Consequently, high-variability households consistently have higher beliefs about the quality of the service compared to those experiencing low-variability. As is evident from the trajectory of the posterior mean for households in segment 1, who start with a much lower prior belief of 1.65, this result holds whenever the prior belief is higher than the actual quality. However, at high quality levels, the benefit from learning deterrence is very small, and is likely to be offset by the higher risk aversion.

A typical test for the validity of a model is its ability to replicate key patterns in the data. As motivated throughout the paper, a key pattern of interest is the differential effect of variability in service quality on termination rates amongst households receiving low versus high quality. In order to verify that the estimates from both versions of the model can indeed replicate these patterns, we simulated the termination probabilities for the same groups of customers. We present the results from this simulation along with the corresponding patterns in the raw data in Figure 6. These results suggest that our model estimates are able to replicate the interaction effect between signal quality mean and variability on termination.

5.2 Alternative Specifications and Model Extensions

5.2.1 Alternative Specifications

We compare the proposed model with two other alternative specifications. The first alternative specification is based on Bolton, Lemon, and Bramlett (2006), who propose that variability can increase customer retention if the variation results in extreme positive experiences. This is an appealing explanation for the lower termination rates among high-variability households that experience low quality, on average. In order to test if this alternative explanation can yield a superior fit, we estimated a model where we specified the utility that a household would derive from subscribing to the service at time t as a function of (a) the average quality experienced by a household until the previous period, (b) the number of extreme positive signals (defined as > 2) received till time t , and (c) the number of extreme negative signals (defined as < 0.5) received till time t . In addition, as in the proposed model, we included controls for demographic characteristics and seasonality. Note that the specification with the average quality experienced until the previous period is similar to a learning model with a very diffuse prior belief about the quality of the service. The alternative model fit the data worse than proposed model in terms log likelihood (-3104.63 vs. -3008.51) and BIC (6474.58 vs. 6293.0). In addition, we estimated a more refined version of the above specification by including the “gain” and “loss” of the most recent signal received relative to the mean signal quality till date. The rationale is that if households are learning about the quality of the service based on their experience, the average quality till date would capture such learning. The positive (negative) deviations can be construed as gains (losses) relative to this perceived quality. Even with the added flexibility, this model could not explain the data as well as our proposed specification (LL = -3037.70, BIC = 6320.72).

The second alternative rationalization for the lower (higher) termination among high-variability households experiencing low (high) average quality is that the customers are risk-seeking at low quality levels and risk-averse at the higher end of the quality spectrum.²¹ Consequently, a convex utility at lower quality levels and a concave specification at high quality levels should be able to explain the data patterns. In order to verify if this alternative specification can explain the data better than the proposed model, we estimated a flexible version of the model with risk-aversion but without learning (as in Equation 4). We impart the additional flexibility by allowing the risk aversion parameter, γ , to vary based on whether the household receives low/high quality (defined based on a median split), on average. As in the above specification, we

²¹We thank Victor Aguirregabiria and Sanjog Misa for suggesting variations of this explanation.

characterize the household's belief about the quality of the service in any given period using the average quality experienced by a household until the previous period. This alternative specification with flexible risk-aversion did not fit the data as well as the proposed model (LL = -3121.16, BIC = 6487.634). Although our discussion thus far has been on the relative fit of these two alternative explanations vis-a-vis our proposed structure, it is worth noting that these alternative accounts also cannot generate the specific implications that our model provides. Thus, the value of the proposed model is more than just being able to rationalize the data better.

5.2.2 Model Extensions

(a) Consideration of Option Value

The presence of learning implies that as a household accumulates more experience with the service, its belief about the quality would get more precise (i.e., $\sigma_{Q_{it}}^2 \rightarrow 0$ in Equation 6). As a result, in addition to the contemporaneous expected utility from subscribing, the household would also derive an option value from continuing with the service. The option value arises because the household can gather more information about the service by subscribing and hence make a more informed decision in the next period (Erdem and Keane (1996), Hitsch (2006)). The total expected utility (including the option value) that the household would derive from subscribing can be written in the form of a Bellman equation. We present details of the model that accounts for option value as well as the appropriate estimation strategy in Appendix A.

We estimated the forward-looking model by fixing the discount factor at 0.99. We discuss the results from this analysis in the appendix. Notably, the forward-looking model fit the data slightly better compared to the myopic version (log-likelihood of -2995.16 for the forward-looking model vs. -3008.5 for the myopic model). Furthermore, the forward-looking model yielded results that were similar to the myopic version in terms of direction of the parameter estimates and the composition and characteristics of the two segments (please see the appendix for results). More importantly, the forward-looking model was able to replicate the key data patterns as the myopic version. Nevertheless, as the myopic version of the model was adequate in describing the data patterns, we appeal to model parsimony and focus on the myopic version of the model.

(b) Accounting for Bounded Signal Quality

As in many other scenarios where researchers have considered learning, the entity about which households

update their beliefs, i.e., mean signal quality, is bounded; in our context, the signal quality (number of new movies downloaded) is bounded between zero and 2.4 (40 per month).²² However, in view of empirical tractability, we follow the precedence in the Bayesian learning literature by assuming a normal distribution for the realized signal quality. While this side-steps the issue of bounded nature of the data, basing our model on more general non-conjugate distributions would render it empirically and analytically intractable.²³ In view of this tradeoff, we formulated an empirical strategy that accounted for the bounded nature of signal quality, while retaining conjugacy. We accomplish this by defining an underlying latent signal quality that is unbounded and can thus be treated as normally distributed. Households learn about this latent variable through their experience. However, their subscription decision would depend only on the corresponding signal quality that is bounded between 0 and 2.4. One can impose these bounds by treating the realized signal as a censored outcome of the underlying latent variable. The approach requires us to transform the latent posterior belief into the corresponding expected quality and variance that would eventually affect the subscription decision. We provide details of this model formulation and estimation in Appendix B.

We present the results from the myopic version of this model in Appendix B. Overall, the key results regarding the presence of two distinct segments and evidence of risk aversion are similar to those from the model that does not account for the bounded nature of our data. Furthermore, the estimates were able to replicate the key aspects of our data, namely, the differential effect of variability in service quality on termination rates amongst households receiving low vs. high quality. However, this model yields inferior fit (log-likelihood of -3016.193) compared to the model that does not account for the bounds (log-likelihood = -3008.511). Therefore, the rest of our discussion below is based on the latter estimates.

(c) *Uncertainty and Learning about Signal Quality Variability*

Our empirical specification assumes that the perceived precision of the information that household accumulates as they gain experience with the service. Our model-free evidence suggests that the rate at which households respond to the cumulative information about the quality of the service is inversely related to the actual variance they receive. In view of this evidence, we fix the signal precision for each household,

²²In our data, there are 4715 (out of 27486) observations when the signal quality is potentially bounded from below, i.e., we observe 0 signal quality. On the other hand, there are only 18 observations where the signal quality is 2.4. Therefore, being bounded from above is less of an issue.

²³In some instances, we can use some naturally bounded conjugate distributions such as beta-binomial to characterize the bounded nature of our data. However, the expressions for the mean and variance in this family are related to each other. This runs counter to our data wherein we have households experiencing different combinations of mean and variance.

$1/\sigma_h^2$, as the inverse of the actual variance in signal quality received by them. Nevertheless, we have made a specific assumption that the precision of the information that the customers perceive is the same as the actual variability they experience. As a robustness check, we estimated a model that allowed households to be uncertain about the variance (or precision) of the information that they accumulate. We operationalized this uncertainty about variability by using a discrete distribution with two supports.²⁴ The results from the myopic version of the model revealed that the substantive results remained unaltered with this augmentation.

The above model assumes that households do not learn about their true variance over time; rather, they know the distribution of variance and account for that in their behavior. A more complete specification would entail allowing for households to also learn about this variance. A natural approach to accommodating such behavior in a computationally tractable way would entail exploiting the conjugacy of the normal and inverse gamma priors for the mean and variance, respectively (Gelman, Carlin, Stern, and Rubin (2000), Zhao, Zhao, and Helsén (2011)). A downside with this approach is that the rate of learning would not depend on the precision of the information (signals).²⁵ This is inconsistent with our proposed mechanism (motivated by data patterns) that the rate of learning is influenced by the variability in signals received.

5.3 Implications

Earlier, we had laid out three implications of the model. In this section, we demonstrate them numerically based on the parameter estimates discussed above. Implication 1 suggests that the households with low variability should be more responsive to changes in the quality of the service compared to those experiencing high variability. In order to demonstrate this, we simulated the termination probabilities when signal quality is higher by 1%. We present these results for different mean levels and variability in quality in Table 6. Consistent with the positive effect of signal quality, the termination rates decrease for all groups. More importantly, the effect is much stronger for households experiencing lower variability, both in absolute and percentage terms. Notably, the average signal quality elasticity for low-variability households is significantly higher at -1.88 (standard error = 0.22) compared to -0.51 (standard error = 0.13) for households experiencing

²⁴Consistent with the model-free evidence, we assumed that the mean of this distribution of uncertainty for each household is the same as the actual variance that they experienced.

²⁵This is evident from Equation 13 in Zhao, Zhao, and Helsén (2011) wherein the posterior belief is independent of the variance of the signal. Our empirical calibration of such a model could not replicate the differential effect of variability in service quality on termination rates amongst households receiving low versus high quality.

high variability in signal quality.

Implication 2 suggests that the extent to which termination rates are lower among high variability households would depend on their prior belief vis-à-vis the actual quality that they experience. Specifically, since the lower termination rates amongst households experiencing high variability accrues due to learning deterrence, it can only happen if the prior belief is higher than the true quality they experience. We demonstrate this via counterfactual analyses wherein we assume values of parameters that are similar to the model estimates and simulate the termination probabilities for different values of priors. Specifically, like in the data, we assume that (scaled) true quality can be in the range of 0-2.5 with households experiencing quality levels in increments of 0.25. We simulate the termination probabilities under two scenarios: prior belief = 0 and 2.25. We present the results from this analysis in Figure 7. When the mean of the prior belief is 2.25, learning deterrence lowers termination rates amongst low quality households. Thus, high variability households have lower termination rates at low quality levels compared to their low variability counterparts. On the other hand, when the mean of the prior belief is 0, high variability households always have higher termination rates. Nevertheless, the interaction effect arising due to differential responsiveness to service quality still exists.

In order to demonstrate that the lower termination due to learning deterrence occurs only at quality levels lower than the prior, we plot the difference between the learning components of utility (i.e., $\beta_h \bar{Q}_{ht} - \gamma \bar{Q}_{ht}^2$ from Equation 4) for high vs. low variability households in Figure 8. Positive values imply that the learning component for the high variability types is higher than that for the low variability types. The results reveal that when the prior mean is 2.25, the difference in learning component of utility is positive throughout. As the prior decreases, it becomes progressively more negative over a large range of qualities, i.e., learning deterrence will increase termination if households have a lower prior belief compared to their true quality. Together, these results reinforce the intuition that learning deterrence is likely to lower termination only amongst households who believe that the quality of the service is better than what it actually is.

Implication 3 predicts that the range of average quality levels at which the termination rates amongst high variability households is lower than their low variability counterparts would diminish as the level of variability (σ_h^2) increases for all customers. On the other hand, if the variability is lower, we would observe lower termination among high-variability households over a larger range of quality levels. In Figure 9, we present termination probabilities when variability increases/decreases by 50% for all households. When variability decreases by 50%, the two lines intersect at the average quality bin of 1.5-2. Note that this is higher than the

point of intersection under status quo, i.e., the average quality bin of 1-1.5 (as in Figure 6). Therefore, when variability is decreased, termination rates would be lower among high-variability households (compared to their low-variability counterparts) over a larger range of quality levels. On the other hand, when variability is increased by 50%, termination rates are always lower among low-variability households. As an extreme outcome, when variability increases by 82%, we would not see the high-variability households having lower termination rates at low quality levels, although the interaction effect in terms of differential responsiveness to quality among low- and high-variability households would still exist.

5.4 Managerial Implications: Relative Effects of Increasing Mean Quality vs. Lowering Variability on Termination

Our empirical analysis suggests that there are two possible levers that managers can employ to increase customer retention, especially at the higher end of the quality spectrum: increase average quality and/or lower variability. In this section, we perform counterfactual analyses to understand the relative efficacy of these two alternatives in lowering termination.²⁶ Clearly, at low quality levels where lowering variability can increase termination, managers need to focus on increasing average service quality. On the other hand, at intermediate and high quality levels, both lowering variability and increasing average quality are viable options. In our counterfactual analysis, we determine the increase in average number of movies that can accomplish the same reduction in termination as lowering the standard deviation by one movie. We present the results from this analysis in Table 7. These results suggest that at intermediate quality levels, increasing the mean is as effective as lowering the standard deviation. On the other hand, at high quality levels, we need a much larger increase in the mean number of movies to achieve the same lower termination as lowering standard deviation by one movie. In fact, at very high quality levels, the required increase in quality would take the number of movies beyond the feasible range. This happens because of two reasons. First, the concavity of the utility function implies that the marginal effect of increasing the mean decreases as quality increases. Second, the marginal benefit of lowering the standard deviation by one movie is higher

²⁶While service quality mean and variability are exogenous in our context, we perform these analyses to highlight general implications that are possibly applicable to contexts that are beyond our own where managers can indeed influence quality levels. Moreover, for our specific context, one can construe our implications as being more “long term” in nature. After the test market in the 3 cities, the company decided to withdraw the service to re-launch at a later date. If the firm now has the capability to boost quality or to influence its variability, these implications would be appropriate.

at high quality levels. Given that low variability households are (a) more responsive to quality and (b) yield lower benefit from lowering variability, increasing quality is likely to be more effective in lowering their termination than for the high variability households. Overall these results imply that at very low quality levels, increasing mean quality can be effective in lowering termination. On the other hand, at high quality levels, lowering variability is likely to be more effective.

The implications of these results go beyond our application; a recent *Wall Street Journal* article (WSJ 2013), “Video Boom Forces Verizon to Upgrade Network,” highlights the importance of the tradeoff that we investigate. While *Verizon* has been successful in rolling out its new 4G LTE service more widely across the US than other carriers, in certain locations the capacities of the base stations have not been able to keep up with the increased data needs. This had led to speed issues (average quality) and service variability. The implication of our analysis for this case depends on whether the focus is on customers who travel across markets that have 4G LTE service or if one is interested in the segment that is largely confined to one market. For the former group, Verizon’s service quality will have a high mean and low variance - a good situation for the company. For the latter segment, in markets with no congestion, the situation is once again that of high mean and low variance. Now compare Verizon’s situation to another provider that does not have as extensive a 4G LTE network coverage as Verizon. Customers of such a provider who sign up for the provider’s 4G service are likely to experience mean and variance in quality depending once again on whether they belong to the traveling or to the local segment. If these customers travel extensively in 3G markets then they would experience lower-than-4G quality with low variation - a situation detrimental to the carrier. Here, the focus should be on the wider rollout of 4G service that increases the average quality of the service for the customer. However, if the customer is confined to a local market with high congestion then as in the Verizon case, the company should focus on lowering variability. We recognize that the telecom market is far more complex and does not necessarily have a 1-to-1 mapping with our VOD business; nevertheless the general nature of the implications appear to be relevant in that case as well.

6 Conclusion

While the positive relationship between customer retention and the average level of service quality has been well established in the literature, the corresponding link with variability is less clear. Notwithstanding the

dominant view that higher variability will lower retention, we document empirically that, in the context of new services, the effect of variability on customer retention might depend on the average service quality. In particular, we show the presence of an interaction effect: the marginal reduction in termination with increase in average quality is likely to be higher amongst customers experiencing low variability. As an extreme outcome, at low quality levels, lowering variability can lead to an increase in termination rates.

We propose a mechanism involving risk aversion and learning that can rationalize this interaction effect. Based on the model and the empirical results, we establish the boundary conditions under which lowering variability can lead to higher termination. From a managerial perspective, we document the relative tradeoff between lowering variability and increasing the average quality in terms of their ability to retain customers.

There are several avenues in which the model and empirical analysis can be extended in future research. First, our context reflects two aspects of service quality that has been discussed in the literature (e.g., Parasuraman, Zeithaml, and Berry (1985)): (a) service quality is generally inconsistent (i.e., quality of the service varies from customer to customer) and (b) cannot be inventoried (i.e., movies that are not downloaded in a given week due to poor service quality are never available to the household even at a future date) of service quality. However, it does not reflect the notion that service quality is essentially inseparable (i.e., it is realized as a consequence of interaction between the customer and the service provider) and intangible (i.e., cannot be measured or quantified precisely). In this regard, our context is similar to common applications such as cell phone service, where customers might be able to objectively infer the quality that they experience and trace it back to the service provider. Nevertheless, even in such contexts (including ours), there are other aspects of service quality such as billing and conflict resolution wherein the intangible and inseparable nature of service quality would be relevant. These, in turn, would render it difficult to understand how customers perceive and evaluate the service as a whole (Zeithaml (1981)). Therefore, the aspect of service quality that we consider can be viewed as a component of the composite service quality. If we believe that these other aspects of service quality are orthogonal to the measures that we consider we can operationalize these by treating the composite service quality as the sum of the objective quality and some uncorrelated, mean zero, random shock from a parametric distribution such as normal. However, one needs to be careful to ensure that the additional parameter that would need to be estimate, viz., the variance of the shock is identified. On the other hand, if we believe that these shocks are likely to be correlated with the objective measure, we need to consider a more general structure that accounts for such dependence.

Second, in consideration of empirical tractability, we have followed the current literature and specified

the learning mechanism in terms of the normal-normal conjugate distribution. Nevertheless, in our model extensions, we account for the bounded nature of signal quality. It appears from that analysis that the key intuition regarding learning deterrence and risk aversion leading to the interaction effect is likely to be invariant to our distributional assumption. At the same time, exploring alternative approaches to accounting for the bounded nature of the signals may be worthwhile from the perspective of having a general model that deals with such data (see Ching, Erdem, and Keane (Forthcoming) for a discussion of some richer specification of learning models). Finally, one can consider a richer specification of the learning mechanism by incorporating consumer forgetting (Mehta, Rajeev, and Srinivasan (2004)). If the data have additional variation in terms of exogenously different service subscription / renewal frequencies, such as monthly, quarterly, and annual, one can exploit this variation to identify the interaction between learning and forgetting behavior. Managerially, such an analysis might help managers design appropriate contract lengths.

References

- BOLTON, R., K. LEMON, AND M. BRAMLETT (2006): "The Effect of Service Quality Experiences Over Time on a Supplier's Retention of Business Customers," *Management Science*, 52(12), 1811–1823.
- CHAN, T., AND B. HAMILTON (2006): "Learning, Private Information, and the Economic Evaluation of Randomized Experiments," *Journal of Political Economy*, 47(6), 997–1040.
- CHING, A. (2010): "Consumer Learning and Heterogeneity: Dynamics of Demand for Prescription Drugs after Patent Expiration," *International Journal of Industrial Organization*, 28, 619–638.
- CHING, A., T. ERDEM, AND M. KEANE (Forthcoming): "Learning Models: An Assessment of Progress, Challenges and New Developments," *Marketing Science*.
- CHING, A., AND M. ISHIHARA (2010): "The Effects of Detailing on Prescription Decisions Under Quality Uncertainty," *Quantitative Marketing Economics*, pp. 123–165.
- COSCELLI, A., AND M. SHUM (2004): "An Empirical Model of Learning and Patient Spillovers in New Drug Entry," *Journal of Econometrics*, 122(4), 213–246.
- CRAWFORD, G., AND M. SHUM (2005): "Uncertainty and Learning in Pharmaceutical Demand," *Econometrica*, 73(4), 1137–1173.
- ERDEM, T., AND M. KEANE (1996): "Decision Making under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets," *Marketing Science*, 15(3), 1–20.
- GELMAN, A., J. CARLIN, H. S. STERN, AND D. B. RUBIN (2000): *Bayesian Data Analysis*. Chapman and Hall.
- HITSCH, G. (2006): "An Empirical Model of Optimal Dynamic Product Launch and Exit Under Demand Uncertainty," *Marketing Science*, 25(1), 25–50.

- KANNAN, P. K., AND J. PROENCA (2010): "Design of Service Systems Under Variability: Research Issues," *Information Systems E-Business Management*, 8(1), 1–11.
- MEHTA, N., S. RAJEEV, AND K. SRINIVASAN (2004): "Role of Forgetting in Memory-Based Choice Decisions: A Structural Model," *Quantitative Marketing and Economics*, 2(2), 107–140.
- MEYER, R. (1981): "A Model of Multiattribute Judgments Under Attribute Uncertainty and Information Constraint," *Journal of Marketing Research*, 18(4), 428–441.
- NAM, S., P. MANCHANDA, AND P. CHINTAGUNTA (2010): "The Effect of Signal Quality and Contiguous Word of Mouth on Customer Acquisition for a Video-on-Demand Service," *Marketing Science*, 29, 690–700.
- NARAYANAN, S., P. MANCHANDA, AND P. CHINTAGUNTA (2005): "Temporal Differences in the Role of Marketing Communication in New Product Categories," *Journal of Marketing Research*, 42(3), 278–290.
- PARASURAMAN, A., V. A. ZEITHAML, AND L. L. BERRY (1985): "A Conceptual Model of Service Quality and Its Implications for Future Research," *Journal of Marketing*, 49, 41–50.
- (1988): "SERVQUAL: A Multiple-Item Scale for Measuring Customer perceptions of Service Quality," *Journal of Retailing*, 64, 12–40.
- PLESKAC, T. (2008): "Decision-Making and Learning While Taking Sequential Risks," *Experimental Psychology*, 34(1), 167–185.
- RUST, R., J. INMAN, J. JIA, AND A. ZAHORIK (1999): "What you Don't Know About Customer-Perceived Quality: The Role of Customer Expectation Distributions," *Marketing Science*, 18(1), 77–92.
- SUN, M. (2012): "How Does the Variance of Product Ratings Matter?," *Management Science*, 58(4), 696–707.
- ZEITHAML, V. A. (1981): "How Consumer Evaluation Processes Differ between Goods and Services," *Marketing of Services*, pp. 186–190.
- ZHAO, Y., Y. ZHAO, AND K. HELSEN (2011): "Consumer Learning in a Turbulent Market Environment: Modeling Consumer Choice Dynamics After a Product-Harm Crisis," *Journal of Marketing Research*, 48, 255–267.

Table 1: Interaction Effect between Mean and Variability

Termination rate		
	High Signal Variance	Low Signal Variance
High Mean Signal	14.6%	10.0%
Low Mean Signal	28.4%	38.5%

Table 2: Termination Rates at Various Levels of Average Quality

LOW VARIABILITY HOUSEHOLDS				HIGH VARIABILITY HOUSEHOLDS			
Avg. Sig. Quality	Number of Households	% Terminating	Tenure (Months) Conditional on Terminating	Avg. Sig. Quality	Number of Households	% Terminating	Tenure (Months) Conditional on Terminating
0.10	50	46%	4.96				
0.24	68	47%	4.69	0.27	12	33%	8.0
0.41	65	58%	4.97	0.41	75	40%	6.77
0.58	40	70%	5.0	0.58	114	34%	6.67
0.74	36	61%	4.50	0.76	144	33%	7.0
0.92	45	69%	5.42	0.93	186	33%	6.98
1.09	42	50%	4.29	1.09	209	27%	5.96
1.26	58	14%	6.0	1.24	208	24%	6.78
1.42	104	16%	5.71	1.42	220	17%	8.16
1.59	157	13%	6.15	1.58	237	13%	7.52
1.76	222	11%	7.38	1.74	153	15%	7.22
1.92	264	14%	4.71	1.91	84	15%	9.31
2.08	256	9%	5.29	2.05	17	24%	9.75
2.25	214	14%	4.67				
2.36	38	11%	2.50				

Table 3: Results from a termination hazard model with a log-logistic baseline hazard

Dependent variable: Time to termination/end		
Variable	Coefficient	t-value
Mean signal quality	-0.62	-15.80
Signal quality variance	-0.28	-5.28
Interaction effect	0.19	3.51
Demographics	Yes	
Start period dummy	Yes	

Table 4: Model Estimates

PARAMETER	ESTIMATE	STD ERROR
INTERCEPT (SEGMENT 1)	1.386	0.494
INTERCEPT (SEGMENT 2)	-7.013	1.283
SIGNAL QUALITY EFFECT (SEGMENT 1)	4.044	0.357
SIGNAL QUALITY EFFECT (SEGMENT 2)	7.558	0.815
RISK AVERSION (EXP(.))	0.200	0.101
OCTOBER	-0.284	0.127
NOVEMBER	1.725	0.509
DECEMBER	0.176	0.215
JANUARY	0.255	0.196
FEBRUARY	0.608	0.146
MARCH	0.277	0.177
APRIL	0.259	0.170
MAY	0.078	0.156
JUNE	-0.158	0.142
JULY	-0.246	0.131
AUGUST	0.023	0.140
SALT LAKE CITY	0.402	0.094
SPOKANE	0.532	0.146
INCOME	0.397	0.085
CHILD	0.006	0.098
% OLD	0.074	0.095
SIGNAL QUALITY PRIOR VARIANCE (SEGMENT 1) (EXP(.))	-2.463	0.383
SIGNAL QUALITY PRIOR VARIANCE (SEGMENT 2) (EXP(.))	-2.500	0.296
TRANSFORMED PRIOR MEAN (SEGMENT 1)	0.790	0.375
TRANSFORMED PRIOR MEAN (SEGMENT 2)	10.296	0.034
SEGMENT 1 MEMBERSHIP PARAMETER (INTERCEPT)	1.532	0.277
SEGMENT 1 MEMBERSHIP PARAMETER (EFFECT OF INTRINSIC PROPENSITY TO USE THE SERVICE)	-0.221	0.039
LL	-3008.511	

Table 5: Segment Characteristics

	SEGMENT 1	SEGMENT 2
SEGMENT SIZE	2602	644
SIGNAL QUALITY PRIOR MEAN (ESTIMATED)	1.651	2.4
SIGNAL QUALITY PRIOR VARIANCE (ESTIMATED)	0.086	0.090
MEAN MONTHLY RENTAL (USAGE)	3.013	4.364
AVG. SIGNAL QUALITY	1.46	1.19
VARIANCE IN SIGNAL QUALITY	0.447	0.449
% TERMINATING	6%	82%

Table 6: Effect of a 1% Increase in Signal Quality on Termination Rate

Signal Quality	Absolute Change in Termination Rate		% Change in Termination Rate (Elasticity)	
	Low-Variability Households	High-Variability Households	Low-Variability Households	High-Variability Households
0-0.5	-0.38%	-0.07%	-0.78%	-0.17%
0.5-1	-0.44%	-0.09%	-0.88%	-0.26%
1-1.5	-0.23%	-0.12%	-0.98%	-0.42%
1.5-2	-0.21%	-0.12%	-1.52%	-0.57%
> 2	-0.06%	-0.12%	-0.75%	-0.64%
Average	-0.20%	-0.11%	-1.10%	-0.41%

Table 7: Tradeoff between Increasing the Average Quality vs. Lowering Variance

INCREASE IN MEAN NUMBER OF MOVIES REQUIRED TO OBTAIN THE SAME TERMINATION RATE AS LOWERING STD. DEVIATION BY 1 MOVIE		
Signal Quality	LOW VARIABILITY HOUSEHOLDS	HIGH VARIABILITY HOUSEHOLDS
0-0.5	N.A	N.A
0.5-1		1.12
1-1.5		1.7
1.5-2		2.25
> 2	0.7	NEVER

Figure 1: Distribution of Average Signal Quality across Households

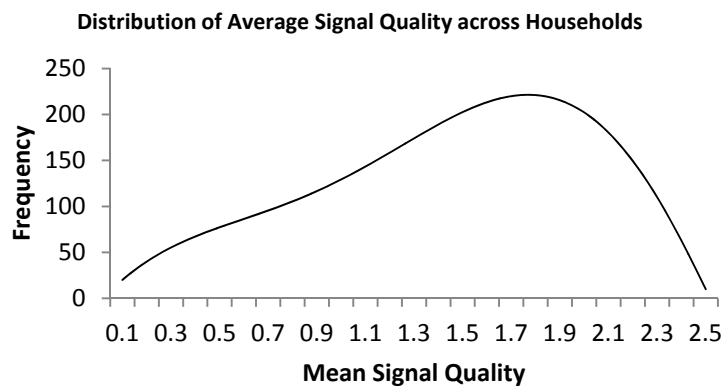


Figure 2: Distribution of Within-Household Variability in Signal Quality

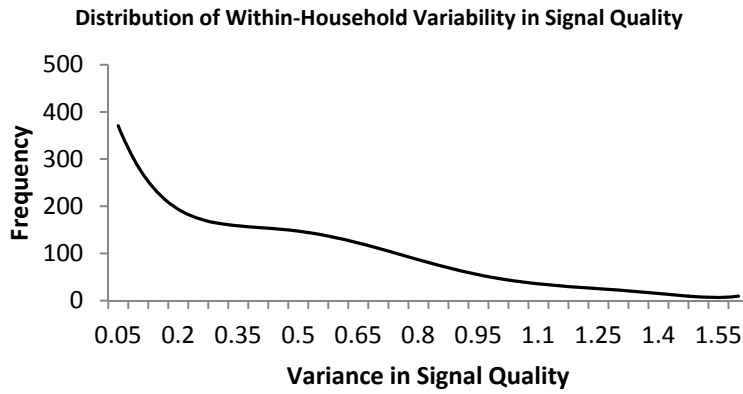


Figure 3: Evidence of Learning: Termination Rates after Receiving a Certain Number of High vs. Low Quality Signals

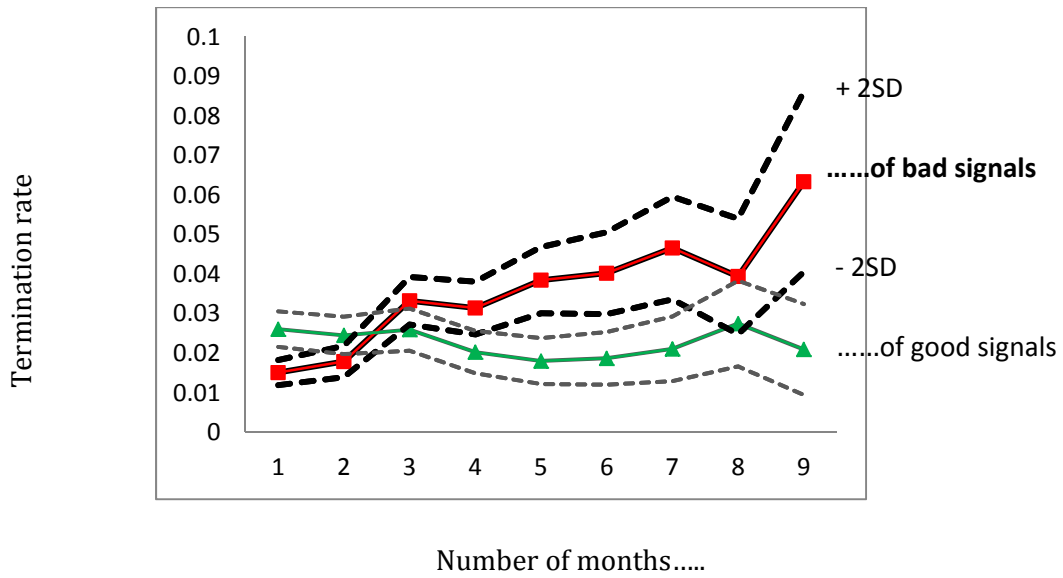


Figure 4: Evidence of Differential Learning Among Households Experiencing Low-vs. High-Variability

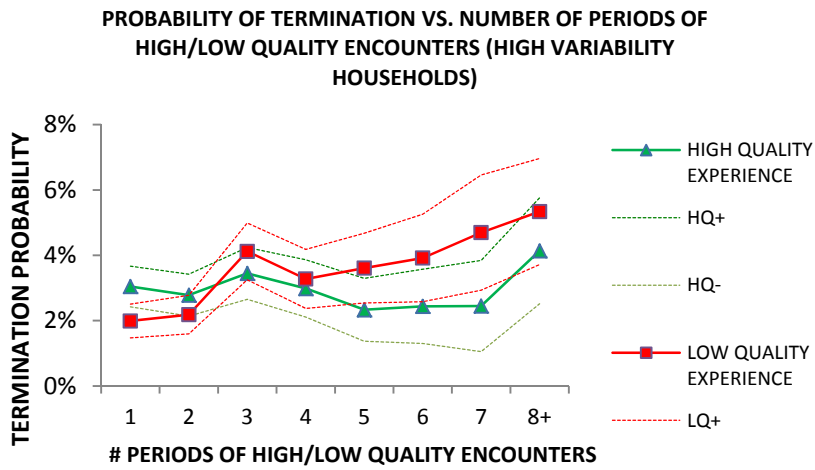
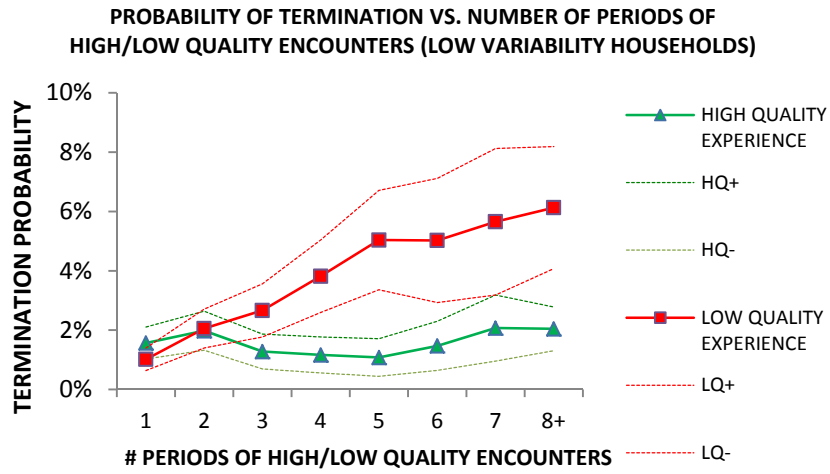


Figure 5: Implication of the Model Estimates

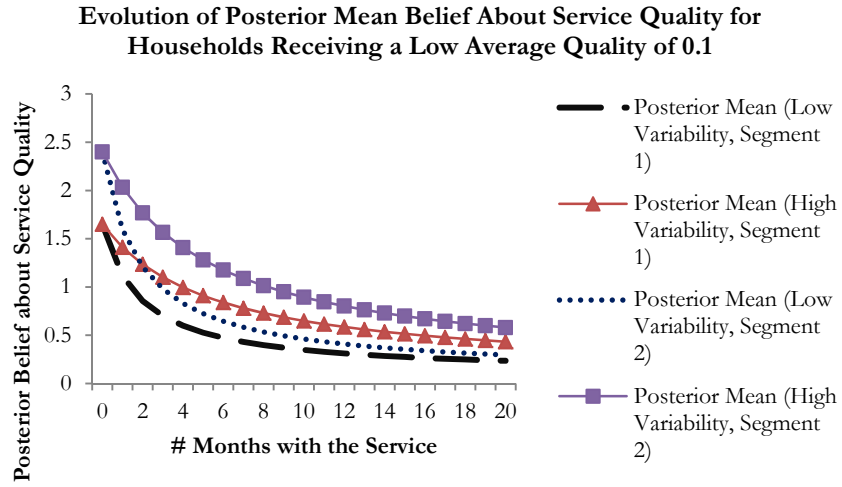


Figure 6: Termination Rates Implied by the Model Estimates

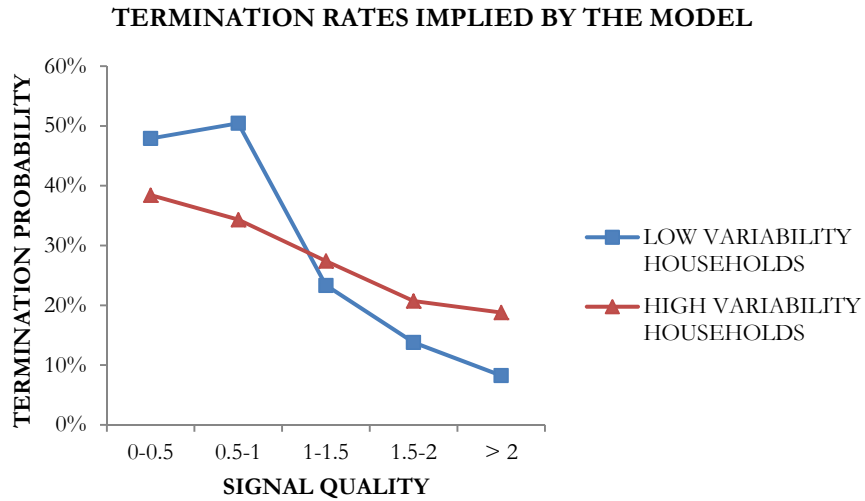


Figure 7: Effect of Prior Belief and Variability on termination Probabilities

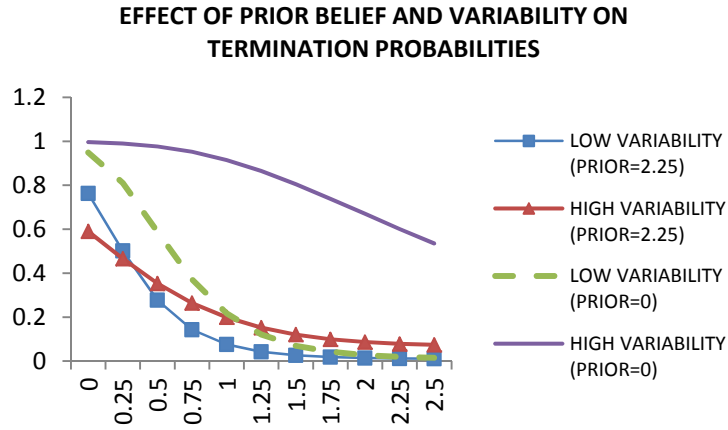


Figure 8: Difference Between the Learning Component of Utility for High and Low Variability Types

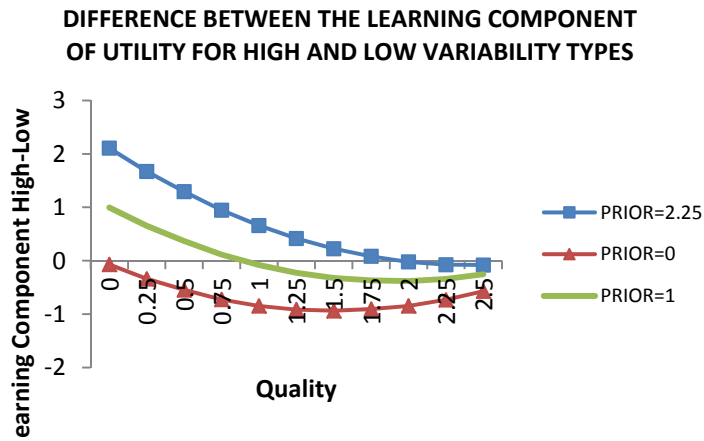
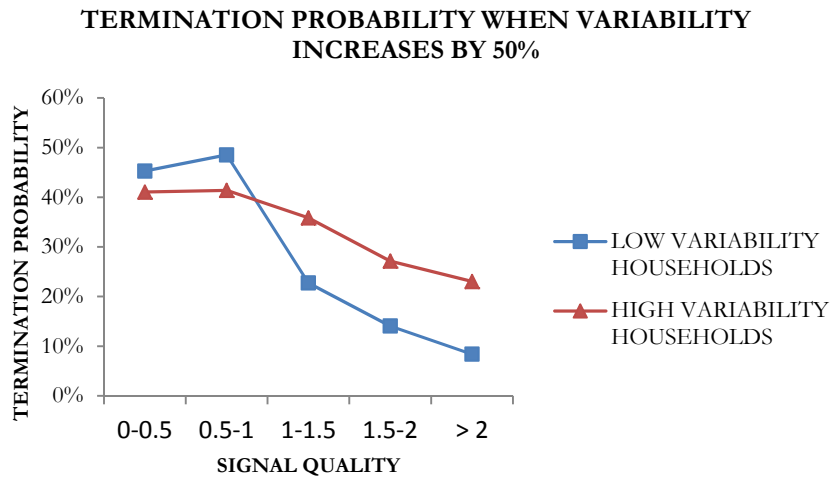
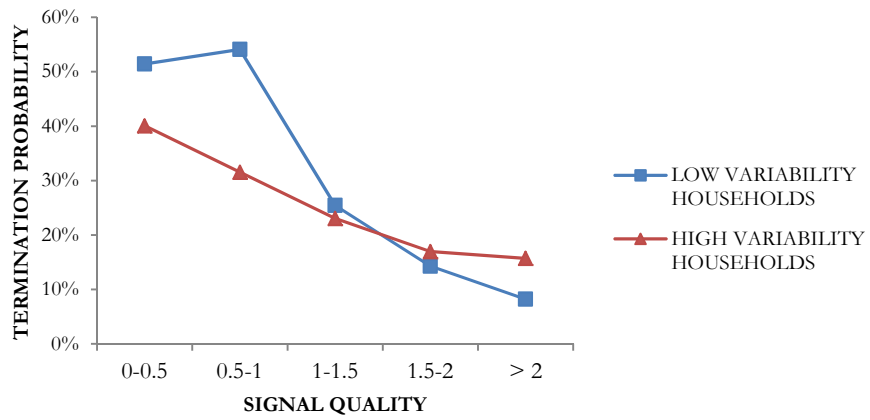


Figure 9: Termination Probabilities when the Overall Variability Changes by 50%



**TERMINATION PROBABILITY WHEN VARIABILITY
DECREASES BY 50%**



Appendices: Model Extensions

Appendix A: Forward-Looking Model with Option Value

Bayesian learning as characterized by the evolution of the posterior belief in Equations 7 and 8 implies that as a household accumulates more experience with the service, its belief about the quality would get more precise (i.e., $\sigma_{Qht}^2 \rightarrow 0$). As a result, in addition to the contemporaneous expected utility from subscribing, the household would also derive an option value from continuing with the service. The option value arises because the household can gather more information about the service by subscribing and hence make a more informed decision in the next period (Hitsch (2006)). The total expected utility (including the option value) that the household would derive from subscribing can be written in the form of a Bellman equation as

$$\begin{aligned}
 W^1(\Omega_{ht}) &= \bar{V}_{1ht} + \epsilon_{1ht} + \delta \int E_{\epsilon}[\max \{W^1(\Omega_{ht+1}), W^0(\Omega_{ht+1})\} | \Omega_{ht}] d(\Omega_{ht+1} | \Omega_{ht}) \\
 &= \bar{V}_{1ht} + \epsilon_{1ht} + \Delta_{1ht} \\
 &= U_{1h}(\Omega_{ht}) + \epsilon_{1ht},
 \end{aligned}
 \tag{A.1}$$

where, Ω_{ht} is the set of state variables (discussed in the estimation section) that influence household h 's utility from the service at time t and δ is the discount factor. The term \bar{V}_{1ht} in Equation A1 captures the per-period flow utility that the household expects to derive from subscribing to the service at time t . The terms $W_h^0(\cdot)$ and $W_h^1(\cdot)$ capture the total utility from terminating and subscribing to the service, respectively, as a function of the state variables and Δ_{1ht} reflects the option value from subscribing to the service with the expectation taken over the distribution of the unobserved component of the utility, ϵ_{1ht} . The integration in Equation (A1) is performed over the distribution of the state variables at time $t+1$ (Ω_{ht+1}) conditional on the observed values at time t (Ω_{ht}) and reflects the stochastic nature of their evolution. The term $U_{1h}(\Omega_{ht})$ is the total observed (by the researcher) component of the utility that household h derives from subscribing to the service given the state variables, Ω_{ht} .

The household will subscribe to the service in a given period if the total expected utility from subscribing (including the option value) during that period exceeds the corresponding expected utility from not subscribing. In the data, we do not observe households reactivating the service after terminating. Consequently, we treat stopping subscription as a terminal decision. Therefore, there is no option value if a household decides to terminate the service. Since we normalize the observed component of the flow utility from terminating the service is 0 (Equation 3), the observed component of the total utility from terminating the service, $U_{1h}(\Omega_{ht})$, is also 0. Therefore, the household would subscribe if

$$W^1(\Omega_{ht}) = U_{1h}(\Omega_{ht}) + \epsilon_{1ht} > \epsilon_{0ht}.
 \tag{A.2}$$

We assume that the household-specific idiosyncratic shocks, ϵ , follow a type I extreme value distribution. Given the above distributional assumption and the normalization of the utility from terminating the service, Equation A1 can be rewritten as

$$\begin{aligned} W^1(\Omega_{ht}) &= \bar{V}_{1ht} + \epsilon_{1ht} + \delta \int \ln \{1 + \exp(U_{1h}(\Omega_{ht+1}))\} |\Omega_{ht}] d(\Omega_{ht+1} | \Omega_{ht}) \\ &= U_{1h}(\Omega_{ht}) + \epsilon_{1ht}. \end{aligned} \tag{A.3}$$

Implications for the Option Value

Option value is typically an increasing function of the per-period flow utility. The extent to which flow utility is an increasing or decreasing function of variability is determined by the tradeoff between the adverse effect via risk aversion and the positive effect of deterring learning, at least for some households. The other driver of option value is the extent to which staying with the service helps the household to make a more informed decision in the future. If the household already has precise knowledge about the average quality it experiences, this benefit is likely to be small. Consequently, households that start with higher uncertainty about the average quality of the service and see larger gains in precision as they accumulate more information are likely to benefit more by staying longer with the service. Thus, one can conjecture that for a given flow utility, high variability is likely to induce customers to stay longer with the service.

Estimation

Given our model specification, there are four sets of state variables: (a) the mean belief about the quality of the service at any period, \bar{Q}_{ht} , (b) uncertainty about this belief $\sigma_{Q_{ht}}^2$, (c) seasonality, and (d) demographic characteristics that act as shifters of the intrinsic preference for the service. Of these, the demographic characteristics are not dynamic but require us to compute the value function for each combination of values that they can take. In order to reduce the computational burden, we dichotomize the five demographic variables based on a median split. Therefore, our estimation involves computing 32 value functions corresponding to all possible combinations of demographic variables. As noted in Equation 1, the seasonality variable captures shifts in the intrinsic propensity to subscribe to the service during different months of the year. Consequently, while this state variable is intrinsically dynamic, its evolution is deterministic.

As in any learning model with option value, the two key state variables in our context are the mean and uncertainty about the unknown entity, the quality of the service. Equation 7 governs the law of motion for the belief about mean quality, \bar{Q}_{ht} . Given the distributional assumptions of normal prior and signals, as in Hitsch (2006), we can characterize the transition density of $\bar{Q}_{ht} | \bar{Q}_{ht-1}$ as $N\left(\bar{Q}_{ht-1}, \frac{\sigma_{Q_{ht-1}}^4}{\sigma_{Q_{ht-1}}^2 + \sigma_h^2}\right)$. Equation 8 characterizes the evolution of the uncertainty about the average quality of the service. Unlike the mean, the uncertainty evolves in a deterministic manner as a function of the signal variance, σ_h^2 . Furthermore, as discussed elsewhere, this uncertainty always declines as a household accumulates more information.

We recast the consumer's decision as a finite horizon problem when we compute the value functions. Specifically, we set $T=30$ as the terminal period and compute the value functions via

backward induction.¹ Together, we have two deterministic state variables (demographic characteristics and seasonality) and two continuous ones. To accommodate the continuous state variables, we evaluate the value functions on a discrete grid of 20 points defined by Chebyshev zeros (Judd (1998)). We used Gaussian quadrature to evaluate the integral in Equation 7. Note that the integration requires us to evaluate the value function outside the set of grid points. To this end, we approximate the value function using a Chebyshev tensor polynomial of degree 4. Similarly, we use Chebyshev interpolation to translate the value functions evaluated at grid points to the actual values of the state variables for individual households over time. We use a discount factor of 0.99 in our estimation.

Results

We present the results from this estimation in Tables A1-A2. The forward-looking model fits the data slightly better than the myopic version (log-likelihood = -2995.162 vs. -3008.51 for the myopic version). The composition of the two segments are similar to those in the myopic version, albeit with some minor differences. In Figure A1, we present the predicted termination probabilities implied by the forward-looking model. As in the case of the myopic model the forward-looking version can replicate the interaction effect that we observe in the data. In Table A3, we present the effect of a 1% increase in signal quality on termination probability. Once again, the results suggest that households that experience high variability are less responsive to improvements in signal quality, both in absolute and percentage terms. Overall, the results from the forward-looking model are consistent with those from the myopic version presented in the paper.

Appendix B: Accounting for the Bounded Nature of Signals

Since the signal quality is bounded between 0 and 2.4, we need to account for this censoring. We can accomplish this by defining an underlying latent signal quality that is unbounded and can thus be treated as normal. Households learn about this latent variable through their experience. However, what matters to their subscription decision is the corresponding signal quality that is bounded between 0 and 2.4. One can impose these bounds by treating the realized signal as a censored outcome of the underlying latent variable. The approach would require us to transform the latent posterior belief into the corresponding expected quality and variance that would eventually affect the subscription decision. We assume that there is an underlying latent signal quality, \tilde{Q}_{ht}^* , that is unbounded, which in turn leads to the realized signal quality, \tilde{Q}_{ht} in Equations 1 and 2. We assume that \tilde{Q}_{ht} is a monotonic transformation of \tilde{Q}_{ht}^* such that

$$\tilde{Q}_{ht} = \begin{cases} \tilde{Q}_{ht}^* & \text{if } 0 < \tilde{Q}_{ht}^* < 2.4 \\ 0 & \text{if } \tilde{Q}_{ht}^* \leq 0 \\ 2.4 & \text{if } \tilde{Q}_{ht}^* \geq 2.4 \end{cases} . \quad (\text{B.1})$$

Further, we assume that at time t, the uncertainty about the household's true latent quality is distributed $N(\tilde{Q}_{ht}^*, \sigma_{\tilde{Q}_{ht}^*}^2)$ and the corresponding mean of the realized signal is \tilde{Q}_{ht} while the variance is $\sigma_{\tilde{Q}_{ht}}^2$. Given the censored normal formulation, we can derive closed form expressions for the moments of the realized signal quality.

¹Recall that our data span 13 periods.

In a similar vein, we assume that the temporal uncertainty is distributed $N(0, \sigma_{*h}^2)$. Therefore, the total uncertainty perceived by the household about the quality of the service in the latent space at time t is the sum of the household's perceived uncertainty about the average quality of the service it receives ($\sigma_{Q^*ht}^2$) and the temporal variation it experiences (σ_{*h}^2), i.e., $\sigma_{ht}^{*2} = \sigma_{Q^*ht}^2 + \sigma_{*h}^2$. Over time, the household accumulates experience with the service and updates its beliefs accordingly in the latent quality space, which in turn would lead to updating of beliefs in the realized quality space. We assume that households are Bayesian learners. Consequently, the mean converges to the true average latent quality received by the household, i.e., Q_h^* and the corresponding uncertainty ($\sigma_{Q^*ht}^2$) converges to 0.

The expression for the expected utility in Equation 2 is

$$\bar{V}_{1ht} = E[\tilde{V}_{1ht}] = \alpha_h + \beta_h \bar{Q}_{ht} - \gamma \bar{Q}_{ht}^2 - \gamma \sigma_{ht}^2 + \psi D_h + \tau S_t + \epsilon_{1ht}. \quad (\text{B.2})$$

In the above equation, \bar{Q}_{ht} is the mean expected quality in the bounded space (i.e., in terms of the number of new movies downloaded) and σ_{ht}^2 is the corresponding total variability as perceived by the household at time t , i.e., σ_{ht}^{*2} . The expressions for \bar{Q}_{ht} and σ_{ht}^2 follow from the properties of the censored normal distribution such that

$$\begin{aligned} \bar{Q}_{ht} &= (\Phi_2 - \Phi_1) \left\{ \bar{Q}_{ht}^* + \frac{\phi_1 - \phi_2}{\Phi_2 - \Phi_1} \sigma_{ht}^* \right\} + 2.4(1 - \Phi_2) \\ \sigma_{ht}^2 &= \left[(\Phi_2 - \Phi_1) \left\{ 1 + \frac{-\bar{Q}_{ht}^* \phi_1 - 2.4 - \bar{Q}_{ht}^* \phi_2}{\sigma_{ht}^* (\Phi_2 - \Phi_1)} - \left(\frac{\phi_1 - \phi_2}{\Phi_2 - \Phi_1} \right)^2 \right\} \sigma_{ht}^{*2} \right] \\ &\quad + \left[\Phi_1 \left(-\bar{Q}_{ht}^* \right)^2 + (1 - \Phi_2) \left(2.4 - \bar{Q}_{ht}^* \right)^2 + (\Phi_2 - \Phi_1) \left\{ \bar{Q}_{ht} - \left(\bar{Q}_{ht}^* + \frac{\phi_1 - \phi_2}{\Phi_2 - \Phi_1} \sigma_{ht}^* \right) \right\}^2 \right] \end{aligned} \quad (\text{B.3})$$

In the above expressions, \bar{Q}_{ht}^* is the expected value of signal quality in the latent space, while $\sigma_{ht}^{*2} = \sigma_{Q^*ht}^2 + \sigma_{*h}^2$ is the total uncertainty as perceived by household h at time t . The terms Φ_1 and Φ_2 correspond to the CDF of the standard normal distribution evaluated at $\frac{-\bar{Q}_{ht}^*}{\sigma_{ht}^*}$ and $\frac{2.4 - \bar{Q}_{ht}^*}{\sigma_{ht}^*}$, respectively. Note that these two points correspond to the normalized lower and upper censoring limits, i.e., 0 and 2.4, respectively. The terms ϕ_1 and ϕ_2 are the corresponding standard normal density functions. The expression of the variance of the bounded signal quality, σ_{ht}^2 in Equation 6 is derived as the sum of the expected value of the variance (the term within the first square brackets) and the variance of the conditional mean (the term within the second square brackets).

As in the literature, we assume that households update their beliefs about the true quality in the latent space of the service in a Bayesian manner. Each household has some prior belief about the latent quality of the service at the time of activation. We assume that this prior belief follows a normal distribution such that $\tilde{Q}_{h0}^* \sim N(Q_0^*, \sigma_{Q^*h0}^2)$. During each period t , household h notes the quality of the service it receives (via the number of new movies in its set-top box) and updates its belief accordingly. As is common in the literature (see, for example Coscelli and Shum (2004)), we assume that the quality of the latent signals that the household receives each period comes from a normal distribution with variance σ_h^{*2} .

Since the model-free evidence suggests that the rate at which a household updates its belief about the quality of the service is inversely related to the variance of the signals it receives, we assume that households know the variability of the latent signals (i.e., σ_h^{*2}) that they receive. We

compute σ_h^{*2} based on the observed mean and variance in the realized quality in the bounded space. We verify that data patterns such as differential learning and the interaction effect (which we had demonstrated based on the variability in the bounded space) exist when we consider the variability in the latent signals.

Given the conjugacy of the prior and signal, the posterior mean belief about the signal quality after t periods of subscribing (i.e., the prior at the beginning of period $t+1$) based on the information it has accumulated till time t can be written as

$$\bar{Q}_{ht}^* = \bar{Q}_{ht-1}^* + \frac{\sigma_{Q^*ht}^2}{\sigma_{Q^*ht}^2 + \sigma_h^{*2}} [Q_{ht}^* - \bar{Q}_{ht-1}^*], \quad (\text{B.4})$$

where, Q_{ht}^* is the actual signal quality in the latent space experienced by the household during period t and the posterior variance at the end of the period t (i.e., prior variance in period $t+1$) is

$$\sigma_{Q^*ht}^2 = \frac{1}{\frac{1}{\sigma_{Q^*ht-1}^2} + \frac{1}{\sigma_h^{*2}}} = \frac{1}{\frac{1}{\sigma_0^2} + \frac{\tau_{ht}}{\sigma_{*h}^2}}. \quad (\text{B.5})$$

In the above expression, τ_{ht} is the number of periods that household h has been with the service as of period t . Since households join the service at different time periods, the number of periods that household h is with the service (τ_h) would not correspond to the number of periods that the service has been active (t). In order to update the mean quality belief in the latent space as in Equation 7, we need to translate the realized signal quality in the 0-2.4 range, Q_{ht} into the corresponding value in the latent space. Given our assumption of a censored normal distribution for the underlying latent variable, when the realized signal quality is in the interior of the (0, 2.4) range, the latent quality would be the same as the realized quality. On the other hand, when the realized signal quality is 0 (2.4), we know that the underlying latent quality is less than 0 (greater than 2.4). Consequently, when the realized signal quality is either 0 or 2.4, we need to integrate Equation 7 over the corresponding range based on the prior distribution, i.e., $N(\bar{Q}_{ht}^*, \sigma_{Q^*ht}^2)$.

Estimation Algorithm

Step 1: In period 0, we need prior beliefs in the latent space, Q_0^* and $\sigma_{Q^*h0}^2$. In the current set up, Q_0 in the bounded space is set to 2.4, while we estimate $\sigma_{Q^*h0}^2$ as a parameter. We need to translate Q_0 to the corresponding value in the latent space, Q_0^* . Alternatively, if we estimate Q_0^* as a parameter directly in the latent space, we don't need to translate it back to the bounded space.

Step 2: The household also knows the variability in the realized signal quality, σ_h^2 . We need to convert this to the latent space, σ_h^{*2} . This translation would depend on the mean of the underlying distribution. We can evaluate this in the first period based on the prior mean, Q_0 .

Step 3: In period t , household receives signal, Q_{ht} in the 0-2.4 range. Convert this to the corresponding quality in the latent space, Q_{ht}^* . We can then use these values in the latent space to compute the updated quality beliefs based on Equation 7. This is fairly easy to do in the logit case.

In the case of the censored normal, we would use the following logic:

$$\begin{aligned} \text{if } 0 < Q_{ht} < 2.4, \text{ then } Q_{ht}^* &= Q_{ht} \\ \text{if } Q_{ht} = 0, \text{ then } Q_{ht}^* &< 0 \\ \text{if } Q_{ht} = 2.4, \text{ then } Q_{ht}^* &> 2.4 \end{aligned} .$$

Therefore, for the realized signal values that lie in the range $0 < Q_{ht} < 2.4$, we can treat the latent quality as the same as the realized quality. For the realized values at the points of censoring, we need to integrate Equation 7 over the distribution of the censored normal. We can accomplish this by making draws from the truncated normal distribution and evaluating Equation 7 for each draw and taking the mean over the draws. We can use Equation 8 to compute the posterior beliefs.

Step 4: We now need to translate the mean and the variance to the corresponding values in the bounded space in order to compute the expected value. We can do this numerically in the logit case. In case of censored normal, we have closed form expressions for this translation.

Results

We present the estimates from this model in Table B1 and B2. The model that accounts for the bounded nature of signals fits the data worse than the version that doesn't (log-likelihood of -3016.184 vs. -3008.511). Nevertheless, the results are fairly consistent across both versions of the model. As seen in Figure B1, the model that accounts for bounded signals can replicate the data patterns found in the data. Moreover, the effect of the 1% increase in termination probabilities are consistent with those in the main model presented in the paper.

References

- COSCELLI, A., AND M. SHUM (2004): "An Empirical Model of Learning and Patient Spillovers in New Drug Entry," *Journal of Econometrics*, 122(4), 213–246.
- HITSCH, G. (2006): "An Empirical Model of Optimal Dynamic Product Launch and Exit Under Demand Uncertainty," *Marketing Science*, 25(1), 25–50.
- JUDD, K. (1998): *Numerical Methods in Economics*. The MIT Press.

Table A1: Estimates from the Forward-Looking Model

PARAMETER	ESTIMATE	STD ERROR
INTERCEPT (SEGMENT 1)	0.618	0.129
INTERCEPT (SEGMENT 2)	-1.169	0.230
SIGNAL QUALITY EFFECT (SEGMENT 1)	0.442	0.057
SIGNAL QUALITY EFFECT (SEGMENT 2)	1.333	0.119
RISK AVERSION (EXP(.))	-1.518	0.085
OCTOBER	-1.605	0.339
NOVEMBER	1.103	0.107
DECEMBER	-0.353	0.173
JANUARY	-0.557	0.160
FEBRUARY	0.087	0.177
MARCH	-0.226	0.133
APRIL	-0.101	0.129
MAY	-0.084	0.165
JUNE	-0.235	0.153
JULY	-0.540	0.140
AUGUST	-0.215	0.180
SALT LAKE CITY	0.057	0.017
SPOKANE	0.076	0.024
INCOME	0.058	0.015
CHILD	-0.001	0.016
% OLD	0.005	0.016
SIGNAL QUALITY PRIOR VARIANCE (SEGMENT 1) (EXP(.))	-2.416	0.067
SIGNAL QUALITY PRIOR VARIANCE (SEGMENT 2) (EXP(.))	-2.451	0.227
TRANSFORMED PRIOR MEAN (SEGMENT 1)	-0.244	0.109
TRANSFORMED PRIOR MEAN (SEGMENT 2)	10.293	0.018
SEGMENT 1 MEMBERSHIP PARAMETER (INTERCEPT)	0.994	0.055
SEGMENT 1 MEMBERSHIP PARAMETER (EFFECT OF INTRINSIC PROPENSITY TO USE THE SERVICE)	-0.210	0.033
LL	-2995.162	

Table A2: Segment Characteristics for the Forward-Looking Model

	SEGMENT 1	SEGMENT 2
SEGMENT SIZE	2312	934
SIGNAL QUALITY PRIOR MEAN (ESTIMATED)	1.055	2.4
SIGNAL QUALITY PRIOR VARIANCE (ESTIMATED)	0.083	0.087
MEAN MONTHLY RENTAL (USAGE)	2.759	4.574
AVG. SIGNAL QUALITY	1.460	1.286
% TERMINATING	4%	65%

Table A3: Effect of a 1% increase in Signal Quality on Termination Rates (Forward-Looking Model)

Signal Quality	Absolute Change in Termination Rate		% Change in Termination Rate (Elasticity)	
	Low-Variability Households	High-Variability Households	Low-Variability Households	High-Variability Households
0-0.5	-0.13%	-0.06%	-0.31%	-0.14%
0.5-1	-0.16%	-0.09%	-0.29%	-0.27%
1-1.5	-0.26%	-0.13%	-1.08%	-0.47%
1.5-2	-0.21%	-0.12%	-1.55%	-0.60%
> 2	-0.22%	-0.05%	-2.06%	-0.29%
Average	-0.21%	-0.11%	-1.42%	-0.43%

Figure A1: Termination Rates Implied by the Model with Option Value

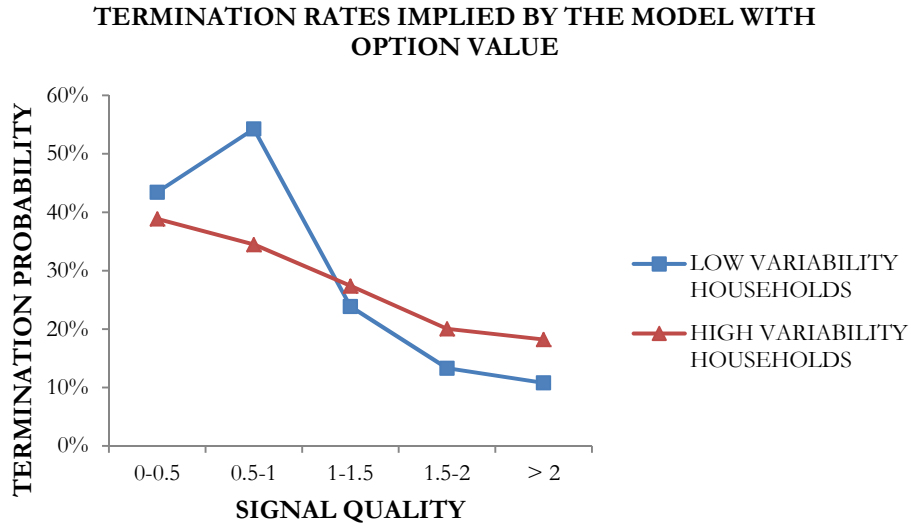


Table B1: Estimates from the Model with Bounded Signals

PARAMETER	ESTIMATE	STD.ERROR	T-VALUE
INTERCEPT (SEGMENT 1)	-4.262	0.035	-122.833
INTERCEPT (SEGMENT 2)	4.342	0.256	16.934
SIGNAL QUALITY EFFECT (SEGMENT 1)	6.586	0.059	111.703
SIGNAL QUALITY EFFECT (SEGMENT 2)	0.713	0.246	2.893
RISK AVERSION (EXP(.))	0.103	0.032	3.178
OCTOBER	-0.271	0.149	-1.817
NOVEMBER	1.022	0.550	1.858
DECEMBER	-0.097	0.284	-0.340
JANUARY	0.006	0.129	0.046
FEBRUARY	0.457	0.244	1.872
MARCH	0.191	0.072	2.434
APRIL	0.189	0.072	1.358
MAY	0.019	0.007	1.923
JUNE	-0.192	0.078	-1.299
JULY	-0.257	0.139	-0.847
AUGUST	0.029	0.010	0.210
SALT LAKE CITY	0.444	0.148	26.218
SPOKANE	0.561	0.303	33.284
INCOME	0.382	0.136	2.802
CHILD	0.049	0.017	2.875
% OLD	0.048	0.017	2.831
SIGNAL QUALITY PRIOR VARIANCE (SEGMENT 1) (EXP(.))	-0.141	0.078	-1.808
SIGNAL QUALITY PRIOR VARIANCE (SEGMENT 2) (EXP(.))	-1.786	0.013	-136.238
PRIOR MEAN (SEGMENT 1)	2.997	0.595	5.042
PRIOR MEAN (SEGMENT 2)	-1.094	0.354	-3.090
SEGMENT 1 MEMBERSHIP PARAMETER (INTERCEPT)	-1.349	0.005	-267.740
SEGMENT 1 MEMBERSHIP PARAMETER (EFFECT OF INTRINSIC PROPENSITY TO USE THE SERVICE)	0.237	0.001	230.483
LL	-3016.184		

Table B2: Segment Characteristics (Model with Bounded Signals)

	SEGMENT 1	SEGMENT 2
SEGMENT SIZE	767	2479
LATENT SIGNAL QUALITY PRIOR MEAN (ESTIMATED)	2.997	-1.094
SIGNAL QUALITY PRIOR VARIANCE	0.87	0.169
MEAN MONTHLY RENTAL (USAGE)	4.578	2.880
AVG. SIGNAL QUALITY	1.21	1.470
% TERMINATING	69%	7%

Table B3: Effect of a 1% increase in Signal Quality on Termination Rates (Model with Bounded Signals)

Signal Quality	Absolute Change in Termination Rate		% Change in Termination Rate (Elasticity)	
	Low-Variability Households	High-Variability Households	Low-Variability Households	High-Variability Households
0-0.5	-0.03%	-0.02%	-0.07%	-0.05%
0.5-1	-0.18%	-0.07%	-0.35%	-0.20%
1-1.5	-0.22%	-0.14%	-0.99%	-0.59%
1.5-2	-0.18%	-0.13%	-1.22%	-0.58%
> 2	-0.18%	-0.14%	-1.29%	-0.57%
Average	-0.17%	-0.11%	-1.03%	-0.45%

Figure B1: Termination Rates Implied by the Model that Accounts for Bounded Signals

