Signaling Quality via Queues

Laurens G. Debo*  
Booth School of Business  
University of Chicago  
Chicago, IL 60637

Christine A. Parlour†  
Haas School of Business  
UC Berkeley  
Berkeley, CA 94720

Uday Rajan‡  
Ross School of Business  
University of Michigan  
Ann Arbor, MI 48109.

November 10, 2010

*Tel: (773) 702-7140, E-mail: laurens.debo@chicagobooth.edu
†Tel: (510) 643-9391, E-mail: parlour@haas.berkeley.edu
‡Tel: (734) 764-2310, E-mail: urajan@umich.edu
Signaling Quality via Queues

Abstract

We show that the length of a queue communicates information about the quality of a product when some consumers are uninformed. In turn, a firm may strategically choose its service rate to signal its quality through the queue. In our setting, a firm may have high or low quality and sells a good to consumers who are heterogeneously informed. The firm may choose a slow or (at a cost) a fast service rate. Consumers arrive and are serviced according to Poisson processes. A consumer who arrives when another consumer is being serviced must join a queue to consume the product. Consumers observe the length of the queue before choosing whether to buy the product. We show that, in equilibrium, informed consumers join the queue if it is below a threshold. The threshold varies with the quality of the good, so an uninformed consumer updates her belief about quality on observing the length of the queue. The strategy of an uninformed consumer has a “hole”: she joins the queue at lengths both below and above the hole, but not at the hole itself. When all consumers are informed, the high-quality firm has a greater incentive to speed up than the low-quality firm. However, the high-quality firm selects a slower service rate than the low-quality firm if there are a lot of queue lengths between the hole in an uninformed consumer’s strategy and the threshold at which informed consumers balk from its queue. Strikingly, if the proportion of informed consumers is low, the high-quality firm may choose the slow service rate even if the technological cost of speeding up is zero. The queue can therefore be a valuable signaling device for a high-quality firm.
1 Introduction

Consumers frequently have to wait before they can consume a product or service. Lines outside nightclubs, rides at amusement parks and waiting lists for new products are part of everyone’s experience. Toys and innovative products, whose value cannot easily be communicated, exhibit similar phenomena. For example, Cabbage Patch Kids in 1983 and Beanie Babies in the 1990s had significant waiting times, and queues formed in front of stores when these products came on the market. Both these goods are mass-produced, and the producers could easily have increased their capacity to reduce consumer waiting times.

Since waiting is costly for consumers, it may appear that the firms were not maximizing profit. Rather, as we show in this paper, in some settings firms have a strategic incentive to manipulate impatient customers’ waiting times to generate greater demand. In other words, the producers benefited from the “buzz” that was created by the high demand.

We develop a model in which a firm sells a good that can be of either high or low quality. Impatient consumers arrive at the market according to a Poisson process. Purchasing the good entails joining a “first-come, first-served” queue. The firm controls the distribution of the queue length by choosing the rate at which it services customers. Some consumers are informed, and know the quality of the good. Others are uninformed, and must infer the quality from the length of the queue. An arriving consumer does not observe the entire history of the game and so does not know how many people arrived before her. She also cannot observe the firm’s service rate choice; she only sees the queue.

The firm chooses either a fast or a slow service rate. We analyze the equilibrium service rate strategies for each of the high and low-quality firms. In particular, we characterize conditions under which the high-quality firm would prefer to slow down (i.e., choose the slow service rate). If the feasible service rates are close to the arrival rate, and the proportion of informed consumers is small, the high-quality firm chooses the slow service rate, even if there is no technological cost to speeding up. Essentially, it uses the length of the queue to signal its quality to uninformed consumers, thereby increasing demand.

As is standard in most queuing models, informed consumers adhere to a “threshold” strategy, where the threshold depends on the quality of the good. That is, they join the queue if it is short enough, and balk at longer queues. In contrast, uninformed agents, who have to infer the quality of the firm from the queue, play a non-threshold strategy. Their Bayesian updating process leads to a “hole” in their joining decision. That is, there is exactly one queue length at which they balk, which lies between the thresholds at which informed consumers balk when faced with low- and high-quality firms. In a pure strategy
equilibrium, they join the queue at both smaller and larger queue lengths. There may also be a mixed strategy equilibrium in which they randomize at some queue lengths below the hole, and join with probability 1 at queue lengths above the hole. Since no uninformed consumer joins the queue at a hole, longer queues can only arise if an informed consumer who knows the firm has high quality were to join at that length. Hence, an uninformed consumer arriving when the queue is above the hole infers that the firm has high quality. The hole thus serves as a natural filter of quality information for uninformed consumers.

The hole in the uninformed consumer’s strategy affects the incentives of each type of firm to speed up (i.e., choose the fast service rate) in different ways. We assume each type of firm has the same profit margin and maximizes the revenue rate per unit time. For the low-quality firm, the queue never rises above the hole, since informed consumers balk at a threshold weakly lower than the position of the hole. Thus, the low-quality firm prefers to keep the queue at lengths below the hole, and naturally has an incentive to choose the fast service rate. The high-quality firm, on the other hand, loses uninformed consumers at the hole, but wins them back at higher queue lengths. If the proportion of informed consumers is high, the high-quality also has an incentive to select a fast service process. However, if there is a large proportion of uninformed consumers, the high-quality firm wishes to avoid having the queue stall at the hole. Enabling the queue to cross the hole allows it to capture the uninformed demand that exists at longer queue lengths. The high-quality firm can do this by selecting the slow service process, even if increasing the service rate comes at a low (or no) cost.

Our paper makes the following contributions. First, we highlight the role of a queue in the process by which uninformed consumers learn about product quality. Second, we provide a complete analysis of the equilibrium in the consumer game in our setting. Third, we show that the informational role played by the queue provides a firm with an incentive to slow down service. More broadly, the queue in our model is simply a device via which a firm communicates information to uninformed consumers about the strength of demand, and hence about their own valuation for the product.

As we show, in Section 3, the inferences consumers can draw from a queue are subtle. If the queue is sufficiently long, uninformed consumers rightly deduce the product has high quality. However, at shorter queue lengths, the posterior probability that a product has high quality may sometimes decrease in the queue length. This happens at low queue lengths if the low-quality firm has a slower service rate than the high-quality firm, and at intermediate queue lengths if the ratio of service rates is less than the proportion of uninformed consumers.
In analyzing the consumer game, we provide conditions for the existence of pure strategy equilibria. We then construct algorithms to identify pure and mixed strategy equilibria, which are detailed in Appendix A. The algorithms provide a constructive proof of the existence of equilibria in the consumer game for any set of parameter values. To illustrate the analytic results and algorithms, we apply them to Example 1 in Section 3, in which we exhibit two pure strategy and three mixed strategy equilibria for given parameter values.

In Section 4, we consider the choice of service rate by each type of firm. With new products, the fraction of informed consumers is often low. There is then a hidden cost to capacity expansion (i.e., to increasing the service rate): It reduces information transmission to uninformed consumers.

It is natural to ask whether a firm may adjust its price to signal the quality of its product. If a firm can choose the price of the product, there are two possibilities. First, both types of firm may pool on price (i.e., choose the same price). Our model corresponds to this case. Second, the high and low quality firms may separate on prices (i.e., choose different prices). In this case, all consumers are effectively informed about quality, since the price reveals quality to uninformed consumers.

There is typically no direct cost to a low-quality firm of simply setting a price equal to that of the high-quality firm. There is an indirect cost: If the price is sufficiently high, the low-quality firm may lose all informed consumers. However, uninformed consumers are still uncertain about the quality of the firm, and will have to learn about quality from the queue. If the proportion of informed consumers is low, the cost to the low-quality firm of imitating the high-quality firm is correspondingly low. Therefore, in this case we expect the low-quality firm to pool on price with the high-quality firm. As a result, our finding that the high-quality firm has an incentive to slow down when there are few informed consumers is robust to firms using the price of the product to signal quality.

Further, in many contexts, firms may have limited ability to provide other signals about quality. For example, Becker (1991) argues that, if a consumer’s utility is increasing in the total number of consumers buying a product, a producer such as a restaurant or Broadway show will prefer to have a low price and excess demand, rather than raise the price. We postulate that, if the good is an experience good that each consumer only purchases once, it is difficult to use price as a signal.¹

Similarly, other signals about quality, such as directly reporting the service rate, may

¹Bagwell and Riordan (1991) provide a two-period model in which a high-quality firm signals with a high price. By the second period, consumers learn the quality of the firm. The repeat purchase feature allows the high-quality firm to separate.
also be infeasible. For example, if a firm reports its service rate, the low-quality firm has an incentive to lie and make the same report as the high-quality firm, so the problem is substantially the same as it is in this paper. However, if the service rate is verifiable, a separating equilibrium may exist in which the quality is known to all consumers. Our model applies to the case in which the service rate is not verifiable.

In the economics literature, Banerjee (1992) and Bikhchandani, Hirshleifer and Welch (1992) show that herd behavior may result; that is, consumers may ignore their own information and adopt the actions of previous agents. There is no signaling by price in these models. Consumers have imperfect information, but observe the actions of all previous consumers. Instead, if each agent observes only her predecessor’s action, beliefs and actions can cycle forever. There are long periods of uniform behavior punctuated by rare switches (Celen and Kariv, 2004). Smith and Sorensen (1997) consider a sequential action model in which agents receive a random sample of the history. They find that the true quality of the good will never be learned, even on the long run. Our model may be interpreted as a model in which each consumer observes a random history that is determined by the queueing process. Chamley (2004) provides a comprehensive review of the herding literature.

In the queueing literature, equilibrium joining strategies are examined by Naor (1969) and the subsequent literature on economic aspects of queueing (see Hassin and Haviv, 2003, for an excellent overview). When there are positive waiting costs, agents play a threshold strategy: they join the queue as long as it is not too long. Beyond some threshold, there is a congestion effect: the waiting costs imply that joining the queue is not worthwhile. The service quality is typically assumed to be common knowledge. Hassin and Haviv (1997) consider a queue with positive externalities in which each agent wishes to “follow the crowd,” and show that threshold behavior continues to obtain.

The literature on service rate decisions with observable queues is sparse (see Hassin and Haviv, Chapter 8). Hassin (1986) considers a firm choosing the profit-maximizing number of single-server facilities, and finds that forcing a firm to make its queue length unobservable reduces social welfare if an observable queue would increase the firm’s profit. Debo and Veeraraghavan (2010) consider a firm with \( N \) servers and no queue, and only uninformed consumers. In the class of equilibria that are analyzed, the price is uninformative, but the congestion level signals quality, as the high and low-quality firms choose different service rates.

We set up our model in Section 2. Equilibrium in the consumer game is analyzed in Section 3. Section 4 considers the selection of service rate by each type of firm. Section 5 provides some concluding remarks. The appendix provides algorithms for constructing
pure and mixed strategy equilibria in the consumer game and the proofs of all results.

2 Model

We consider the market for an experience good, which may be a physical product or a service. The good may have a high \((h)\) or low \((\ell)\) quality. In addition, firms choose the speed at which they provide the good to their customers (“the service rate”).

The sequence of events is as follows. At stage 0, nature chooses the type of the firm. At stage 1, the firm (which privately knows its own type) chooses a Poisson service rate \(\mu \in \{\mu, \bar{\mu}\}\). The mean service time per consumer is then \(\frac{1}{\mu}\). Without loss of generality, we normalize the cost of choosing rate \(\bar{\mu}\) to zero; let \(k \geq 0\) be the technological cost of choosing rate \(\bar{\mu}\). We assume that \(k\) does not depend on the quality of the firm and represents a non-verifiable expenditure of resources. Once chosen, the service rate is held constant for the rest of the game. For \(\theta \in \{h, \ell\}\), let \(\beta_\theta\) denote the probability that a firm of type \(\theta\) chooses service rate \(\bar{\mu}\).

At stage 2, after the firm has chosen its service rate, risk-neutral consumers arrive at the market according to a Poisson process with parameter \(\Lambda\). A proportion \(q\) of consumers is informed, and knows the quality of the good. The remaining proportion \(1 - q\) is uninformed. Uninformed consumers have a prior belief that the good is high quality with probability \(p\). The utility an agent obtains from purchasing and consuming a good of quality \(\theta\) is \(v_\theta\), with \(v_h > v_\ell\). This utility is net of the good’s price, which is not explicitly modeled.

The firm cannot communicate its service rate to the consumers. Hence, upon arrival a consumer does not observe the service rate of the firm but does observe the number of people waiting to be served by the firm. If agents arrive faster than they are serviced, they form a queue. The queue is served on a first-come first-served basis. A consumer suffers a disutility \(c > 0\) per unit of time that she has to wait to obtain the good, starting from her initial arrival to the market.

Each consumer takes an action \(a \in \{\text{join}, \text{balk}\}\), where \(a = \text{join}\) is the decision to acquire the good, or to join the queue. Once she joins the queue, she may not renege; i.e., she cannot leave until she has been served. Joining the queue is therefore synonymous with consuming the good. If she chooses \(a = \text{balk}\) (i.e., not acquire the good), she obtains a reservation utility of zero. Thus, each consumer will join the queue only if the expected utility from joining exceeds zero.

We assume that an informed consumer obtains a positive utility from consuming the low-quality good whenever she finds no one else ahead of her, even if the firm has chosen the
slow service rate. This ensures that the low-quality firm can earn a strictly positive profit. Otherwise, the low-quality firm may exit the market, in which case consumers would know any good offered was of high quality.

**Assumption 1** \( \nu_\ell > c/\mu \).

In what follows, we characterize Markov perfect Bayesian equilibria of the game. Throughout this paper, we restrict attention to the case in which each firm type plays a pure strategy; that is, \( \beta_\theta \in \{0, 1\} \) for each \( \theta \). We refer to the service rate of firm \( \theta \) as \( \mu_\theta \). In the consumer continuation game (i.e., once the firm has chosen its service rate), we consider equilibria in which the strategy of a consumer depends only on whether she is informed and on the length of the queue when she arrives at the market, but not on the exact time at which she arrives or on the number of consumers that have preceded her. Thus, the equilibria are both symmetric across consumers and time-invariant.

### 2.1 Consumers’ Objective Functions

Consider an informed consumer who knows the firm has type \( \theta \) and finds that there are already \( n \) consumers in the queue (including the consumer currently being served) when she arrives at the market. Suppose she joins the queue. The expected service time for each customer is \( \frac{1}{\mu_\theta} \). Therefore, the total expected time to service the \( n \) consumers already in the queue and the new arrival is \( \frac{n+1}{\mu_\theta} \). Hence, her expected utility from joining the queue is 

\[
\ell_i(n, \theta, \mu_\theta) = \nu_\theta - (n + 1) \frac{c}{\mu_\theta}.
\]

Suppose an uninformed consumer who finds \( n \) consumers waiting in line ascribes a posterior probability \( \gamma(n) \) to the firm being of high quality. Then, her expected utility if she joins the queue when there are already \( n \) consumers waiting in line is 

\[
\ell_u(n, \gamma, \mu_h, \mu_\ell) = (1 - \gamma(n))\ell_i(n, \ell, \mu_\ell) + \gamma(n)\ell_i(n, h, \mu_h).
\] (1)

Define \( \bar{N} = \lfloor \nu_h \frac{\bar{v}}{\ell} \rfloor \), where \( \lfloor x \rfloor \) is the largest integer less than or equal to \( x \). Then, \( \bar{N} \) is the largest queue length that can be observed in any equilibrium of the game, because no informed consumer will join a queue of \( \bar{N} \) or longer, so that uninformed consumers will also balk at these queue lengths. Let \( \mathcal{N} = \{0, 1, \cdots, \bar{N}\} \).

In general, a consumer may play a mixed strategy, and join the queue with some probability between 0 and 1. Thus, a mixed strategy for an informed agent is a mapping \( \sigma_i : \{h, \ell\} \times \mathcal{N} \to [0, 1] \), and a mixed strategy for an uninformed agent is a mapping \( \sigma_u : \mathcal{N} \to [0, 1] \). The overall consumer strategy profile may then be represented as 

\[
\sigma = (\sigma_i, \sigma_u).
\]
2.2 Firm’s Objective Function

Each type of firm chooses its own service rate \( \mu_\theta \) to maximize its expected payoff, given \( \sigma \). Recall that the price is not a choice variable for the firm; rather the price is implicit in the values \( v_h, v_\ell \). We therefore assume that each type of firm earns a revenue of \( r \) per consumer that it serves. We assume that \( r \) does not depend on the quality of the firm. The throughput per unit time in the long-run depends on the stationary distribution over queue length induced by \( \sigma \) and \( \mu \).

Let \( X_\theta(t \mid \mu, \sigma) \) be a random variable that denotes the number of consumers in the queue at some time \( t \geq 0 \), given that the firm has quality \( \theta \), the service rate of the firm is \( \mu \), and the strategy profile of consumers is denoted by \( \sigma \). The arrival and service processes, together with the strategies of informed and uninformed consumers, imply that \( X_\theta(t \mid \mu, \sigma) \) follows a Markov process. Since the queue length process is irreducible and aperiodic, it has a stationary distribution with support contained in \( \mathbb{N} \). Let \( \pi_\theta(n, \mu, \sigma) \) denote the stationary probability of observing \( n \) consumers in the queue. That is, \( \pi_\theta(n, \mu, \sigma) = \lim_{t \to \infty} \text{Prob}(X_\theta(t \mid \mu, \sigma) = n) \). By the PASTA (Poisson Arrivals See Time Averages) property of queueing systems (see Wolff, 1982), in the limit a newly-arriving agent will be faced with a distribution over queue length equal to the stationary distribution.

In general, the distribution defined by \( \pi_h(\cdot) \) will be different than the distribution defined by \( \pi_\ell(\cdot) \) for two reasons. First, even if both types of firm choose the same service rate, informed consumers follow a strategy that varies with firm type. Second, the two types of firm may adopt different service rate strategies. Since the distributions over queue lengths induced by the two firms are different, uninformed consumers update their prior beliefs about firm quality on observing the length of the queue when they arrive.

As is standard, the number of consumers in the queue at any given point of time can be represented as a birth-death process, from which the stationary distribution over queue length may be derived. Suppose a firm has type \( \theta \). Let \( s_\theta(n, \sigma) \) be the probability that a consumer (of either type) who plays strategy \( \sigma \) joins a queue that already has \( n \) agents. Then,

\[
s_\theta(n, \sigma) = (1 - q)\sigma_u(n) + q\sigma_i(\theta, n) \quad \text{for } \theta \in \{h, \ell\}.
\]

If there are \( n \) agents already in the queue or in service, the rate at which a new agent joins is \( \Lambda s_\theta(n, \sigma) \). The rate at which agents leave the queue is equal to the service rate \( \mu \). The stationary probability for each queue length \( n \), \( \pi_\theta(n, \mu, \sigma) \), is then derived in a standard manner from the resulting flow balance equations.
Lemma 1 Suppose the firm has type $\theta \in \{h, \ell\}$ and chooses service rate $\mu$, and all agents follow the strategy profile $\sigma$. Then, the stationary probabilities over different queue lengths are given by:

$$\pi_{\theta}(0, \mu, \sigma) = \frac{1}{1 + \sum_{n=1}^{N} \left(\frac{\Lambda}{\mu}\right)^{n} \prod_{j=0}^{n-1} s_{\theta}(j, \sigma)}$$

(3)

$$\pi_{\theta}(n, \mu, \sigma) = \pi_{\theta}(0, \mu, \sigma) \left(\frac{\Lambda}{\mu}\right)^{n} \prod_{j=0}^{n-1} s_{\theta}(j, \sigma).$$

(4)

Let $R_{\theta}(\mu, \sigma)$ be the firm $\theta$’s revenue per unit of time when the consumer strategy profile is $\sigma$ and the firm chooses service rate $\mu$. Under these conditions, the firm is busy with stationary probability $1 - \pi_{\theta}(0, \mu, \sigma)$. Since the expected time to service a consumer is $\frac{1}{\mu}$, the revenue per unit of time is $r\mu$. Hence, the expected revenue of firm $\theta$ when it chooses service rate $\mu$ is $R_{\theta}(\mu, \sigma) = r\mu(1 - \pi_{\theta}(0, \mu, \sigma))$.

2.3 Definition of Equilibrium

We consider Markov-perfect Bayesian equilibria of the game: both types of firm and both types of consumer maximize their respective expected payoffs, and, where possible, the beliefs of uninformed consumers must obey Bayes’ rule. Note that we define equilibria in which firms play pure strategies.

Definition 1 A Markov-perfect Bayesian equilibrium in which each type of firm plays a pure strategy is defined by a 4-tuple $(\mu_{h}, \mu_{\ell}, \sigma, \gamma)$ that satisfies the following properties.

(i) Each type of firm maximizes its expected payoff:

$$\mu_{\theta} \in \arg \max_{\mu \in \{\mu, \bar{\mu}\}} R_{\theta}(\mu, \sigma) - k1_{\{\mu = \bar{\mu}\}} \text{ for } \theta \in \{h, \ell\},$$

(5)

where $1_{\{\mu = \bar{\mu}\}}$ is an indicator variable that takes the value one if $\mu = \bar{\mu}$ and zero otherwise.

(ii) Each type of consumer maximizes her expected utility:

$$\sigma_{i}(\theta, n) \in \arg \max_{\tilde{\sigma} \in [0, 1]} \tilde{\sigma}w_{i}(n, \mu_{\theta}, \beta) \text{ for } \theta \in \{h, \ell\} \text{ and each } n \in \mathcal{N},$$

(6)

$$\sigma_{u}(n) \in \arg \max_{\tilde{\sigma} \in [0, 1]} \tilde{\sigma}w_{u}(n, \gamma, \mu_{h}, \mu_{\ell}) \text{ for each } n \in \mathcal{N}. $$

(7)
Where possible, the beliefs \(\gamma\) are derived from the strategies using Bayes’ rule. That is, whenever the denominator is strictly positive,

\[
\gamma(n) = \frac{p \pi_h(n, \mu_h, \sigma)}{p \pi_h(n, \mu_h, \sigma) + (1 - p) \pi_\ell(n, \mu_\ell, \sigma)}.
\] (8)

Suppose firms have chosen their service rates at stage 1. At stage 2, consumer arrive; we refer to this stage as the “consumer game.” In this game, a subset of consumers does not know the type of the firm, and so is playing a game with incomplete information. As is standard, we refer to the consumer game as a “continuation game.” Each consumer has a finite set of actions and the set of feasible states (i.e., queue lengths observed by a newly-arrived consumer) is also finite. As equilibria exist in finite games, it is immediate that an equilibrium (possibly in mixed strategies) exists in the consumer game, regardless of the service rate chosen by each type of firm. Further, each type of firm has only two service rates to choose from, so it is immediate to conclude that an equilibrium (again, possibly in mixed strategies) exists in the firms’ game. Hence, existence of an equilibrium in the overall game follows. That is, for any values of the parameters \((\Lambda, \mu_h, \mu_\ell, p, q, c, v_h, v_\ell)\), there is at least one Markov-perfect Bayesian equilibrium in the overall game.

In what follows, when we consider the firms’ choice of service rates, we focus on pure strategy equilibria. We discuss existence of pure strategy equilibria in the consumer game in the next section. Equilibria in which firms mix over service rates are non-generic in \(k\). That is, for firm \(\theta\) to mix, it must be that \(k = R_\theta(\bar{\mu}, \sigma) - R_\theta(\mu, \sigma)\), so that a small increase (decrease) in \(k\) leads to the firm strictly preferring \(\mu\) (\(\bar{\mu}\)).

3 Equilibria in the Consumer Game

We first consider equilibria in the consumer game that arises after each firm type \(\theta\) has chosen its service rate strategy. Let \(\mu_\theta\) denote the service rate chosen by the firm of type \(\theta\).

The optimal strategy of an informed consumer is straightforwardly characterized as a threshold strategy, which is standard in the queueing literature (see, for example, Naor, 1969, and Hassin and Haviv, 2003). She knows the quality of the good, and hence joins the queue whenever it is short enough. That is, an informed consumer joins as long as the queue length \(n < n_\theta = \lceil v_\theta \frac{\mu_h}{c} \rceil\). Further, if \(n_\theta \frac{\mu_h}{\mu_\theta} < v_\theta\), it is a unique best response for an informed consumer to balk at \(n_\theta\). For the rest of the paper, we assume that this condition holds.
For analytic convenience, we further assume that $n_\ell < n_h$. That is, regardless of the service rate strategies chosen by the firms, the threshold at which informed consumers balk from the high-quality firm strictly exceeds the corresponding threshold for the low-quality firm. Part (ii) of Assumption 2 essentially implies that the range in consumer valuations $(v_h - v_\ell)$ has a greater impact on the consumer’s choice than the range in service rates $(\mu - \bar{\mu})$. If the assumption is satisfied, then, for any pair of strategies $(\mu_h, \mu_\ell)$ chosen by the different types of the firm, it will be the case that $n_h > n_\ell$.

**Assumption 2**

(i) $n_\theta \frac{\bar{v}_h}{\mu_\theta} < v_\theta$, for each $\theta = h, \ell$.
(ii) $[v_\ell(\bar{\mu}/c)] < [v_h(\mu/c)]$.

We show that the equilibrium strategy of an uninformed consumer is typically not a threshold strategy. Rather, it is characterized by a “hole.” That is, there exists exactly one queue length $\hat{n}$ between $n_\ell$ and $n_h$ at which the uninformed consumer does not join the queue. At every other queue length between 0 and $n_h$, she follows the strategy of an informed consumer who knows the firm has high quality; that is, she joins the queue.

**Proposition 1** *In any pure strategy equilibrium of the consumer game, there exists a $\hat{n} \in \{n_\ell, n_\ell + 1, \cdots, n_h\}$ such that $\sigma_u(\hat{n}) = 0$ and $\sigma_u(n) = \sigma_i(h, n)$ for all $n \in \mathcal{N} \setminus \hat{n}$.*

To understand the existence of a hole in the uninformed consumer’s strategy, consider her Bayesian updating problem. If the queue length she observes is below $n_\ell$, her beliefs are irrelevant to her optimal action: even a fully informed agent would join such a queue, so it is optimal for the uninformed agent to also join. However, at $n_\ell$, informed agents no longer join the queue for the low-quality firm. Thus, any queue length strictly above $n_\ell$ the beliefs of the uninformed consumer over the quality of the firm are critical in determining her action. As long as she believes it is sufficiently likely the firm has high quality, she will join. Suppose there is a queue length $\hat{n} \geq n_\ell$ at which uninformed consumers do not join the queue. Then, any queue length strictly above $\hat{n}$ can only be reached if an informed consumer joins the queue at $\hat{n}$. But informed consumers have perfect information, so that observing a queue of length strictly greater than $\hat{n}$, an uninformed consumer must believe the firm has high quality with probability 1. Thus, the uninformed consumer again joins the queue at lengths between $\hat{n} + 1$ and $n_h - 1$, leading to a hole at $\hat{n}$.

Next, consider the posterior beliefs of an uninformed consumer.
Lemma 2 Suppose each type of firm plays a pure strategy, with type $\theta$ choosing service rate $\mu_\theta \in \{\bar{\mu}, \mu\}$. Suppose further that a pure strategy equilibrium $\sigma$ results in the consumer game, in which informed consumers balk at $n_\theta$ for a type $\theta$ firm, and uninformed consumers join at every queue length $n \leq n_\theta$ except at $\hat{n} \geq n_{\ell}$. Then, the uninformed consumer’s posterior that the firm has high quality is given by $\gamma(n) = \frac{p}{p+(1-p)\phi(n,\mu_h,\mu_{\ell},\sigma)}$, where

$$\phi(n,\mu_h,\mu_{\ell},\sigma) = \begin{cases} 
\phi(0,\mu_h,\mu_{\ell},\sigma) \left(\frac{\mu_h}{\mu_{\ell}}\right)^n & \text{if } n \in \{0, \cdots, n_{\ell}\} \\
\phi(0,\mu_h,\mu_{\ell},\sigma) \left(\frac{(1-q)\mu_h}{\mu_{\ell}}\right)^n & \text{if } n \in \{n_{\ell}+1, \cdots, \hat{n}\} \\
0 & \text{if } n \in \{\hat{n}+1, \cdots, n_h\}
\end{cases}$$

with $\phi(0,\mu_h,\mu_{\ell},\sigma) = \frac{\pi_\ell(0,\mu_h,\sigma)}{\pi_\ell(0,\mu_h,\sigma)}$.

The expression $\phi(\cdot)$ may be interpreted as the likelihood the firm has low quality given the queue length $n$, but before the prior belief $p$ is taken into account. At queue lengths $n$ between 0 and $n_{\ell} - 1$, both informed and uninformed consumers join. That is, $s_\theta(n,\sigma) = 1$ for each $\theta$. From equation (4) in Lemma 1, it follows that the probability of observing $n \in \{1, \cdots, n_{\ell}\}$ is proportional to $(\Lambda/\mu_\theta)^n$. The likelihood the firm has low quality based only on the queue length, $\phi(\cdot)$, therefore depends on whether $\Lambda/\mu_h$ is greater or less than on $\Lambda/\mu_{\ell}$ and hence on the ratio $\frac{\mu_h}{\mu_{\ell}}$. It increases (decreases) in $n$ when the high-quality firm is faster (slower) than the low-quality firm.

When the queue length $n$ is in the range $n_{\ell}, \cdots, \hat{n}$, the rate of change of an uninformed consumer’s posterior belief depends critically on the ratio $\frac{(1-q)\mu_h}{\mu_{\ell}}$. The intuition is as follows. At the length $n_{\ell}$, only uninformed consumers join the queue for the low-quality firm. Thus, the arrival rate for firm $\ell$ falls discretely to $(1-q)\Lambda$, whereas the arrival rate for the high-quality firm remains $\Lambda$. Consumers leave firm $\theta$ at the rate $\mu_\theta$. From Equation (4) in Lemma 1, it follows that the probability of observing $n \in \{n_{\ell}, \cdots, \hat{n}\}$ is proportional to $((1-q)\Lambda/\mu_{\ell})^{n-n_{\ell}}$ for the low-quality firm and to $(\Lambda/\mu_h)\eta_{n-n_{\ell}}$ for the high-quality firm. Hence, beyond the queue length $n_{\ell}$, if $(1-q)\Lambda/\mu_{\ell} > \Lambda/\mu_h$, a higher queue length implies it is less likely that the firm has high-quality. As a result, the consumer’s expected consumption value for the good falls at higher queue lengths. Conversely, if $(1-q)\Lambda/\mu_{\ell} < \Lambda/\mu_h$, the likelihood a firm has high quality increases in queue length beyond $n_{\ell}$. The relationship between $\mu_{\ell}$ and $(1-q)\mu_h$ plays an important role in our subsequent analysis in this section.

When will a pure strategy equilibrium exist in the consumer game? We provide two results on this below. First, we consider the existence of a pure strategy equilibrium in which the hole in the uninformed consumer’s strategy is at any queue length between $n_{\ell}$ and $n_h$. Then, we turn to restrictions on the parameters that ensure a pure strategy equilibrium.
with a hole at a generic queue length \( n \). In the next section, we show that holes at low queue lengths (i.e., close to \( n_\ell \)) are of particular interest in terms of the high-quality firm’s decision on service rate. Therefore, at the end of the section, we highlight the condition under which there exists a pure strategy equilibrium in the consumer game with a hole exactly at the queue length \( n_\ell \).

We first consider the existence of any pure strategy equilibrium. In Appendix A.1, we provide an algorithm to construct pure strategy equilibria. The algorithm relies on a mapping that maps a conjectured likelihood ratio at the empty queue to pure strategy best responses of informed and uninformed consumers, from which the likelihood ratio at the empty queue can be constructed. A fixed point of this mapping represents a pure strategy equilibrium. In particular, it reduces the problem of finding an \(|\bar{N}|\)-dimensional fixed point (in strategies) to finding a single-dimensional fixed point (based on the likelihood ratio at the empty queue). Analyzing this mapping, we are able to derive conditions under which it will have at least one fixed point.

As in the algorithm, define

\[
\Phi(n_\ell) = \frac{\sum_{m=0}^{n_\ell} \left(\frac{\Lambda}{\mu_h}\right)^m + q \sum_{m=n_\ell+1}^{n_h} \left(\frac{\Lambda}{\mu_h}\right)^m}{\sum_{m=0}^{n_\ell} \left(\frac{\Lambda}{\mu_\ell}\right)^m}.
\]

We show that the ratio of service rates, \( \mu_\ell/\mu_h \) helps determine the number of pure strategy equilibria in the game.

**Proposition 2** (i) Suppose \( \mu_\ell < (1 - q)\mu_h \) and \( \left(\frac{\mu_\ell}{\mu_h}\right)^{n_\ell+1} < \Phi(n_\ell) \). Then, the consumer game has at least one pure strategy equilibrium.

(ii) Conversely, if \( \mu_\ell > (1 - q)\mu_h \) and \( \left(\frac{\mu_\ell}{\mu_h}\right)^{n_\ell+1} > \Phi(n_\ell) \), the consumer game has at most one pure strategy equilibrium.

In case (ii) of the proposition, the consumer game may not have any pure strategy equilibrium. As argued in Section 2.3, an equilibrium always exists, so in this case there must be a mixed strategy equilibrium. In Section A.2, we provide an algorithm to determine any mixed strategy equilibrium in the game.

Observe that the conditions in Proposition 2 do not depend on the parameter \( p \), the prior probability the firm has high quality.\(^2\) By imposing conditions on \( p \), we can be more precise about the nature of the pure strategy equilibria in the game. In particular, for each

\(^2\)Since \( n_\theta \) depends on \( v_\theta, \mu_\theta \) and \( c \) for each \( \theta \in \{\ell, h\} \), the conditions use all the other parameters of the model.
queue length $n$ we identify upper and lower thresholds for $p$, the prior probability the firm has high quality. These thresholds $\bar{p}(n)$ and $\underline{p}(n)$ depend on all other parameters in the game, $(\Lambda, \mu_h, \mu_\ell, q, c, v_h, v_\ell)$. For convenience, we suppress this dependence in the notation.

If $p \in (\underline{p}(n), \bar{p}(n))$, there exists a a pure strategy equilibrium in the consumer game with a hole exactly at the queue length $n$ if either an added condition on the parameters is satisfied (condition (a)) in Proposition 3 below) or $n$ is close to either $n_\ell$ or $n_h$. However, if condition (a) below is violated and $n$ is sufficiently far from both $n_\ell$ and $n_h$, a pure strategy equilibrium with a hole at $n$ cannot be supported, regardless of the value of $p$.

**Proposition 3** There exist threshold queue lengths $n \geq n_\ell$ and $\bar{n} \leq n_h$ and threshold probabilities $\underline{p}(n), \bar{p}(n)$ for each $n \in \{n_\ell, \cdots, n_h\}$ such that $0 \leq \underline{p}(n) < \bar{p}(n) \leq 1$, $\underline{p}(n_\ell) = 0$, $\bar{p}(n_h) = 1$, and:

(i) If $p \in [\underline{p}(n), \bar{p}(n)]$ and either (a) $\frac{\mu_\ell}{(1-q)\mu_h} < \frac{v_h\mu_h-v_\ell\mu_\ell+2c}{v_h\mu_h-v_\ell\mu_\ell-2c}$ or (b) $n \in \{n_\ell, \cdots, \bar{n}\} \cup \{\bar{n}, \cdots, n_h\}$, there is a pure strategy equilibrium in the consumer game with a hole at $n$.

(ii) If $\frac{\mu_\ell}{(1-q)\mu_h} < \frac{v_h\mu_h-v_\ell\mu_\ell+2c}{v_h\mu_h-v_\ell\mu_\ell-2c}$ and $n \in \{\bar{n}+1, \cdots, \bar{n}-1\}$, there does not exist a pure strategy equilibrium in the consumer game with a hole at $n$ for any $p \in [0, 1]$.

Importantly, there always exist parameter values which there is an equilibrium in which an uninformed consumer’s strategy has a hole at $n_\ell$. In such an equilibrium, uninformed consumers do not join at the same queue length at which informed consumers balk when the firm has low quality. In Section 4, we focus on this equilibrium.

The thresholds $p$ and $\bar{p}$ are increasing in $n$. However, they may overlap (i.e., it is possible that $\underline{p}(n+1) < \bar{p}(n)$). Therefore, for the same parameter values, there may be multiple pure strategy equilibria in the consumer game. For completeness, we provide an example in which we compute all equilibria in the consumer game. In the example, we find two pure strategy and three mixed strategy equilibria. The example is non-generic in the sense that small perturbations in any of the parameters qualitatively preserve each of the equilibria. Therefore, the example also highlights that multiple equilibria may occur in this model.

**Example 1**

Let $p = 0.5$, $q = 0.45$, $v_h = 1.70$, $v_\ell = 0.75$, $c = 0.35$, $\Lambda = 1$, $\mu_h = 1.40$, and $\mu_\ell = 0.75$. Then, $n_h = 6$ and $n_\ell = 1$. That is, in every equilibrium, the informed consumer joins the queue at all queue lengths less than 6 if the firm has high quality and only at the queue length 0 if the firm has low quality. An equilibrium can then be characterized by the strategy of an uninformed consumer.
Consider the condition in Proposition 3(i)(a): we have \( \frac{\mu_c}{(1-q)\mu_h} = 0.974 < \frac{v_h\mu_h - v_h\mu_c + 2c}{v_h\mu_h - v_h\mu_c - 2c} = 2.2528 \). Therefore, for every \( n \in \{1, 2, \cdots, 6\} \), there exists a range of \( p \) such that there is a pure strategy equilibrium in which the uninformed consumer’s strategy has a hole at \( n \).

Using the expressions shown in the proof of Proposition 3, we compute the probability thresholds \( p(n) \) and \( \bar{p}(n) \). The thresholds are shown in Table 1. As can be observed, when \( p = 0.5 \), there exist two pure strategy equilibria in this game:

(P1) The hole in the uninformed consumer’s strategy is at \( n = 2 \), so the uninformed consumer joins at queue lengths \( n = 0, 1, 3, 4, 5 \).

(P2) The hole in the uninformed consumer’s strategy is at \( n = 3 \), so the uninformed consumer joins at queue lengths \( n = 0, 1, 2, 4, 5 \).

<table>
<thead>
<tr>
<th>Queue length ( n )</th>
<th>( p(n) )</th>
<th>( \bar{p}(n) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0.2245</td>
</tr>
<tr>
<td>2</td>
<td>0.1858</td>
<td>0.5120</td>
</tr>
<tr>
<td>3</td>
<td>0.4812</td>
<td>0.6895</td>
</tr>
<tr>
<td>4</td>
<td>0.6735</td>
<td>0.8237</td>
</tr>
<tr>
<td>5</td>
<td>0.8166</td>
<td>0.9302</td>
</tr>
<tr>
<td>6</td>
<td>0.9280</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 1: Probability Thresholds in Example 1

Finally, we apply the algorithm to construct mixed strategy equilibria in Appendix A.2 to this example. Details of the application are illustrated in Section A.3.

We find three mixed strategy equilibria. In each of these, the uninformed consumer joins the queue with probability 1 at lengths \( n = 0, 1 \) and 3, and with probability 0.0924 at \( n = 2 \), and balks at any queue length greater than or equal to 6. The equilibria are further characterized as follows:

(M1) The uninformed consumer joins the queue with probability 0.0866 at \( n = 4 \) and probability 1 at \( n = 5 \).

(M2) The uninformed consumer joins the queue with probability 0.4934 at \( n = 4 \) and balks at \( n = 5 \) or higher.

(M3) The uninformed consumer joins the queue with probability 0.1843 and \( n = 4 \) and probability 0.7588 at \( n = 5 \).

Observe that the algorithm is sophisticated enough to identify multiple queue lengths.
at which the uninformed consumer randomizes. For example, in equilibrium (M3), the
uninformed consumer mixes between joining and balking at queue lengths 2, 4 and 5.

4 Equilibrium Choice of Service Rates

When firms choose their service rates, they take the consumer strategy as given. For any
c consumer strategy \( \sigma = (\sigma_i, \sigma_u) \), let \( \hat{n}(\sigma) = \min\{n \mid \sigma_u(n) = 0\} \) denote the queue length at
which \( \sigma_u \) has a hole. If the firm has low-quality, its queue never extends beyond \( \hat{n}(\sigma) \). At
queue lengths below \( n_\ell \), both informed and uninformed agents join the queue, and at queue
lengths between \( n_\ell \) and \( \hat{n}(\sigma) - 1 \), only uninformed agents join. In contrast, the high quality
firm has all agents joining the queue at all queue lengths below \( n_h \), except for the queue
length \( \hat{n}(\sigma) \), at which only informed agents join.

For each \( \theta = h, \ell \), let \( \tilde{n}_\theta(\sigma) = \min\{n \mid \sigma_i(\theta, n) = 0\} \) be the first queue length at
which informed consumers balk. We say a consumer pure strategy \( \sigma \) satisfies the necessary
conditions of equilibrium in the consumer game if \( \tilde{n}_h(\sigma) \geq \tilde{n}_\ell(\sigma) \) and \( \hat{n}(\sigma) \in \{\tilde{n}_\ell(\sigma), \tilde{n}_\ell(\sigma) + 1, \cdots, \tilde{n}_h(\sigma)\} \). As part (ii) of the next Lemma shows, conditional on consumers playing a
pure strategy, both types of firm would prefer to have the hole in the uninformed consumer’s
strategy at as high a queue length as possible.

Lemma 3 Fix a pure strategy for consumers, \( \sigma \), that satisfies the necessary conditions for
equilibrium in the consumer game. Suppose a firm of type \( \theta \in \{h, \ell\} \) chooses service rate \( \mu \),
and consumers play \( \sigma \). Then,

(i) The expected revenue per unit of time for each type of firm is

\[
R_h(\mu, \sigma) = r\mu \left( 1 - \frac{1}{\sum_{j=0}^{\hat{n}(\sigma)} (\Lambda/\mu)^j + q \sum_{j=\hat{n}(\sigma)+1}^{\tilde{n}_h(\sigma)} (\Lambda/\mu)^j} \right),
\]

\[
R_\ell(\mu, \sigma) = r\mu \left( 1 - \frac{1}{\sum_{j=0}^{\hat{n}(\sigma)} (\Lambda/\mu)^j + \sum_{j=\tilde{n}_\ell(\sigma)+1}^{\hat{n}(\sigma)} (1-q)j^{-\hat{n}(\sigma)}(\Lambda/\mu)^j} \right).
\]

(ii) Fixing \( \tilde{n}_h(\sigma) \) and \( \tilde{n}_\ell(\sigma) \), the expected revenue \( R_\theta \) increases in \( \hat{n}(\sigma) \) for \( \theta \in \{h, \ell\} \).

Therefore, all else equal, each type of firm will prefer that the hole in an uninformed
c consumer’s strategy, \( \hat{n} \), occur at a high rather than low queue length. The intuition for
our results in this section is that, if the hole occurs at a low queue length (specifically, \( n_\ell \)),
the high-quality firm has an incentive to slow down (i.e., choose the slow service rate \( \mu_\ell \)) to
ensure that a relatively large proportion of time is spent at queue lengths above the hole.
One complication that arises is that speeding up (i.e., choosing the fast service rate $\mu$) increases $n_h$, the threshold at which informed consumers balk, thereby potentially increasing the revenue of the firm. We begin with an extended example of equilibrium in the overall game that fully addresses this complication. We then turn to analytic results in Section 4.2, where we consider the more restricted case in which service rates are sufficiently close to each other to leave the thresholds $n_h, n_\ell$ unaffected by the choices of each type of firm.

4.1 Example 2: Service Rate Equilibria

Let $v_h = 20, v_\ell = 1, c = 0.4005, \Lambda = 1, \mu = 1.15$ and $\bar{\mu} = 0.85$. We consider two cases for $q$, the proportion of informed consumers: $q = 0.05$ (Case I) and $q = 0.9$ (Case II). Observe that we have not set value of $p$ and $k$; rather, we will analyze how the equilibrium varies with $p$ and $k$.

We begin by analyzing the consumer game at stage 2. The equilibrium in the consumer game does not depend on $k$, the cost of speeding up.

4.1.1 Consumer Game at Stage 2

Table 2 indicates the possible equilibria in the consumer game, across all values of $p$. Each row in the table describes a different pair of service rate choices by the firms. In the table, we refer to the service rate $\bar{\mu} = 1.15$ as “Fast,” and the service rate $\mu = 0.85$ as “Slow.” The service rates in turn determine $n_h$ and $n_\ell$, the thresholds at which informed consumers balk. We also exhibit the queue lengths at which a hole in the uninformed consumer’s strategy can occur in equilibrium, with different queue lengths being supported as holes for different values of $p$.

Consider the condition (a) in part (i) of Proposition 3, $\frac{\mu_\ell}{(1-q)\mu_h} < \frac{v_h\mu_h + v_\ell\mu_\ell + 2c}{v_h\mu_h + v_\ell\mu_\ell - 2c}$. It is straightforward to check that for Case I ($q = 0.05$) this condition is satisfied given the service rate choices in the first three rows of the table, but not when firm $h$ is slow and firm $\ell$ is fast. Therefore, in Case I, whenever $\mu_h = 1.15$ or $\mu_h = \mu_\ell = 0.85$, for every $n \in \{n_\ell, n_\ell + 1, \ldots, n_h\}$ there exists a range of $p$ such that there is a pure strategy equilibrium in which the uninformed consumer’s strategy has a hole at $n$. However, if $\mu_h = 0.85$ and $\mu_\ell = 1.15$, the condition is violated. Here, in Case I, $n = 5$ and $\bar{n} = n_h = 42$. Therefore, a pure strategy equilibrium with a hole at $n$ can be supported only if $n \in \{2, 3, 4, 5, 42\}$. There cannot be an equilibrium at which the uninformed consumer’s strategy has a hole at any other value of $n$.

In Case II (with $q = 0.9$), the condition is violated for any pair of service rate choices.
As may be seen from the table, in this case (i.e., when the proportion of informed consumers is high), a hole in the uninformed consumer’s strategy can be supported in equilibrium only if the hole is either at queue length \( n_\ell \) or at \( n_h \), but no queue lengths in-between those two. That is, here \( n = n_\ell \) and \( \bar{n} = n_h \).

4.1.2 Equilibria in Service Rates

There are four possible pairs of service rate outcomes in the model, since each type of firm could choose either a fast or a slow service rate. Given the firms’ service rate choices \( \mu_h \) and \( \mu_\ell \), \( n_h \) and \( n_\ell \) are defined as in Table 2. We focus on values of \( p \) that support a pure strategy equilibrium in which the uninformed consumer’s strategy has a hole at \( n_\ell \); that is, we consider \( p \leq \bar{p}(n_\ell) \) in each case. We then set the uninformed consumer’s strategy to have a hole at \( n_\ell \).

Next, we then use the expressions in Lemma 3 part (i) to compute the revenue of each type of firm for each of the four pairs of service rate choices. The difference in revenues for firm \( \theta \) between the high and low service rates, \( R_\theta(\bar{\mu}, \sigma) - R_\theta(\mu, \sigma) \), then determines the maximum the firm is willing to pay to increase its service rate from \( \underline{\mu} \) to \( \bar{\mu} \). In any pure strategy equilibrium of the overall game, if firm \( \theta \) chooses the fast service rate \( \bar{\mu} \), it must be that \( R_\theta(\bar{\mu}, \sigma) - R_\theta(\mu, \sigma) \leq k \), with the inequality being reversed if it chooses the slow rate \( \underline{\mu} \). Thus, the revenue expressions allow us to construct the range of \( k \) which supports each of the four possible pairs of service rate choices as an equilibrium choice.

Table 3 provides the ranges of \( p \) (the prior probability the firm has high quality) and \( k \) (the cost of speeding up) for which each type of equilibrium is sustained. That is, in each
row, for all values of \( p \) and \( k \) in the specified range, it is an equilibrium for firms to choose the service rates mentioned in column 1 of the same row and for uninformed consumers to play a strategy that has a hole at \( n_\ell \). For each type of equilibrium, \( n_\ell = 2 \), \( \hat{n} = 2 \), and \( n_h = 57 \) when \( \mu_h = 1.15 \) and \( 42 \) when \( \mu_h = 0.85 \).

<table>
<thead>
<tr>
<th>Service Rate in Equilibrium Firm</th>
<th>Supporting Range of ( q = 0.05 )</th>
<th>Case I ( q = 0.05 )</th>
<th>Case II ( q = 0.9 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fast   Fast</td>
<td>( p )</td>
<td>( p \leq 0.006 )</td>
<td>( k \leq 0.101 )</td>
</tr>
<tr>
<td></td>
<td>( k )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slow  Slow</td>
<td>( p )</td>
<td>( p \leq 0.660 )</td>
<td>( p \leq 0.972 )</td>
</tr>
<tr>
<td></td>
<td>( k )</td>
<td>( k \geq 0.101 )</td>
<td>( k \geq 0.139 )</td>
</tr>
<tr>
<td>Fast  Slow</td>
<td>( p )</td>
<td>( p \leq 0.074 )</td>
<td>( k \in [0.101, 0.139] )</td>
</tr>
<tr>
<td></td>
<td>( k )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slow  Fast</td>
<td>( p )</td>
<td>( p \leq 0.135 )</td>
<td>( k \leq 0.101 )</td>
</tr>
<tr>
<td></td>
<td>( k )</td>
<td></td>
<td>( k \leq 0.101 )</td>
</tr>
</tbody>
</table>

This table shows the range of \( k \) and \( p \) that supports different kinds of service rate equilibria for two sets of parameters, when the equilibrium in the continuation game has \( \hat{n} = n_\ell = 2 \).

Table 3: Equilibrium service rate choices

Consider Case I first, in which the proportion of informed consumers is low, with \( q = 0.05 \). Here, we find that \( R_h(\bar{\mu}, \sigma) - R_\ell(\underline{\mu}, \sigma) < 0 \). That is, even if it is costless to improve its technology, firm \( h \) prefers the slow service rate. Therefore, there is no equilibrium in which firm \( h \) chooses the fast service rate \( \bar{\mu} \). The intuition follows from the hole in the uninformed consumer’s strategy being at a low queue length (\( \hat{n} = 2 \)). Since there are many uninformed consumers, it is valuable for firm \( h \) to signal its quality by ensuring that the queue generally remains greater than 2 in length.

For the low-quality firm in Case I, we find that \( R_\ell(\bar{\mu}, \sigma) - R_\ell(\underline{\mu}, \sigma) = 0.101 \) in all cases. Therefore, if \( k \leq 0.101 \), the firm chooses the fast service rate, and if \( k \geq 0.101 \), it chooses to slow down. Thus, in equilibrium firm \( h \) is slow and firm \( \ell \) is either fast or slow, depending
on the cost of speeding up. For each of these two sets of equilibria, the maximal value of \( p \) under which the equilibrium can be sustained is given by \( \bar{p}(n_\ell) \).

In Case II, with a high proportion of informed consumers, it is no longer important to the high-quality firm that uninformed consumers do not join at queue length 2. Regardless of service rate choices, the revenue gain from speeding up is greater for the high-quality firm than the low-quality firm. Therefore, it is not feasible to sustain an equilibrium in which the high-quality firm is slow and the low-quality firm is fast, although the converse case is now feasible. Now, when \( k \) is large both firms are slow, and when \( k \) is sufficiently small, over a narrow range of \( p \), both firms are fast. It may be noted that there are values of \( p \) and \( k \) at which multiple equilibria exist in the service rate game. For example, when \( p = 0.05 \) and \( k = 0.139 \), there is an equilibrium in which both firms are slow and an equilibrium in which the high-quality firm is fast and the low-quality firm is slow. Whenever there are multiple equilibria in service rates, one or both firms must be indifferent across the two service rates.

It is important to appreciate that the service rate equilibria in Table 3 depend on the equilibrium in the consumer game. As shown in Table 2, in Case I, the hole in the uninformed agent’s strategy may be at any value of \( n \). Therefore, there exist other ranges of \( p \) and \( k \) that support different kinds of equilibria. Suppose, for example, \( p \geq 0.907 \) and \( k \leq 0.101 \). Then, in both cases (i.e., whether \( q \) is high or low), there exists an equilibrium in which the both firms choose the fast service rate, and the hole in the uninformed consumer’s strategy is at \( n_h = 57 \).

4.2 Analytic Results on Service Rate Equilibria

The overall intuition from the previous example is that if \( q \), the proportion of informed consumers, is low and the hole in the uninformed agent’s strategy is at a low queue length (which relies in turn on a sufficiently low prior probability the firm has high quality, \( p \)), the high-quality firm has a greater incentive than the low-quality firm to choose the slow service rate. We now formalize this intuition. To obtain the same results analytically, we have to make some additional assumptions.

Recall that \( n_\theta = \lfloor v_\theta \frac{\mu}{c} \rfloor \) is the threshold at which informed consumers balk when the firm has type \( \theta \). We assume that \( n_h \) and \( n_\ell \) are invariant to the service rate chosen by each type of firm. This assumption takes away one incentive the high-quality firm may have to speed up, which is increasing the maximal length of the queue.

**Assumption 3** \( [v_\ell(\bar{\mu}/c)] = [v_\ell(\mu/c)] \) and \( [v_h(\bar{\mu}/c)] = [v_h(\mu/c)] \).
Assumption 3 effectively implies that the fast and slow service rates ($\bar{\mu}$ and $\mu$) are close to each other. In other words, if a firm speeds up the rate at which it services consumers, it does so through an incremental operational improvement. Let $\mu_0 > 0$ be the mean of the fast and slow service rates. Then, for any $\epsilon \in (0, \mu_0)$, define $\bar{\mu} = \mu_0 + \epsilon$ and $\underline{\mu} = \mu_0 - \epsilon$.

We focus on equilibria in the consumer game in which uninformed consumers play a pure strategy with a hole at exactly $n_\ell$. Let $\hat{\sigma}$ denote the consumer strategy in which informed consumers join the queue of a firm of type $\theta$ at all queue lengths up to and including $n_\theta - 1$, and uninformed consumers join the queue at all queue lengths up to $n_h - 1$ except at the length $n_\ell$. Observe that for a given service rate choice $\mu_h, \mu_\ell$, if $p \leq \bar{p}(n_\ell)$ (where $\bar{p}(n_\ell)$ is identified precisely in the proof of Proposition 3), such an equilibrium exists. Since $\bar{p}$ depends on $\mu_h$ and $\mu_\ell$, we can write it more generally as $\bar{p}(n_\ell | \mu_h, \mu_\ell)$. Then, define

$$p_0 = \min_{\mu_h, \mu_\ell \in \{\underline{\mu}, \bar{\mu}\}} \bar{p}(n_\ell | \mu_h, \mu_\ell).$$

If $p \leq p_0$, regardless of the service rate strategy chosen by each type of firm, there is an equilibrium of the consumer continuation game in which the hole in an uninformed consumer’s strategy is exactly at $n_\ell$.

Now, consider a firm of type $\theta$. Let $\Delta_\theta$ denote the marginal revenue the firm gains from choosing the fast service rate $\bar{\mu} = \mu_0 + \epsilon$ rather than the slow service rate $\underline{\mu} = \mu_0 - \epsilon$, given that consumers are playing the strategy $\hat{\sigma}$. Then,

$$\Delta_\theta(\mu_0, \epsilon) = R_\theta(\mu_0 + \epsilon, \hat{\sigma}) - R_\theta(\mu_0 - \epsilon, \hat{\sigma}).$$

Observe that, from Lemma 3, for each $\theta$ the revenue rate $R_\theta$ depends on $q$, the proportion of informed consumers. Hence, $\Delta_\theta$ also depends on $q$. For any $\mu_0$, the firm will strictly prefer the fast service rate $\mu_0 + \epsilon$ over the slow service rate $\mu_0 - \epsilon$ if and only if $\Delta_\theta(\mu_0, \epsilon) > k$, the marginal cost of installing added capacity (i.e., of choosing the fast service rate).

In the remainder of the paper, we consider two cases in which the high-quality firm prefers to slow down in equilibrium. We first consider service rates $\bar{\mu}$ and $\underline{\mu}$ that surround the arrival rate, $\Lambda$, and show that the incentives of the high-quality firm to speed up depend on the fraction of informed consumers, $q$. We then assume both service rates are strictly greater than, but close to, $\Lambda$, and show that the high-quality firm has a strict incentive to slow down, regardless of $q$.

### 4.3 Service Rates on Either Side of Arrival Rate

Suppose $\mu_0 = \Lambda$. Then, the slow service rate $\underline{\mu}$ is strictly less than $\Lambda$, and the fast service rate $\bar{\mu}$ is strictly greater than $\Lambda$. We consider the relative incentive each type of firm has...
to choose the fast service rate $\bar{\mu} = \Lambda + \epsilon$ over the slow service rate $\underline{\mu} = \Lambda - \epsilon$. We obtain clean analytic expressions for the derivative of revenue of type $\theta$ with respect to $\mu$, $\frac{\partial R_\theta}{\partial \mu}(\mu, \hat{\sigma})$, in the limit as $\epsilon$ goes to zero (i.e., $\underline{\mu}$ and $\bar{\mu}$ each approach $\Lambda$). Using these limiting analytic expressions as our starting point, we are then able to provide comparisons in a neighborhood of service rates around $\Lambda$.

We first provide the limiting expressions for the rate of change of revenue with respect to $\mu$, when the service rate exactly equals the arrival rate.

**Lemma 4** Suppose $\mu = \Lambda$ and $p \leq p_0$. Then, the rate of change of revenue with respect to $\mu$ is:

$$\frac{\partial R_h}{\partial \mu}(\Lambda, \hat{\sigma}) = \frac{r [2(n_h - n_\ell)^2 q^2 - (n_h - 3n_\ell - 1)(n_h - n_\ell)q + n_\ell(n_\ell + 1)]}{2[n_\ell + 1 + q(n_h - n_\ell)]^2}$$

$$\frac{\partial R_\ell}{\partial \mu}(\Lambda, \hat{\sigma}) = \frac{r n_\ell}{2(n_\ell + 1)}.$$

Consider the rate of change of the revenue of the high-quality firm in the limit, when $\mu = \Lambda$. First, suppose that $q = 1$; i.e., all consumers are informed. Then, $\frac{\partial R_h}{\partial \mu}(\Lambda, \hat{\sigma}) = \frac{r n_\ell}{2(n_\ell + 1)}$, which exceeds $\frac{\partial R_\ell}{\partial \mu}(\Lambda, \hat{\sigma}) = \frac{r n_\ell}{2(n_\ell + 1)}$ since $n_h > n_\ell$. That is, when all consumers are informed, the high-quality firm has a greater incentive to speed up than the low-quality firm. Next, suppose that $q = 0$, so that all consumers are uninformed. Then, the two derivatives in Lemma 4 are each exactly equal to $\frac{r n_\ell}{2(n_\ell + 1)}$. This is intuitive: if there are no informed consumers, the two types of firm have exactly the same stationary distribution over queue length. Hence, each type of firm has the same incentive to speed up.

In equilibrium, a firm of type $\theta$ will strictly prefer the fast service rate $\Lambda + \epsilon$ whenever $\Delta_\theta(\epsilon) > k$; that is, the increase in revenue from speeding up the service rate exceeds the cost of the technology. We have a closed-form expression for $R_\theta$ only in the limit, $\mu$ and $\bar{\mu}$ are each equal to $\Lambda$ (so that $\epsilon = 0$). To the extent that a firm has a strict incentive to speed up (i.e., choose $\bar{\mu} = \Lambda + \epsilon$) or slow down (i.e., choose $\underline{\mu} = \Lambda - \epsilon$), results obtained when $\mu = \Lambda$ will continue to hold for service rates in an $\epsilon$-neighborhood of $\Lambda$.

Define a threshold proportion of informed consumers, $\hat{q}$ as follows:

$$\hat{q} = \left(1 - \frac{1}{n_h - n_\ell}\right) \frac{n_\ell + 1}{n_\ell + 2}. \quad (11)$$

It is immediate that the threshold $\hat{q}$ is increasing in $n_h - n_\ell$.

Our main result is described as follows. Consider any $q > \hat{q}$. Then, if $\underline{\mu}$ and $\bar{\mu}$ are sufficiently close to $\Lambda$, the marginal revenue of the high-quality firm exceeds the marginal...
revenue of the low-quality firm; that is, \( \Delta_h(\Lambda, \epsilon) > \Delta_\ell(\Lambda, \epsilon) \). The high-quality firm therefore has a greater incentive to speed up than the low-quality firm. Thus, there is a range of costs at which the high-quality firm chooses the fast service rate \( \bar{\mu} \) and the low-quality firm chooses the slow service rate \( \mu \). Similarly, if \( q < \hat{q} \) and \( \mu \) and \( \bar{\mu} \) are sufficiently close to \( \Lambda \), the low-quality firm has a greater incentive to speed up than the high-quality firm; that is, \( \Delta_h(\Lambda, \epsilon) < \Delta_\ell(\Lambda, \epsilon) \). Thus, there is another range of costs at which the high-quality firm chooses the slow service rate \( \mu \) and the low-quality firm chooses the fast service rate \( \bar{\mu} \).

**Proposition 4** Suppose \( \bar{\mu} = \Lambda + \epsilon \) and \( \mu = \Lambda - \epsilon \). For every \( q \in (0, 1) \), there exists an \( \hat{\epsilon}(q) > 0 \) such that if \( \epsilon \in (0, \hat{\epsilon}(q)] \) and \( p \leq p_0 \):

(i) If \( q < \hat{q} \), then \( \Delta_h(\Lambda, \epsilon) < \Delta_\ell(\Lambda, \epsilon) \). Hence, there exists a range for \( k \) such that in equilibrium the high-quality firm chooses the slow service rate \( \Lambda - \epsilon \) and the low-quality firm chooses the fast service rate \( \Lambda + \epsilon \). However, the converse service rate strategies cannot be sustained in equilibrium.

(ii) If \( q > \hat{q} \), then \( \Delta_h(\Lambda, \epsilon) > \Delta_\ell(\Lambda, \epsilon) \). Hence, there exists a range for \( k \) such that in equilibrium the high-quality firm chooses the fast service rate \( \Lambda + \epsilon \) and the low-quality firm chooses the slow service rate \( \Lambda - \epsilon \). However, the converse service rate strategies cannot be sustained in equilibrium.

Consider the trade-off faced by the high-quality firm as it chooses its service rate. Assume that \( n_\ell < n_h \), i.e. informed consumers will balk at strictly higher queue length when the quality is high than when the quality is low. Slowing down (i.e., choosing the slow service rate) implies that the average waiting time for consumers is longer. From the firm’s viewpoint, slowing down implies that higher queue lengths occur more often. Uninformed consumers join the queue at all lengths except at the hole, i.e., at queue length \( n_\ell \). Suppose the queue length is exactly \( n_\ell \) when an informed consumer arrives. Recall that \( n_\ell < n_h \), hence, the informed consumer will join, and the queue length will become \( n_\ell + 1 \). Therefore, if the proportion of informed consumers is relatively high, the hole is irrelevant, since it will be crossed whenever an informed consumer arrives. However, if the proportion of informed consumers is low, the hole is unlikely to be crossed simply by an arriving consumer. In this case, choosing the slow service rate is potentially valuable, since a greater amount of time is then spent at queue lengths above the hole.

Intuitively, the benefit of slowing down for the high-quality firm will depend on the number of queue lengths above the hole; i.e., on \( n_h - n_\ell \). If \( n_h \) is substantially greater than
there is a greater incentive to slow down and reach the higher queue lengths. Conversely, if \( n_h \) is close to \( n_\ell \), the incentive to slow down is weaker. Indeed, it may be observed that the threshold proportion of informed consumers in equation (11), \( \hat{q} \), is strictly increasing in \((n_h - n_\ell)\). Thus, the range of \( q \) for which the high-quality firm has a stronger incentive to slow down, compared to the low-quality firm, increases in \((n_h - n_\ell)\).

Further, note that when \( n_h = n_\ell + 1 \), \( \hat{q} = 0 \). Thus, at the minimal distance between \( n_h \) and \( n_\ell \), the high-quality firm always has a stronger incentive to speed up, compared to the low-quality firm. It is only when the distance between \( n_h \) and \( n_\ell \) increases to two or more that the high-quality firm may have a weaker incentive to speed up.

Next, as a special case, consider \( v_h \) becoming large, keeping \( v_\ell \) fixed. Observe that, as \( v_h \) (and hence \( n_h \)) becomes large, the threshold \( \hat{q} \) (as defined in equation 11) goes to \( \frac{n_\ell + 1}{n_\ell + 2} \). Thus, it follows from Proposition 4 that for large values of \( v_h \), if \( q > \frac{n_\ell + 1}{n_\ell + 2} \), there is a range of \( k \) for which the high-quality firm speeds up and the low-quality firm slows down. Conversely, if \( q < \frac{n_\ell + 1}{n_\ell + 2} \), there is a range of \( k \) for which the high-quality firm slows down and the low-quality firm speeds up.

However, as \( v_h \) becomes large, we obtain even stronger results for the high-quality firm. Strikingly in this case, if the proportion of informed consumers is sufficiently low (in particular, \( q < \frac{1}{2} \)), the high-quality firm actually obtains a lower revenue when it speeds up to the fast service rate, compared to staying at the slow service rate. Hence, even if there is no cost to increasing the service rate, the high-quality firm will prefer to provide slow service.

Proposition 5 There exist an \( \bar{c}(q) \) and a threshold value for the high-quality good \( \bar{v}(q) \) with the following properties: Suppose \( v_h \geq \bar{v}(q) \), \( \epsilon \in (0, \bar{c}(q)] \), \( \bar{\mu} = \Lambda + \epsilon \), \( \mu = \Lambda - \epsilon \), and \( p \leq p_0 \). Then, \( \Delta_h(\Lambda, \epsilon) < 0 \) if \( q < \frac{1}{2} \) and \( \Delta_h(\Lambda, \epsilon) > 0 \) if \( q > \frac{1}{2} \). Hence, if \( q < \frac{1}{2} \), there is no value of \( k \) (including \( k = 0 \)) at which the high-quality firm chooses the fast service rate in equilibrium.

The intuition behind our results as follows. Speeding up service has two potential benefits for the high-quality firm. First, it may increase the threshold \( n_h \) at which informed consumers balk. Assumption 3 rules out this effect since \( n_h \) is fixed at the same level at both service rates \( \bar{\mu} \) and \( \mu \). Second, it shifts the stationary distribution of consumers toward the lower parts of the queue, in particular to queue lengths below the length at which the uninformed consumers have a hole in their strategy. When the proportion of informed consumers is low, speeding up service is less valuable, since the queue is less likely to cross
the length at which uninformed consumers have a hole.

When the proportion of informed consumers is relatively high, however, even at that queue length, the likelihood is that the next consumer is an informed consumer. Hence, queue lengths above the hole are more likely to be reached, making speeding up more valuable. This value is enhanced when \( v_h \) is large, so that there are many queue lengths above the length at which the uninformed consumers’ strategy has a hole.

4.4 Both Service Rates Greater than Arrival Rate

In the previous subsection, the slow service rate was below the consumer arrival rate \( \Lambda \) and the fast service rate was above the arrival rate. In this subsection, we consider the case in which both service rates, slow and fast, are strictly greater than the arrival rate \( \Lambda \). Suppose \( \mu_0 > \Lambda \), and as before let \( \mu = \mu_0 - \epsilon \) with \( \bar{\mu} = \mu_0 + \epsilon \). Here, we assume that \( \epsilon \in (0, \mu_0 - \Lambda] \), to ensure that \( \bar{\mu} \) is weakly greater than \( \Lambda \). We continue to assume that, for any service rate strategy of each type of firm, there is a pure strategy equilibrium \( \hat{\sigma} \) with a hole in the uninformed consumer’s strategy exactly at \( n_\ell \).

We show that, if the mean of feasible service rates, \( \mu_0 \), is sufficiently close to \( \Lambda \), and \( v_h \) is sufficiently high, the high quality firm will never choose the fast service rate, even if there is no cost to speeding up. This result mirrors the result in Proposition 5. Importantly, the slowing down by the high-quality firm in Proposition 5 occurs only when \( q < \frac{1}{2} \), whereas in Proposition 6 below, it occurs for any value of \( q \).

For any \( n \geq 1 \), define \( \rho(n) \) to be the unique solution to the equation

\[
\sum_{j=0}^{n} \rho^j + q \sum_{j=n+1}^{\infty} \rho^j = n + 1. \tag{12}
\]

Observe that the left-hand side equals 0 when \( \rho = 0 \), and is infinite when \( \rho = 1 \). Further, it is strictly increasing in \( \rho \). Hence, for any fixed value of \( n \) and any \( q \in (0, 1) \), \( \rho(n) \) exists, is unique, and is strictly less than one. Further, note that \( \lim_{n \to \infty} \rho(n) = 1 \).

**Proposition 6** Consider any \( n \in \{1, 2, \cdots \} \). Suppose \( \bar{\mu} = \mu_0 + \epsilon \) and \( \mu = \mu_0 - \epsilon \), where \( \mu_0 \in (\Lambda, \Lambda/\rho(n)) \) and \( cn/\mu_0 < v_\ell < c(n+1)/\mu_0 \). Then, there exists an \( \epsilon'(n) < \mu_0 - \Lambda \) such that, if \( \epsilon \in (0, \epsilon'(n)] \) and \( p \leq p_0 \), then for any choice of service rate by each type of firm, there is a pure strategy equilibrium in the consumer continuation game in which \( n_\ell = n \) and there is a hole in the uninformed consumer’s strategy at exactly \( n_\ell \). Further, there exists a \( v_h(n) \) such that if \( v_h \geq v_h(n) \), for any value of \( k \) (including \( k = 0 \)), the high-quality firm chooses the slow service rate \( \mu = \mu_0 - \epsilon \).
Proposition 6 shows that if $v_h$ is sufficiently high, the high-quality firm will never have an incentive to speed up, even if the superior service technology comes at no cost. The expression $\Lambda/\rho(n)$ tends toward $\Lambda$ as $n$ becomes large. Thus, the higher the position of the hole in the uninformed consumer’s strategy, the narrower the range of service rates for which the high-quality firm will not speed up.

The intuition underlying this proposition again relates to the number of feasible queue lengths beyond the length at which the uninformed consumer’s strategy has a hole. Recall that we fix the hole to be exactly at $n_\ell$. If $v_h$ is very high, there are a lot of queue lengths above the hole. Intuitively, speeding up increases the proportion of time the high-quality firm finds itself facing a queue below the length at which the hole exists. Thus, slowing down is valuable when there are many queue lengths above the hole at which uninformed consumers join. When $n_h$ is sufficiently high, the high-quality firm actually loses revenue by speeding up, and thus prefers to choose the slow service rate.

5 Conclusion and Discussion

Our characterization of the conditions under which firms of differing quality choose either fast or slow service rates translates quite naturally into predictions on how different types of firms strategically manipulate queues. In a pure strategy equilibrium, these conditions all depend on the queue length at which the equilibrium hole in an uninformed consumer’s strategy appears. For the uninformed, queues below the hole are short enough such that low waiting costs make joining rational. At the hole, they balk, and above the hole, uninformed consumers perfectly infer that the good has high quality.

Both firm types lose the uninformed consumers at the hole. The high-quality firm knows it will win back these customers at queue lengths above the hole. However the hole will only be filled (or crossed) if an informed consumer arrives at the market. Therefore, a high-quality firm is worst off facing a market with few informed consumers and in which the uninformed ascribe a low prior probability to it being of high quality. In this case, the hole is at low queue lengths, with little likelihood it will be filled by an informed customer. It is therefore in the high-quality firm’s best interest to keep the queue above the hole, leading it to choose a slow service rate. Quite naturally, the revenue-destroying effect of a fast service rate in these circumstances is at its highest when the slow service rate is slightly above the consumer arrival rate, and when there are few informed consumers that could fill the hole. Intuitively, the high quality firm can substitute for the effect of informed consumers by slowing down the service rate, ensuring a higher probability that the queue is above the
hole.

In our model, the queue serves to communicate information to uninformed consumers about the strength of demand, and hence about their own valuation for the product. More broadly, our (infinite horizon) queuing model can be considered as an approximation of the initial phase of a product life-cycle. The informational effect of excess demand will be strongest at this point as there are likely to be few informed consumers in the market. Later phases of the product life-cycle can also be approximated with our model, with a larger proportion of informed consumers. Our analysis suggests that it may be optimal for such firms to gradually increase the service rate over time as the fraction of informed consumers in the market increases. In case such firms cover the demand rate too early in the product life cycle, when the fraction of informed consumers is low, further service rate expansion would stall until the fraction of informed consumers is high enough. Indeed, if excess demand provides an informational externality that generates even more demand, it is no longer true that supply creates its own demand. Rather, the lack of supply may lead to an increase in demand.
Appendix

A Algorithms to Identify Equilibria in the Consumer Game

In this appendix, we develop algorithms to identify pure and mixed strategy equilibria in the consumer game. We assume throughout that the firms play pure strategies, so that $\beta_0 = \mu_0$ for each $\theta$.

Suppose a consumer arrives and sees queue length $n$. An informed consumer who knows the firm has type $\theta$ joins as long as $n < n_{\theta}$. That is, informed consumers play a pure strategy. Let $\sigma_i$ denote this strategy.

It is a best response for an uninformed consumer to join if $w_u(n, \gamma, \beta) \geq 0$, where $w_u(\cdot)$ is defined in equation (1). Let $\mu = (\mu_h, \mu_\ell)$. Since each type of firm plays a pure strategy, with a slight abuse of notation, we let $\pi_{\theta}(n, \mu, \sigma)$ be the stationary probability of queue length $n$ at a firm of type $\theta$, given that the consumer strategy is $\sigma$. Then, from equation (8),

$$\gamma(n) = \frac{p \pi_h(n, \mu, \sigma)}{p \pi_h(n, \mu, \sigma) + (1-p) \pi_\ell(n, \mu, \sigma)}$$

so that $1 - \gamma(n) = \frac{(1-p) \pi_\ell(n, \mu, \sigma)}{p \pi_h(n, \mu, \sigma) + (1-p) \pi_\ell(n, \mu, \sigma)}$. The condition $w_u(n, \beta, \sigma) \geq 0$ then reduces to

$$\left(1 - p\right) \pi_\ell(n, \mu, \sigma) \left[v_\ell - \frac{(n+1)c}{\mu_\ell}\right] + p \pi_h(n, \mu, \sigma) \left[v_h - \frac{(n+1)c}{\mu_h}\right] \geq 0 \quad (13)$$

Now, for $n < n_\ell$, we know the uninformed consumer joins. Consider $n \in \{n_\ell, \cdots, n_h\}$. Then, $v_\ell < \frac{(n+1)c}{\mu_\ell}$. Therefore, the inequality in (13) can be expressed as

$$\frac{\pi_\ell(n, \mu, \sigma)}{\pi_h(n, \mu, \sigma)} \leq \frac{p}{1-p} \frac{v_h - \frac{(n+1)c}{\mu_h}}{v_\ell - \frac{(n+1)c}{\mu_\ell}}. \quad (14)$$

Suppressing the dependence of the left-hand side on $\mu$ and $\sigma$ for convenience, let $\phi(n) = \frac{\pi_\ell(n, \mu, \sigma)}{\pi_h(n, \mu, \sigma)}$ be the likelihood the firm has low quality, given only that the queue length is $n$. Let

$$V(n) = \frac{v_h - \frac{(n+1)c}{\mu_h}}{\frac{(n+1)c}{\mu_\ell} - v_\ell}. \quad (15)$$

Then, $V(n_\ell) > 0 > V(n_h)$, and $V(n)$ clearly declines in $n$. It is a best response for the uninformed consumer to join if $\phi(n) \leq \frac{p}{1-p} V(n)$.

Observe further that, from equation (4), it follows that $\phi(n) = \phi(0) \left(\frac{\mu_h}{\mu_\ell}\right)^n \prod_{j=0}^{n-1} s_j(n, \sigma)$. Here, $\phi(0)$ depends on consumer actions at all queue lengths, whereas the third term depends only on consumer actions up to queue length $n-1$. We exploit this structure to transform the characterization of an equilibrium from an $|\bar{N}|$-dimensional problem to a single-dimensional fixed point problem.
A.1 Pure Strategy Equilibria

Suppose uninformed agents also play a pure strategy, and let $\hat{n}$ denote the queue length at which the strategy has a hole. Then, $s_\ell(j, \sigma) = 1$ for $n < n_\ell$ and $1 - q$ for $j \in \{n_\ell, \ldots, \hat{n}\}$, with $s_h(j, \sigma) = 1$ for all $j < \hat{n}$. We use these properties in our algorithm, which involves the following steps.

Step 1. Let $\varphi \in \mathbb{R}^+$ be a conjecture for $\phi(0)$, and define

$$
\Psi(n, \varphi) = \begin{cases} 
\varphi \left( \frac{\mu_h}{\mu_\ell} \right)^n, & 0 \leq n \leq n_\ell - 1 \\
\varphi \left( \frac{1 - q}{q} \right)^n \left( \frac{(1 - q)\mu_h}{\mu_\ell} \right)^n, & n_\ell \leq n \leq n_h
\end{cases}
$$

(16)

Step 2. Since $\Psi(n, \varphi) \geq 0$ and $V(n_h) < 0$, it must be that $\Psi(n_h, \varphi) > \frac{p}{1-p} V(n_h)$. Define

$$
\hat{n}(\varphi) = \min\{n \leq n_h | \Psi(n, \varphi) > \frac{p}{1-p} V(n)\}.
$$

(17)

The queue length $\hat{n}(\varphi)$ represents the hole in the uninformed consumer’s best response strategy. The consumer joins for all $n = 0, 1, \ldots, n_h - 1$, except at $\hat{n}$.

Step 3. For $n \in \{n_\ell, \ldots, n_h\}$, define

$$
\Phi(n) = \frac{1 + \sum_{m=1}^n \left( \frac{\Delta}{\mu_h} \right)^m + q \sum_{m=n+1}^{n_h} \left( \frac{\Delta}{\mu_h} \right)^m}{1 + \sum_{m=n_\ell}^n \left( \frac{\Delta}{\mu_\ell} \right)^m + \sum_{m=n_\ell+1}^{n_h} (1 - q)^m - n_\ell \left( \frac{\Delta}{\mu_\ell} \right)^m}.
$$

(18)

Then, $\Phi(n)$ is the likelihood ratio (specifically, the likelihood the firm has low quality) at the empty queue when the uninformed consumer’s strategy has a hole at $n$.

By definition, every equilibrium in the consumer game is characterized as a fixed point of the mapping $\varphi \mapsto \Phi(\hat{n}(\varphi))$. That is, for any pure strategy equilibrium $\sigma^*$, there exists a $\varphi^*$ that satisfies $\varphi^* = \Phi(\hat{n}(\varphi^*))$. Conversely, for any $\varphi^*$ such that $\varphi^* = \Phi(\hat{n}(\varphi^*))$, there is a pure strategy equilibrium in which the uninformed consumer’s strategy has a hole at $\hat{n}(\varphi^*)$.

Step 4. Define $\underline{\varphi} = \min_{n \in \{n_\ell, \ldots, n_h\}} \Phi(n)$ and $\bar{\varphi} = \max_{n \in \{n_\ell, \ldots, n_h\}} \Phi(n)$. Then, any pure strategy equilibrium $\sigma^*$ must satisfy $\varphi^* \in [\underline{\varphi}, \bar{\varphi}]$.

Step 5. Consider a fine grid for $\varphi \in [\underline{\varphi}, \bar{\varphi}]$, and compute $\Phi(\hat{n}(\varphi))$ over the grid. Any points at which $\varphi^* = \Phi(\hat{n}(\varphi^*))$ correspond to pure strategy equilibria. At all such points, an informed consumer’s equilibrium strategy is the same, and is straightforwardly determined as a threshold strategy in which the consumer balks at queue lengths $n_\theta$ and higher, for $\theta = h, \ell$. An uninformed consumer’s equilibrium strategy will vary across different values of $\varphi^*$, and is computed as in Step 3 above.
A.2 Mixed Strategy Equilibria

In this section, we extend the algorithm for finding the pure strategy equilibria in the consumer to find all mixed strategy equilibria. The algorithm exploits the property that, given the likelihood ratio at the zero queue, $\varphi$, both the expected benefit and expected waiting cost at any queue length $n$ depend only on the uninformed consumer’s actions at queue lengths 0 through $n - 1$.

Given our assumptions, an informed consumer continues to play a pure strategy in which she balks at $n_\theta$ and all higher queue lengths, for $\theta = h, \ell$. However, the uninformed consumer may randomize at some queue lengths. In any mixed strategy equilibrium as well, it must be that the likelihood ratio at the empty queue satisfies $\varphi^* \in [\underline{\varphi}, \overline{\varphi}]$.

Observe that $\hat{n}(\varphi)$ is, by construction, a weakly decreasing step function. Across all pure strategy equilibria, the hole in the uninformed consumer’s strategy can be only at a finite number of queue lengths between $n_\ell$ and $n_h$. That is, there are at most $n_h - n_\ell + 1$ pure strategy equilibria. Each such equilibrium is characterized by a downward discontinuity in the mapping $\varphi \mapsto \Phi(\hat{n}(\varphi))$. Let $\mathcal{F} = \{\varphi \mid \hat{n}(\varphi + \epsilon) > \hat{n}(\varphi - \epsilon) \text{ for every } \epsilon > 0\}$ be the set of such downward discontinuities. In any mixed strategy equilibrium, it must be that $\varphi^* \in \mathcal{F}$. Further, define $\underline{\Phi}(\varphi) = \min\{\lim_{\varepsilon \to 0^+} \Phi(\hat{n}(\varphi + \varepsilon)), \Phi(\hat{n}(\varphi - \varepsilon))\}$ and $\overline{\Phi}(\varphi) = \max\{\lim_{\varepsilon \to 0^+} \Phi(\hat{n}(\varphi + \varepsilon)), \Phi(\hat{n}(\varphi - \varepsilon))\}$.

For each $\varphi \in \mathcal{F}$ such that $\varphi \in [\underline{\Phi}(\varphi), \overline{\Phi}(\varphi)]$, we iterate through the following steps to determine whether it can support a mixed strategy equilibrium.

Step 1. Let $t = 1$. Set $\hat{n}_1 = \hat{n}(\varphi + \epsilon)$, and consider the following strategy for the uninformed consumer: $\sigma_1(n, a_1) = 1$ for $n \leq \hat{n}_1 - 1$, $\sigma_1(n, a_1) = a_1$ for $n = \hat{n}_1$, and $\sigma_1(n, a_1) = 1$ for $n \in \{\hat{n}_1 + 1, \ldots, n_h - 1\}$. For all $n \geq n_h$, $\sigma_1(n, a_1) = 0$.

Observe that when $a_1 = 0$, $\sigma_1(\cdot, a_1)$ is a pure strategy with a hole at $\hat{n}_1$. Define $\hat{\Psi}_0(n) = \Psi(n, \varphi)$ for each $n$. Note that, by definition of $\mathcal{F}$, $\hat{\Psi}_0(\hat{n}_1) = \frac{p}{1-p} V(\hat{n}_1)$.

Step 2. Slowly increase $a_t$ from 0. When $a_t$ is strictly positive, $s_\ell(\hat{n}_t, \sigma_1) = (1 - q)a_t$, and $s_h(\hat{n}_t, \sigma_1) = q + (1 - q)a_t$. For all $n \in \{\hat{n}_t + 1, \ldots, n_h\}$, $s_\ell(n, \sigma_1) = 1 - q$ and $s_h(n, \sigma_1) = 1$.

For each value of $a_t$,

(i) Define

$$
\hat{\Psi}_t(n \mid a_t) = \begin{cases} 
\hat{\Psi}_{t-1}(n) & \text{if } n \leq \hat{n}_t \\
\hat{\Psi}_{t-1}(n) \frac{(1-q)a_t}{q+(1-q)a_t} (1-q)^{n-\hat{n}_t-1} & \text{if } \hat{n}_t + 1 \leq n \leq n_h.
\end{cases}
$$

(19)
For values of \(a_t\) sufficiently close to zero, we have \(\hat{\Psi}_t(n \mid a_t) < V_n\) for all \(n \geq \tilde{n}_t + 1\). Observe also that since \(\phi(n)\) depends only on consumer behavior up to queue length \(n - 1\), it remains the case that \(\hat{\Psi}_t(\tilde{n} \mid a_t) = \frac{p}{1-p} V(\tilde{n}_t)\) for all values of \(a_t\).

(ii) Define a best response strategy for the uninformed consumer at \(n > \tilde{n}_t\) as follows. Let \(\hat{n}_{t+1}(a_t) = \min\{n \geq \tilde{n}_t + 1 \mid \hat{\Psi}_t(n \mid a_t) > \frac{p}{1-p} V(n)\}\). Then, \(\sigma_t(n, a_t) = 1\) for all \(n \in \{\tilde{n}_t + 1, \ldots, \hat{n}_{t+1}(a_t) - 1\}\), \(\sigma_t(\hat{n}_{t+1}(a_t), a_t) = 0\) and \(\sigma_t(n, a_t) = 1\) for all \(n \in \{\hat{n}_{t+1}(a_t) + 1, \ldots, n_h - 1\}\). Of course, \(\sigma_t(n_h, a_t) = 0\).

(iii) Next, we compute the likelihood ratio at the empty queue as:

\[
\hat{\Phi}(\sigma_t) = \frac{1 + \sum_{n=1}^{n_h} \left(\frac{1}{\mu_h}\right)^n \prod_{m=0}^{n-1} (q\sigma_t(h, m) + (1 - q)\sigma_t(m, a_t)(m))}{1 + \sum_{n=1}^{n_h} \left(\frac{1}{\mu_t}\right)^n \prod_{m=0}^{n-1} (q\sigma_t(h, m) + (1 - q)\sigma_t(m, a_t))}.
\]

We write \(\hat{\sigma}_t(\varphi, a_t) = \sigma_t(\cdot, a_t)\).

**Step 3.** As \(a_t\) is increased, one of the following two events must happen.

(i) For some value(s) of \(a_t\), \(\hat{\Phi}(\hat{\sigma}_t(\varphi, a_t)) = \varphi\). Then \(\varphi\) characterizes a mixed strategy equilibrium in which the uninformed consumer plays the strategy \(\sigma_t(\cdot, a_t)\).

(ii) For some value(s) of \(a_t\), there is an \(n' > \tilde{n}_t\) such that \(\hat{\Psi}(n', a_t) = \frac{p}{1-p} V(n')\). Let \(A_t = \{a_t \mid \text{ there exists } n' > \tilde{n}_t \text{ such that } \hat{\Psi}(n', a_t) = \frac{p}{1-p} V(n)\}\).

For each \(a_t \in A_t\), define \(\bar{\Phi}_t = \min[\lim_{\epsilon \to 0^+} \{\hat{\Phi}(\hat{\sigma}_t(\varphi, a_t + \epsilon)), \hat{\Phi}(\hat{\sigma}_t(\varphi, a_t - \epsilon))\}]\) and \(\bar{\Phi}_t = \max[\lim_{\epsilon \to 0^+} \{\hat{\Phi}(\hat{\sigma}_t(\varphi, a_t + \epsilon)), \hat{\Phi}(\hat{\sigma}_t(\varphi, a_t - \epsilon))\}]\). By construction, we have \(\varphi \in [\bar{\Phi}_t, \bar{\Phi}_t]\).

Increase \(t\) to \(t + 1\), and set \(\hat{n}_{t+1} = n'\). Define \(\sigma_{t+1}(\cdot, a_{t+1})\) as follows:

For \(n \leq \tilde{n}_t\), set \(\sigma_{t+1}(n, a_{t+1}) = \sigma_t(n, a_t)\). In particular, at queue length \(\tilde{n}_t\), the uninformed consumer joins with probability \(a_t\) (and balks with probability \(1 - a_t\)). Set \(\sigma_{t+1}(n, a_{t+1}) = 1\) for all \(n \in \{\tilde{n}_{t+1} + 1, \ldots, n_h - 1\}\), and \(\sigma_{t+1}(n_h, a_{t+1}) = 0\).

Now, repeat from Step 2 onward. The queue length \(\tilde{n}_{t+1}\) thus represents a second queue length at which the uninformed consumer may randomize.

Observe that \(\tilde{n}_t\) is strictly increasing in \(t\). Therefore, the algorithm will either find a mixed strategy equilibrium at a lower \(t\), or for some finite value of \(t\) we will find that \(\tilde{n}_t = n_h - 1\). In the latter case, the best response strategy in Step 2 (ii) for \(n > \tilde{n}_t\) must have the consumer not joining at \(n_h\). Therefore, \(\hat{\Phi}(\sigma_t)\) is continuous in \(a_t\). Further, since \(\phi \in [\Phi_t, \hat{\Phi}_t]\), there must in this case be an equilibrium with mixing at \(\tilde{n}_t\). Therefore, the algorithm always finds a mixed strategy equilibrium at the relevant value of \(\varphi\).
A.3 Illustration of Mixed Strategy Algorithm in Example 1

Recall that the parameters for example 2 are \( p = 0.5, q = 0.45, v_h = 1.70, v_\ell = 0.75, \)
\( c = 0.35, \Lambda = 1, \mu_h = 1.40 \) and \( \mu_\ell = 0.75. \) We go through the following steps to identify
the mixed strategy equilibria in this example.

Step 1. Construct the mapping \( \varphi \mapsto \Phi(\hat{n}(\varphi)) \) mentioned in Step 3 of the pure strategy
algorithm. The set of downward discontinuities in this mapping, \( \mathcal{F} \) is as defined in
the mixed strategy algorithm.

We wish to identify a \( \varphi \in \mathcal{F} \) such that \( \varphi \in [\Phi(\varphi), \bar{\Phi}(\varphi)]. \) We find that this
property holds only at \( \varphi = 0.7626, \) with \( \Phi(\varphi) = 0.7074 \) and \( \bar{\Phi}(\varphi) = 0.8001. \)

For \( \varphi = 0.7626, \) the uninformed consumer strictly wishes to join at queue
lengths 0 and 1, and is indifferent between join and balk at \( n = 2. \)

Step 2. Set \( t = 1. \) Let \( a_1 \) denote the probability the uninformed consumer joins at \( n = 2. \)
We vary \( a_1 \) between 0 and 1, and compute the mapping \( \hat{\Phi}(\sigma_1) \) defined in equation
(20). This mapping is depicted in the sub-figure on the left in Figure 1.

\[ 0 \quad 0.1 \quad 0.2 \quad 0.3 \quad 0.4 \quad 0.5 \quad 0.6 \quad 0.7 \quad 0.8 \quad 0.9 \quad 1 \]
\[ 0.7 \quad 0.71 \quad 0.72 \quad 0.73 \quad 0.74 \quad 0.75 \quad 0.76 \quad 0.77 \quad 0.78 \quad 0.79 \quad 0.8 \]

\( \Phi(\sigma_1) \)

\[ \text{Probability uninformed consumer joins at } n=2 \]

Randomization at \( n = 2 \)

\[ 0 \quad 0.1 \quad 0.2 \quad 0.3 \quad 0.4 \quad 0.5 \quad 0.6 \quad 0.7 \quad 0.8 \quad 0.9 \quad 1 \]
\[ 0.75 \quad 0.752 \quad 0.754 \quad 0.756 \quad 0.758 \quad 0.76 \quad 0.762 \quad 0.764 \quad 0.766 \quad 0.768 \quad 0.77 \]

\( \Phi(\sigma_2) \)

\[ \text{Probability uninformed consumer joins at } n=4 \]

Randomization at \( n = 4 \)

Figure 1: \( \hat{\Phi} \) mappings in Example 2
We find that at $a_1 = 0.0924$, $\Phi_1 = 0.7616$ and $\bar{\Phi}_1 = 0.7686$, so that $\varphi = 0.7626 \in [\Phi_1, \bar{\Phi}_1]$, as in Step 3 (ii) of the mixed strategy algorithm. We therefore set $\sigma_u(2) = 0.0924$; that is, the uninformed consumer joins with probability 0.0924 at $n = 2$.

Further, at this value of $\sigma_u(2)$, the uninformed consumer is indifferent between join and balk at $n = 4$. Thus, $n' = 4$, and we move to $t = 2$. So far we have the following strategy for the uninformed consumer: join at $n = 0$ and 1, join with probability 0.0924 at $n = 2$, and join at $n = 3$.

**Step 3.** Let $a_2$ denote the probability that the uninformed consumer joins at $n = 4$. We vary $a_2$ between 0 and 1 and construct the mapping $\hat{\Phi}(\hat{\sigma}_2)$ defined in equation (20). This mapping is depicted in the right sub-figure of Figure 1.

We find there are two values of $a_2$ at which $\hat{\Phi}(\hat{\sigma}_2) = \varphi = 0.7626$. These are $a_2 = 0.0866$ and $a_2 = 0.4934$.

**Step 4.** When $a_2 = 0.0866$, the uninformed consumer joins with probability 1 at $n = 5$ and balks at queue lengths 6 or higher. This corresponds to mixed strategy equilibrium (M1).

When $a_2 = 0.4934$, the uninformed consumer balks at queue lengths 5 or higher. This corresponds to mixed strategy equilibrium (M2).

Finally, at $a_2 = 0.1843$, we find that $\Phi_2 = 0.7590$ and $\bar{\Phi} = 0.7638$, so that $\varphi = 0.7626 \in [\Phi_2, \bar{\Phi}_2]$. At this value of $a_2$, the uninformed consumer is indifferent between join and balk at $n = 5$; that is, $n' = 5$. To analyze this last case, set $\sigma_u(4) = 0.1843$, set $t = 3$, and continue with the algorithm.

**Step 5.** Let $a_3$ be the probability the uninformed consumer joins at $n = 5$. Vary $a_3$ between 0 and 1 and construct the mapping $\hat{\Phi}(\hat{\sigma}_3)$ defined in equation (20).

When $a_3 = 0.7588$, we find that $\hat{\Phi}(\hat{\sigma}_3) = \varphi = 0.7626$. Thus, we have mixed strategy equilibrium (M3), in which the uninformed consumer joins with probability 1 at $n = 0, 1, 3$, and with probability 0.0924 at $n = 2$, probability 0.1843 at $n = 4$, and probability 0.7588 at $n = 5$. She balks at all queue lengths 6 and higher.
B Proofs

Proof of Lemma 1

\( \pi(\theta, n, \mu, \sigma) \) is the stationary probability of observing a queue of length \( n \) when the firm has quality \( \theta \) and chooses service rate \( \mu \), and agents play the strategy profile \( \sigma \). This is the long run probability of a birth-death process. The birth rate at which agents join the queue when the queue length is \( n \) is \( \Lambda_s(\theta, n, \mu, \sigma) \). Once in the queue, consumers leave at the rate \( \mu \) (the death rate). Thus, given firm quality \( \theta \) and agents’ strategy \( \sigma \), the flow balance equations are

\[
\pi(n, \mu, \sigma) = \pi(n - 1, \mu, \sigma)\Lambda_s(n - 1, \mu, \sigma) + \pi(n + 1, \mu, \sigma)\mu = \pi(n, \mu, \sigma)[\Lambda_s(n, \mu, \sigma) + \mu] \quad \text{for } n \geq 1.
\]

Further, it must be that \( \sum_{n=0}^{\infty} \pi(n, \mu, \sigma) = 1 \).

Recursively solving this system of equations yields the expressions in the statement of the Lemma.

Proof of Proposition 1

Suppose there is a pure strategy equilibrium \( \sigma \) in the consumer game. Consider a newly-arrived uninformed consumer. Suppose first that \( n < n_h \). Since an informed consumer joins the queue of the low-quality firm for all \( n < n_h \), it must be that \( v_h(n) > (n + 1)\frac{c}{\mu_h} \). Now, by Assumption 2, it follows that \( v_h > (n + 1)\frac{c}{\mu_h} \). Hence, for any value of \( \gamma(n) \in [0, 1] \), \( w_u(n, \gamma, \beta) > 0 \). Thus, the uninformed consumer should join the queue whenever \( n < n_h \), so that in equilibrium \( \sigma_u(n) = 1 \) for \( n \) in this range.

Now, suppose \( n \geq n_h \). In a pure strategy equilibrium, \( \sigma_u(n) = 1 \) or \( \sigma_u(n) = 0 \) for all \( n \). Suppose first that \( \sigma_u(n) = 1 \) for all \( n < n_h \). By definition of \( n_h \), \( v_h(n) > (n_h + 1)\frac{c}{\mu_h} \). Hence, regardless of her belief \( \gamma(n_h) \), the uninformed consumer does not join at \( n_h \), so that \( \sigma_u(n_h) = 0 \).

Next, suppose there exists an \( n \in \{n_E, n_h + 1, \ldots, n_h - 2\} \) at which \( \sigma_u(n) = 0 \). Then, \( s_e(n, \sigma) = 0 \), since an informed agent does not join the queue of a low-quality firm when \( n > n_h \). However \( s_h(n, \sigma) = q > 0 \). Now, consider queue length \( n + 1 \). Since \( n + 1 < n_h \), it follows that \( \sigma_i(h, n + 1) = 1 \). Then, from Lemma 1, it follows that \( \pi_i(n + 1, \mu, \sigma) = 0 \) and \( \pi_h(n + 1, \mu, \sigma) > 0 \). Therefore, from Bayes’ rule, it must be that \( \gamma(n + 1) = 1 \); that is, the uninformed consumer believes the firm has high quality with probability 1. Since the best response of an informed consumer is strict for every \( n \), it must be that \( \sigma_u(n + 1) = \sigma_i(n + 1) \).

The same reasoning applies to any queue length \( \tilde{n} > n + 1 \); at any such queue length,
\[ \gamma(\hat{n}) = 1, \text{ so that } \sigma_u(\hat{n}) = \sigma_i(h, \hat{n}). \]

Finally, suppose that \( \sigma_u(n_h - 1) = 0 \). The queue length \( n_h \) is never observed in equilibrium, so an uninformed consumer’s beliefs are arbitrary at that queue length. However, since \( \sigma_i(h, n_h) = 0 \), regardless of beliefs, \( w_u(n_h, \gamma, \beta) < 0 \), so that in equilibrium it must be that \( \sigma_u(n_h) = 0 = \sigma_i(h, n_h) \). \hfill \blacksquare

**Proof of Lemma 2**

Suppose a firm of type \( \theta \) chooses a pure strategy \( \mu_\theta \), and there is a pure strategy \( \sigma \) in the consumer game. Then, the overall probability a consumer joins the queue of the low-quality firm when there are already \( j \) consumers in the queue is given by

\[
\begin{align*}
s_\ell(\theta, \sigma) &= \begin{cases} 
1 & \text{if } j \in \{0, \cdots, n_\ell - 1\} \\
(1 - q) & \text{if } j \in \{n_\ell, \cdots, \hat{n} - 1\} \\
0 & \text{otherwise}
\end{cases}
\end{align*}
\]

Similarly, for the high-quality firm,

\[
\begin{align*}
s_h(\theta, \sigma) &= \begin{cases} 
1 & \text{if } j \in \{0, \cdots, \hat{n} - 1\} \text{ or } j \in \{n_\ell + 1, \cdots, n_h - 1\} \\
q & \text{if } j = \hat{n}
\end{cases}
\end{align*}
\]

Now, for an uninformed consumer, the posterior probability the firm has high quality may be written as

\[
\gamma(n, \mu_h, \mu_\ell) = \frac{p \pi_h(n, \mu_h, \sigma)}{p \pi_h(n, \mu_h, \sigma) + (1 - p) \pi_\ell(n, \mu_\ell, \sigma)} = \frac{p}{p + (1 - p) \frac{\pi_\ell(n, \mu_\ell, \sigma)}{\pi_h(n, \mu_h, \sigma)}}.
\]

Consider the expression for \( \pi_\ell(n, \mu, \sigma) \) in equation (4) in the statement of Lemma 1. Let \( \phi_0 = \frac{\pi_\ell(0, \mu_\ell, \sigma)}{\pi_h(0, \mu_h, \sigma)} \). Then, it follows that

\[
\frac{\pi_\ell(n, \mu, \sigma)}{\pi_h(n, \mu, \sigma)} = \begin{cases} 
\phi_0(\mu_h, \mu_\ell, \sigma)(\mu_h/\mu_\ell)^n & \text{if } n \in \{0, \cdots, n_\ell\} \\
\phi_0(\mu_h, \mu_\ell, \sigma)(\mu_h/\mu_\ell)^n(1 - q)^{n-n_\ell} & \text{if } n \in \{n_\ell + 1, \cdots, \hat{n}\} \\
0 & \text{otherwise}
\end{cases}
\]

The statement of the lemma now follows. \hfill \blacksquare

**Proof of Proposition 2**

Consider the algorithm for determining pure strategy equilibria in Section A.1. From the definition of \( \hat{n} \) in equation (17), it is easy to see that \( \hat{n}(\phi) \) is decreasing in \( \phi \).

Now, consider the mapping \( \Phi(n) \) introduced in equation (17). We proceed to find conditions under which \( \Phi(n) \) is increasing in \( n \) and decreasing in \( \ell \).
Let \( a_n = 1 + \sum_{m=1}^{n} \left( \frac{A}{\mu_h} \right)^m + q \sum_{m=n+1}^{n_h} \left( \frac{A}{\mu_h} \right)^m \) and \( b_n = 1 + \sum_{m=1}^{n} \left( \frac{A}{\mu_e} \right)^m + \sum_{m=n_{\ell}+1}^{n} (1-q)^{m-n_{\ell}} \left( \frac{A}{\mu_e} \right)^m \). Then, \( \Phi(n) = \frac{a_n}{b_n} \), and we can write \( \Phi(n+1) = \frac{a_n+c_{n+1}}{b_n+d_{n+1}} \), where \( c_{n+1} = (1-q) \left( \frac{A}{\mu_e} \right)^{n+1} \) and \( d_{n+1} = (1-q)^{n_{\ell}+1-n_{\ell}} \left( \frac{A}{\mu_e} \right)^{n+1} \).

It is immediate that if \( c_{n+1} > d_{n+1} \) above, then \( \frac{a_n+c_{n+1}}{b_n+d_{n+1}} > \frac{a_n}{b_n} \). Evaluating these terms at \( n = n_{\ell} \), suppose

\[
\left( \frac{\mu_{\ell}}{\mu_h} \right)^{n_{\ell}+1} > \frac{1 + \sum_{m=1}^{n_{\ell}} \left( \frac{A}{\mu_h} \right)^m + q \sum_{m=n_{\ell}+1}^{n_h} \left( \frac{A}{\mu_h} \right)^m}{1 + \sum_{m=1}^{n_{\ell}} \left( \frac{A}{\mu_e} \right)^m}.
\]

Then, \( \Phi(n_{\ell}+1) > \Phi(n_{\ell}) \).

Now, \( \Phi(n+2) \) can be written generically as \( \frac{a_n+c_{n+1}+c_{n+2}}{b_n+d_{n+1}+d_{n+2}} \), where \( a_n, b_n, c_{n+1}, d_{n+1} \) are as defined above, \( c_{n+2} = c_{n+1} \frac{A}{\mu_h} \) and \( d_{n+2} = d_{n+1}(1-q) \frac{A}{\mu_e} \). In particular, \( \frac{c_{n+2}}{d_{n+2}} > \frac{c_{n+1}}{d_{n+1}} \) if and only if \( \mu_{\ell} > (1-q)\mu_h \), or \( q > 1 - \frac{\mu_{\ell}}{\mu_h} \). Whenever \( \Phi(n+1) > \Phi(n) \), it follows that \( \frac{c_{n+1}}{d_{n+1}} > \frac{a_n}{b_n} \) and hence \( \frac{c_{n+2}}{d_{n+2}} > \frac{c_{n+1}}{d_{n+1}} \) implies that \( \frac{c_{n+2}}{d_{n+2}} > \frac{a_n}{b_n} \). That is, under these conditions, it follows that \( \Phi(n+2) > \Phi(n+1) \). Further, if \( \mu_{\ell} > (1-q)\mu_h \), it must be that \( \frac{c_{n+2}}{d_{n+2}} > \frac{c_{n+1}}{d_{n+1}} \) for any \( n \), so that \( \Phi(n) \) is monotonically increasing in \( n \).

Finally, observe that if inequality \( (21) \) is reversed and \( \mu_{\ell} < (1-q)\mu_h \), the argument above is entirely reversed, and \( \Phi(n) \) is monotonically decreasing in \( n \).

(i) Now, consider part (i) of the proposition. Observe that under the assumptions in the statement, \( \Phi(n) \) is monotonically decreasing in \( n \). Further, \( \hat{n}(\varphi) \) is weakly decreasing in \( \varphi \). Therefore, the mapping \( \varphi \mapsto \Phi(\hat{n}(\varphi)) \) is weakly increasing in \( \varphi \). We know that at least one equilibrium exists in the game. Further, if \( \varphi^* \) supports a mixed strategy equilibrium, then it must be that \( \lim_{\epsilon \to 0} \Phi(\hat{n}(\varphi^* - \epsilon)) \leq \varphi^* \) and \( \lim_{\epsilon \to 0} \Phi(\hat{n}(\varphi^* + \epsilon)) \geq \varphi^* \). Therefore, there must exist at least one point \( \varphi' \) at which \( \Phi(\hat{n}(\varphi')) = \varphi' \). That is, there is at least one pure strategy equilibrium in the consumer game.

(ii) Under the assumptions in the statement of part (ii) of the proposition, \( \Phi(n) \) is monotonically increasing in \( n \). Since \( \hat{n}(\varphi) \) is weakly decreasing in \( \varphi \), it follows that, the mapping \( \varphi \mapsto \Phi(\hat{n}(\varphi)) \) is weakly decreasing in \( \varphi \). As a result, there can be at most one point \( \varphi^* \) at which \( \Phi(\hat{n}(\varphi^*)) = \varphi^* \). That is, there is at most one pure strategy equilibrium in the consumer game.

\[ \hat{n}(\varphi) = \min \left\{ n_{\ell} \leq n \leq n_h \mid \left( \frac{\mu_{\ell}}{1 - q \mu_h} \right)^n V(n) < \frac{1 - p}{p} (1-q)^{n_{\ell}} \varphi \right\}. \]
Fix all other parameters except \( p \), and consider the inequality \( \left( \frac{\mu_\ell}{(1 - q)\mu_h} \right)^n V(n) < \frac{p}{1 - p} (1 - q)^n \). Suppose \( \left( \frac{\mu_\ell}{(1 - q)\mu_h} \right)^n V(n) \) is strictly decreasing in \( n \). Then, for each \( \varphi \) and each \( n \in \{ n_\ell, \ldots, n_h \} \), we can find a range of \( p \) such that \( \hat{n}(\varphi) = n \). We then set \( \varphi = \Phi(n) \) (where \( \Phi(n) \) is defined in equation (18)) to identify a range of \( p \) which supports a pure strategy equilibrium in which the uninformed consumer’s strategy has a hole at \( n \).

We begin by exploring sufficient conditions to ensure that \( \left( \frac{\mu_\ell}{(1 - q)\mu_h} \right)^n V(n) \) is strictly decreasing in \( n \). Observe that

\[
\left( \frac{\mu_\ell}{(1 - q)\mu_h} \right)^n V(n) > \left( \frac{\mu_\ell}{(1 - q)\mu_h} \right)^{n+1} V(n + 1)
\]

\[
\iff \frac{\mu_\ell}{(1 - q)\mu_h} < \frac{V(n)}{V(n + 1)}.
\]

(23)

Consider the right-hand side of inequality (23), \( \frac{V(n)}{V(n+1)} \). We establish the following claim:

Claim: Suppose \( n \leq n_h - 3 \), so that \( V(n + 2) > 0 \). Then, \( \frac{V(n)}{V(n+1)} \leq \frac{V(n+1)}{V(n+2)} \iff n \leq \frac{v_h \mu_h + v_\ell \mu_\ell}{2c} - 2 \).

Proof of Claim: Observe that \( \frac{V(n)}{V(n+1)} > \frac{V(n+1)}{V(n+2)} \iff V(n)V(n+2) > V(n+1)^2 \). The latter inequality is true if and only if

\[
\frac{\{v_h \mu_h - (n + 1)c\} \{v_h \mu_h - (n + 3)c\}}{(n + 1)c - v_\ell \mu_\ell} \frac{\{v_h \mu_h - (n + 2)c\}^2}{\{(n + 2)c - v_\ell \mu_\ell\}^2} > \frac{\{v_h \mu_h - (n + 2)c\}^2}{\{(n + 3)c - v_\ell \mu_\ell\}^2}.
\]

(24)

Write \( z = v_h \mu_h, y = v_\ell \mu_\ell, f = (n + 1)c, g = (n + 2)c \) and \( h = (n + 3)c \). Then, we can write (24) as

\[
\frac{(z - f)(z - h)}{(f - y)(h - y)} > \frac{(z - g)^2}{(g - y)^2},
\]

or \( (z - f)(z - h)(g - y)^2 > (f - h)(h - y)(z - g)^2 \). Multiply out fully, and eliminate common terms on both sides (using the relationship \( 2g = f + h \)). Rearranging the remaining terms, the inequality reduces to

\[
(z^2 - y^2)(g^2 - fh) > (z - y)g([f + h]g - 2fh).
\]

(25)

Now, \( z - y = v_h \mu_h - v_\ell \mu_\ell > 0 \) by assumption (2). Further, since \( f + h = 2g \), we can write the right-hand side as \( 2(z - y)g(g^2 - fh) \). Here, \( g^2 - fh = (n + 2)^2 - (n + 1)(n + 3) = n^2 + 4n + 4 - n^2 - 4n - 3 = 1 > 0 \). Divide both sides of inequality (25) by \( (z - y)(g^2 - fh) \), leaving \( z + y > 2g \), or

\[
v_h \mu_h + v_\ell \mu_\ell > 2(n + 2)c,
\]

36
which reduces to \( n < \frac{v_h\mu_h + v_i \mu_i}{2c} - 2 \).

Therefore, \( \frac{V(n)}{V(n+1)} > \frac{V(n+1)}{V(n+2)} \) if and only if \( n < \frac{v_h\mu_h + v_i \mu_i}{2c} - 2 \). Following through the same steps starting with \( \frac{V(n)}{V(n+1)} < \frac{V(n+1)}{V(n+2)} \), it is easy to see that the latter inequality holds if and only if \( n > \frac{v_h\mu_h + v_i \mu_i}{2c} - 2 \). Finally, it also follows that \( \frac{V(n)}{V(n+1)} = \frac{V(n+1)}{V(n+2)} \) if and only if \( n = \frac{v_h\mu_h + v_i \mu_i}{2c} - 2 \). \( \blacksquare \)

Now, we return to the proof of part (i) of the proposition. It follows from the claim that if the quantity \( \frac{v_h\mu_h + v_i \mu_i}{2c} - 2 \) is an integer, the expression \( \frac{V(n)}{V(n+1)} \) attains a global minimum at this queue length. A lower bound for the value of \( \frac{V(n)}{V(n+1)} \) is thus \( \frac{v_h\mu_h - v_i \mu_i}{v_h\mu_h - v_i \mu_i - 2c} \). Therefore, if \( \frac{\mu_i}{(1-q)\mu_h} < \frac{v_h\mu_h - v_i \mu_i}{v_h\mu_h - v_i \mu_i - 2c} \), the expression \( \left( \frac{\mu_i}{(1-q)\mu_h} \right)^n V(n) \) is strictly decreasing in \( n \).

Now, let \( \delta(n) = (1 - q)^{n_\ell} \left( \frac{\mu_i}{(1-q)\mu_h} \right)^n V(n) \), and for \( n \in \{n_\ell, \ldots, n_h\} \), define \( \bar{p}(n) = \min \left\{ \frac{\Phi(n)}{\Phi(n)+\delta(n-1)}, 1 \right\} \), with \( \underline{p}(n) = \max \left\{ \frac{\Phi(n)}{\Phi(n)+\delta(n-1)}, 0 \right\} \). Since \( \delta(n) < \delta(n-1) \), it follows that \( \bar{p}(n) > \underline{p}(n) \).

Now, for a given \( n \in \{n_\ell+1, \ldots, n_h\} \), suppose \( p \in [\underline{p}(n), \bar{p}(n)] \). Then, \( \frac{\bar{p}(n)}{1-p} \delta(n) \leq \Phi(n) \leq \frac{\underline{p}(n)}{1-p} \delta(n) - 1 \). It follows that there exists a pure strategy equilibrium in the consumer game in which the uninformed consumer’s strategy has a hole at exactly \( n \). Recognizing that \( \underline{p}(n) = 0 \), the same argument extends to an equilibrium in which the uninformed consumer’s strategy has a hole at \( n_\ell \).

(ii) Next, suppose \( \frac{\mu_i}{(1-q)\mu_h} > \frac{v_h\mu_h - v_i \mu_i + 2c}{v_h\mu_h - v_i \mu_i - 2c} \). Then, the expression \( \left( \frac{\mu_i}{(1-q)\mu_h} \right)^n V(n) \) is not strictly decreasing in \( n \). Therefore, it may be for some \( n \) that \( \bar{p}(n) < \underline{p}(n) \), so that there cannot be a pure strategy equilibrium in which the uninformed consumer’s strategy has a hole at \( n \).

The claim in part (i) establishes that \( \frac{V(n)}{V(n+1)} \) is a U-shaped function, strictly decreasing for \( n < \frac{v_h\mu_h + v_i \mu_i}{2c} - 2 \), and strictly increasing for \( n > \frac{v_h\mu_h + v_i \mu_i}{2c} - 2 \). Therefore, the set of queue lengths at which \( \frac{\mu_i}{(1-q)\mu_h} > \frac{V(n)}{V(n+1)} \), may be represented as \( \{n + 1, n + 2, \ldots, \bar{n} - 2, \bar{n} - 1\} \), where \( \bar{n} = \min_{n \in \{n_\ell, \ldots, n_h-1\}} \{n \mid \delta(n) < \delta(n+1)\} \) and \( \bar{n} = \max_{n \in \{n_\ell, \ldots, n_h-1\}} \{n \mid \delta(n) < \delta(n+1)\} - 1 \). Then, following the same steps as in part (i) of the proposition, for each \( n \in \{n_\ell, \ldots, \bar{n}\} \) and \( \bar{n}, \ldots, n_h \), there exist values of \( p \) for which, in equilibrium, the uninformed consumer’s strategy has a hole at \( n \).

Finally, observe that \( \bar{p}(n_\ell) > p(n_\ell) = 0 \), so that for \( p \leq \bar{p}(n_\ell) \), there is a pure strategy equilibrium with the uninformed consumer failing to join at queue length \( n_\ell \). Therefore, it must be that \( n \geq n_\ell \). Further, \( \delta(n_h) < 0 < \delta(n) \) for any \( n \in \{n_\ell, \ldots, n_h-1\} \). Therefore, in equilibrium the queue length \( n_h \) can also be supported as a hole in the uninformed consumer’s strategy; that is, \( \bar{n} \leq n_h \). \( \blacksquare \)
Proof of Lemma 3

Suppose a firm with type $\theta$ chooses a service rate $\mu$, and suppose consumers play $\sigma$, where $\sigma$ is a pure strategy that satisfies the necessary conditions for equilibrium in the continuation game. Then, $s_{h}(j, \sigma) = 1$ for $j \leq \hat{n}(\sigma) - 1$ and $j \in \{\hat{n}(\sigma) + 1, \ldots, \hat{n}_{h}(\sigma) - 1\}$, with $s_{h}(\hat{n}(\sigma)) = q$ (since only informed consumers join at $\hat{n}(\sigma)$). Substituting for $s_{h}(j, \sigma)$ in equation (3) in Lemma 1 yields

$$\pi_{h}(0, \mu, \sigma) = \frac{1}{\sum_{j=0}^{\hat{n}(\sigma)}(\Lambda/\mu)^{k} + q\sum_{j=\hat{n}(\sigma)+1}^{\hat{n}_{h}(\sigma)}(\Lambda/\mu)^{k}}.$$ \hspace{1cm} (26)

By inspection, we observe that regardless of whether $\Lambda/\mu$ is greater or less than 1, $\pi_{h}(0, \mu, \sigma)$ declines in $\hat{n}(\sigma)$ when $\hat{n}_{h}(\sigma)$ is kept fixed.

Further, $s_{q}(j, \sigma) = 1$ for $j \leq \hat{n}_{q}(\sigma) - 1$, with $s_{q}(j, \sigma) = 1 - q$ for $j \in \{\hat{n}_{q}, \ldots, \hat{n}(\sigma) = 1\}$ (since only uninformed consumers join at these queue lengths) and $s_{q}(j, \sigma) = 0$ for $j \geq \hat{n}(\sigma)$. Substituting for $s_{q}(j, \sigma)$ in equation (3) in Lemma 1 yields

$$\pi_{q}(0, \mu, \sigma) = \frac{1}{\sum_{j=0}^{\hat{n}_{q}(\sigma)}(\Lambda/\mu)^{k} + \sum_{j=\hat{n}(\sigma)+1}^{\hat{n}_{h}(\sigma)}(1-q)^{k-\hat{n}_{q}(\sigma)}(\Lambda/\mu)^{k}}.$$ \hspace{1cm} (27)

By inspection, we observe that $\pi_{q}(0, \mu, \sigma)$ declines in $\hat{n}(\sigma)$, keeping $\hat{n}_{q}(\sigma)$ fixed.

The expected revenue of firm $\theta$ is $R_{\theta}(\mu, \sigma) = r\mu(1 - \pi_{\theta}(0, \mu, \sigma))$, from which the expressions in the statement of the Lemma follow. Since $\pi_{\theta}(0, \mu, \sigma)$ decreases as $\hat{n}(\sigma)$ increases (keeping fixed $\hat{n}_{h}$ and $\hat{n}_{q}$), it follows that for each $\theta$, the expected revenue $R_{\theta}(\mu, \sigma)$ increases as $\hat{n}(\sigma)$ increases.

Proof of Lemma 4

Observe that under Assumption 3 (iii), regardless of firm’s choices over service rate, there is a pure strategy equilibrium in which the strategy of the uninformed consumers has a hole at $n_{q}$. Further, given Assumption 3 (ii), $n_{q}$ and $n_{h}$ are invariant to firms’ service rate choices.

Substitute $\hat{n} = n_{q}$ into the revenue rate expressions in Lemma 3. We obtain

$$R_{h}(\mu, \sigma) = r\mu \left(1 - \frac{1}{\sum_{j=0}^{n_{q}}(\Lambda/\mu)^{j} + q\sum_{j=n_{q}+1}^{n_{h}}(\Lambda/\mu)^{j}}\right),$$ \hspace{1cm} (28)

$$R_{q}(\mu, \sigma) = r\mu \left(1 - \frac{1}{\sum_{j=0}^{n_{q}}(\Lambda/\mu)^{j}}\right).$$ \hspace{1cm} (29)

First, consider the high-quality firm. We have

$$\frac{\partial R_{h}}{\partial \mu} = r \left[1 - \frac{1}{\sum_{j=0}^{n_{q}}(\Lambda/\mu)^{j} + q\sum_{j=n_{q}+1}^{n_{h}}(\Lambda/\mu)^{j}} - \frac{\sum_{j=0}^{n_{q}}(\Lambda/\mu)^{j} - q\sum_{j=n_{q}+1}^{n_{h}}(\Lambda/\mu)^{j}}{[\sum_{j=0}^{n_{q}}(\Lambda/\mu)^{j} + q\sum_{j=n_{q}+1}^{n_{h}}(\Lambda/\mu)^{j}]^{2}}\right].$$
Now, when $\mu = \Lambda$,
\[
\frac{\partial R_h}{\partial \mu} = r \left[ 1 - \frac{1}{n_\ell + 1 + q(n_h - n_\ell - 1)} - \frac{n_\ell(n_\ell + 1)/2 - q(n_h(n_h + 1)/2 - (n_\ell + 1)(n_\ell + 2)/2)}{[n_\ell + 1 - q(n_h - n_\ell - 1)]^2} \right].
\]
Collecting terms and simplifying yields the expression in the statement of the Lemma.

Similarly, for the low-quality firm, we have
\[
\frac{\partial R_\ell}{\partial \mu} = r \left[ 1 - \frac{1}{\sum_{j=0}^{n_\ell} (\Lambda/\mu)^j} - \frac{\sum_{j=0}^{n_\ell} j(\Lambda/\mu)^j}{\left(\sum_{j=0}^{n_\ell}(\Lambda/\mu)^j\right)^2} \right].
\]
Hence, when $\mu = \Lambda$,
\[
\frac{\partial R_\ell}{\partial \mu} = r \left[ 1 - \frac{1}{n_\ell + 1} - \frac{n_\ell(n_\ell + 1)/2}{(n_\ell + 1)^2} \right]
= \frac{rn_\ell}{2(n_\ell + 1)}.
\]

**Proof of Proposition 4**

Consider the expressions for the rate of change of revenue (with respect to $\mu$) for each type of firm in Lemma 4, when $\mu = \Lambda$. Write $n_h = n_\ell + m$, where $m \geq 1$ is an integer. Then,
\[
\frac{\partial R_h}{\partial \mu} (\Lambda, \hat{\sigma}) = \frac{r[2m^2q^2 + m(2n_\ell + 1 - m)q + n_\ell(n_\ell + 1)]}{2(n_\ell + 1 + qm)^2}.
\]
Let $\Gamma = \frac{1}{r} \left[ \frac{\partial R_h}{\partial \mu} (\Lambda, \hat{\sigma}) - \frac{\partial R_\ell}{\partial \mu} (\Lambda, \hat{\sigma}) \right]$. Then,
\[
\Gamma = \frac{2m^2q^2 + m(2n_\ell + 1 - m)q + n_\ell(n_\ell + 1)}{2(n_\ell + 1 + qm)^2} - \frac{n_\ell}{2(n_\ell + 1)}
= \frac{(n_\ell + 1)[2m^2q^2 + m(2n_\ell + 1 - m)q + n_\ell(n_\ell + 1)]}{(n_\ell + 1 + qm)^2(n_\ell + 1)} - (n_\ell + 1 + qm)^2 n_\ell
\]
\[
2\Gamma = \frac{(n_\ell + 1)[2m^2q^2 + m(2n_\ell + 1 - m)q + n_\ell(n_\ell + 1)] - (n_\ell + 1 + qm)^2 n_\ell}{2(n_\ell + 1 + qm)^2(n_\ell + 1)}.
\]

The sign of $\Gamma$ equals the sign of numerator in equation (31), since the denominator is strictly positive. The numerator may be evaluated as:
\[
(n_\ell + 1)[2m^2q^2 + m(2n_\ell + 1 - m)q + n_\ell(n_\ell + 1)] - (n_\ell + 1 + qm)^2 n_\ell
= 2(n_\ell + 1)m^2q^2 + (2n_\ell + 1)(n_\ell + 1)mq - m^2(n_\ell + 1)q + n_\ell(n_\ell + 1)^2
- n_\ell[(n_\ell + 1)^2 + 2(n_\ell + 1)mq + m^2q^2]
= (n_\ell + 2)m^2q^2 - (n_\ell + 1)mq(m - 1) = mq[(n_\ell + 2)mq - (n_\ell + 1)(m - 1)].
\]
Hence, the sign of the numerator is equal to the sign of the expression $[(n_\ell + 2)mq - (n_\ell + 1)(m - 1)]$. 

39
Now, recall that  \( \hat{q} = \left(1 - \frac{1}{n_h - n_\ell}\right) \frac{n_\ell + 1}{n_h + 2} = \left(\frac{m - 1}{m}\right) \frac{n_\ell + 1}{n_h + 2}. \) Hence, it follows that the numerator in equation (31) is positive if \( q > \hat{q} \) and negative if \( q < \hat{q} \). Therefore, if \( q > \hat{q} \), \( \frac{\partial R_h}{\partial \mu}(\Lambda, \hat{\sigma}) > \frac{\partial R_h}{\partial \mu}(\Lambda, \hat{\sigma}) \), and if \( q < \hat{q} \), \( \frac{\partial R_h}{\partial \mu}(\Lambda, \hat{\sigma}) < \frac{\partial R_h}{\partial \mu}(\Lambda, \hat{\sigma}) \).

We now turn to the two parts of the statement of the proposition.

(i) Suppose \( q < \hat{q} \). Observe that for each \( \theta = h, \ell \), \( \lim_{\epsilon \to 0} \Delta \hat{q}(\Lambda, \epsilon) = \frac{\partial R_h}{\partial \mu}(\Lambda, \hat{\sigma}) \). Since \( \frac{\partial R_h}{\partial \mu} < \frac{\partial R_h}{\partial \mu} \), it follows that there exists an \( \hat{\epsilon}(q) > 0 \) such that, for all \( \epsilon \in (0, \hat{\epsilon}] \), \( \Delta_h(\Lambda, \epsilon) < \Delta_\ell(\Lambda, \epsilon) \).

Now, for any \( k \in (\Delta_h(\Lambda, \epsilon), \Delta_\ell(\Lambda, \epsilon)) \), the high-quality firm chooses the slow service rate \( \Lambda - \epsilon \) and the low-quality firm chooses the fast service rate \( \Lambda + \epsilon \).

Further, since \( \Delta_h(\Lambda, \epsilon) < \Delta_\ell(\Lambda, \epsilon) \), there cannot exist a cost \( k \) such that the high-quality firm chooses the fast service rate \( \Lambda + \epsilon \) and the low-quality firm chooses the slow service rate \( \Lambda - \epsilon \).

(ii) The argument when \( q > \hat{q} \) is exactly similar to the argument in (i) above. \( \blacksquare \)

**Proof of Proposition 5**

(i) Consider the expression for \( \frac{\partial R_h}{\partial \mu} \) when \( \mu = \Lambda \), as shown in equation (30) in the proof of Proposition 4. Recall that \( n_h = n_\ell + m \) in this expression. Consider the limit as \( n_h \to \infty \), or, equivalently, \( m \to \infty \).

In the limit as \( n_h \to \infty \), we obtain

\[
\lim_{n_h \to \infty} \frac{\partial R_h}{\partial \mu}(\Lambda, \hat{\sigma}) = \frac{r(2q^2 - q)}{2q^2} = r\left(1 - \frac{1}{2q}\right). \tag{32}
\]

The last expression is strictly negative whenever \( q < \frac{1}{2} \). Thus, for any \( q < \frac{1}{2} \), there exists a threshold \( \hat{n}(q) \) such that if \( n_h \geq \hat{n}(q), \frac{\partial R_h}{\partial \mu}(\Lambda, \hat{\sigma}) < 0 \). Similarly, for any \( q > \frac{1}{2} \), there exists a threshold \( \hat{n}(q) \) such that if \( n_h \geq \hat{n}(q), \frac{\partial R_h}{\partial \mu}(\Lambda, \hat{\sigma}) > 0 \). Define \( \hat{\nu}(q) = \hat{n}(q) \frac{\epsilon}{\mu_0 - \epsilon} \). Then, the condition \( n_h \geq \hat{n}(q) \) is equivalent to \( \nu_h \geq \hat{\nu}(q) \).

Now, \( \lim_{\epsilon \to 0} \Delta_h(\Lambda, \epsilon) = \frac{\partial R_h}{\partial \mu}(\Lambda, \hat{\sigma}) \). Suppose \( q < \frac{1}{2} \) and \( \nu_h \geq \hat{\nu}(q) \). Then, if \( \epsilon \) is sufficiently close to zero, if must be the case that \( \Delta_h(\Lambda, \epsilon) < 0 \). Formally, there exists some \( \hat{\epsilon}(q) \) such that if \( \epsilon \in (0, \hat{\epsilon}(q)] \), then \( \Delta_h(\Lambda, \epsilon) < 0 \). Hence, for parameters in this range, at any value of \( k \), including \( k = 0 \), the high-quality firm will choose the slow service rate \( \Lambda - \epsilon \).

Finally, suppose \( q > \frac{1}{2} \) and \( \nu_h \geq \hat{\nu}(q) \). Then, there exists some \( \hat{\epsilon}(q) \) such that if \( \epsilon \in (0, \hat{\epsilon}(q)] \), then \( \Delta_h(\Lambda, \epsilon) > 0 \). \( \blacksquare \)

**Proof of Proposition 6**

Consider the expression for \( R_h \) in equation (28), in the proof of Lemma 4. Let \( n_h \to \infty \),
and denote $\rho = \Lambda/\mu$. Let $\bar{R}_h = \frac{1}{r} \lim_{n_h \to \infty} R_h(\mu, \hat{\sigma})$. Then, we obtain

$$
\bar{R}_h(\mu, \hat{\sigma}) = \mu \left( 1 - \frac{1}{\sum_{j=0}^{n_h} \rho^j + q \lim_{n_h \to \infty} \sum_{j=n_{\ell}+1}^{n_h} \rho^j} \right)
$$

Hence, if $\rho > 1$ (i.e., $\mu < \Lambda$), $\bar{R}_h = \mu$. If $\rho < 1$,

$$
\bar{R}_h(\mu, \hat{\sigma}) = \mu \left( 1 - \frac{1 - \rho^{n_{\ell}+1} + q \rho^{n_{\ell}+1}}{1 - (1 - q)\rho^{n_{\ell}+1}} \right) = \mu \left( 1 - \frac{1 - \rho^{n_{\ell}+1}}{1 - (1 - q)\rho^{n_{\ell}+1}} \right)
$$

It follows that for $\rho < 1$,

$$
\frac{\partial \bar{R}_h}{\partial \rho} = \frac{\Lambda}{[1 - (1 - q)\rho^{n_{\ell}+1}]^2} \left\{ - (1 - \rho)n_{\ell} + \rho[1 - (1 - q)\rho^{n_{\ell}}] \right\}
$$

Let $\Psi(\rho) = \rho[1 - (1 - q)\rho^{n_{\ell}}] - (1 - \rho)n_{\ell}$. Then, the sign of $\frac{\partial \bar{R}_h}{\partial \rho}$ is equal to the sign of $\Psi(\rho)$.

Now, $\Psi(0) = -n_{\ell} < 0$ and $\Psi(1) = q > 0$. Further,

$$
\Psi'(0) = 1 - (1 - q)\rho^{n_{\ell}} - n_{\ell}(1 - q)\rho^{n_{\ell}} + n_{\ell} = (n_{\ell} + 1)[1 - (1 - q)\rho^{n_{\ell}}] > 0.
$$

Hence, there exists a unique $\hat{\rho}(n)$ which solves the equation

$$
\rho[1 - (1 - q)\rho^{n}] - (1 - \rho)n = 0, \text{ or, } \frac{\rho}{1 - \rho}[1 - (1 - q)\rho^{n}] = n.
$$

Further, for any $n_{\ell}$, if $\rho \in (\hat{\rho}(n_{\ell}), 1)$, it follows that $\frac{\partial \bar{R}_h}{\partial \rho}$ is increasing in $\rho$. Since $\rho = \frac{\Lambda}{n}$, it follows immediately that if $\rho \in (\hat{\rho}(n_{\ell}), 1)$, then $\frac{\partial \bar{R}_h}{\partial \mu} < 0$. That is, $\bar{R}_h$ is decreasing in $\mu$.

That is, if $\mu \in (1, \frac{\Lambda}{\hat{\rho}(n_{\ell})})$, it follows that $\frac{\partial \bar{R}_h}{\partial \mu} < 0$. Now, observe that for any finite $v_h$,

$$
\frac{\partial \bar{R}_h}{\partial \mu}(\mu_0, \hat{\sigma}) = \lim_{\epsilon \to 0} \Delta(\mu_0, \epsilon). \text{ Hence, } \frac{\partial \bar{R}_h}{\partial \mu}(\mu_0, \hat{\sigma}) = \lim_{n_h \to \infty} \lim_{\epsilon \to 0} \Delta(\mu_0, \epsilon). \text{ Hence, there exists an } n_h \text{ large enough (alternatively a } v_h \text{ large enough) and an } \epsilon' \text{ small enough such that the statement of the proposition follows.}
$$
References


