"We Will be Right with You:" Managing Customer Expectations with Vague Promises and Cheap Talk

Gad Allon

Kellogg School of Management, 2001 Sheridan Road Evanston, IL 60208 , g-allon@kellogg.northwestern.edu

Achal Bassamboo

Kellogg School of Management, 2001 Sheridan Road Evanston, IL 60208 , a-bassamboo@kellogg.northwestern.edu

Itay Gurvich

Kellogg School of Management, 2001 Sheridan Road Evanston, IL 60208, i-gurvich@kellogg.northwestern.edu

September 6, 2011

Delay announcements informing customers about anticipated service delays are prevalent in service-oriented systems. How delay announcements can influence customers in service systems is a complex problem which depends on both the dynamics of the underlying queueing system and on the customers' strategic behavior. We examine this problem of information communication by considering a model in which both the firm and the customers act strategically: the firm in choosing its delay announcement while anticipating customer response, and the customers in interpreting these announcements and in making the decision about when to join the system and when to balk. We characterize the equilibrium language that emerges between the service provider and her customers. The analysis of the emerging equilibria provides new and interesting insights into customer-firm information sharing. We show that even though the information provided to customers is non-verifiable, it improves the profits of the firm and the expected utility of the customers. The robustness of the results is illustrated via various extensions of the model. In particular, studying models with incomplete information on the system parameters allows us also to highlight the role of information provision in managing customer expectations regarding the congestion in the system. Further, the information could be as simple as "High Congestion"/"Low Congestion" announcements, or could be as detailed as the true state of the system. We also show that firms may choose to shade some of the truth by using *intentional vagueness* to lure customers.*

1. Introduction

Delay announcements that inform customers about anticipated delays are prevalent in service systems. Call centers often use recorded announcements to inform callers of the congestion in the system and encourage them to wait for an available agent. While some of these announcements do not provide much information - such as the common message, "Due to high volume of calls, we are unable to answer your call immediately," some call centers go as far as providing the customer with an estimate of his waiting time or his place in the queue. In many service systems where the real state of the system is not visible to customers, delay announcements affect customers' behavior and may, in turn, have significant impact on system performance. Consequently, the service provider can use the delay announcements as a tool to induce the appropriate customers' behavior, stirring the system so as to maximize her own profits. How to use delay announcements to manage the service system in an efficient manner is a complex problem, and its answer depends on both the dynamics of the underlying queueing system and the customer behavior.

Previous papers addressing this issue analyzed two categories of information provided to the customer: (i) full state information - either exact waiting time information (if such information is available to the service provider) or the state of the system when the customer arrives (or an estimate based on the state), and (ii) no information - where no information is provided, and customers must base their decisions on their expectation regarding the system performance. The main assumption made in these papers is that customers treat the information provided in the delay announcements as a-priori verified (i.e., credible), and make their decisions whether to join or leave the system, accordingly. The two main issues with this assumption are: (i) Customers may not be naive in their attitude towards any information provided by interested parties, and thus take such announcements with a "grain of salt." Moreover, under the assumption of "naivety," it makes sense for the firm to deviate from the truth-telling policy. The option that the firm might lie, given that the customer always believes the firm, is never explored in the literature. (ii) Further, prior work implicitly assumes that the announcements have a quantifiable meaning in terms of place-in-queue or average waiting time. However, as stated above, many service providers use verbal messages that need to be further processed in order for customers to make the decision whether to join or not. For

example, without processing, it is not clear what "high volume of calls" means. This problem is clearly a consequence of the first issue, since without processing, only quantifiable announcements are possible. Thus the customers are able to evaluate their utility only for simple (i.e., no-information or full-state-information) announcement and only those are discussed. In practice, however, we observe a much richer variety of announcements used by service providers.

This paper addresses these issues by proposing a model in which customers treat information provided by the service provider as unverified and non-binding (that is non-contractable). The model treats customers as strategic in the way they process information (as well as in making the decisions whether to join or balk) and the service provider as strategic in the way she provides the information. The customers and the service provider are assumed to be self-interested in making their decisions: the service provider in choosing which announcements to make, and the customers in interpreting these and making their decisions.

This allows us to characterize the equilibrium language that emerges between the service provider and her customers. (We use the term equilibrium language, in accordance with the usage in the cheap talk literature, to refer to the equilibrium emerging in the cheap talk game. The phrase "language" is used to emphasize the signaling rule and its associated actions.) By doing that, not only do we relax the assumption that customers may be naive in their treatment of the announcements, but we also demonstrate that many of the commonly used announcements arise in *equilibrium* in such a model. The spectrum of possible equilibria will range from announcements that are analogous to the non-quantifiable type that describe the volume of arriving customers as high or low, to the more quantifiable ones that provide waiting time announcements, both of which are common in different systems. While prior work focused on systems with either full state information or no information, our model shows that a richer language, which usually includes intentional vagueness, arises as an equilibrium outcome in a game played between the service provider and her customers. By intentional vagueness we mean that under any equilibrium the firm provides the same announcement on multiple states of the system. In developing the model and characterizing the emerging equilibrium language, we will account both for the strategic nature of the interested parties – customers and service provider – as well as for the queueing dynamics prevalent to service systems.

This paper proposes what appears to be the first model to deal with the strategic nature of the information transmission in a practical operational setting, where unverifiable[1], non-committal, real-time information is provided by a self-interested firm to selfish customers.

We treat the announcements made by the system manager as "cheap talk," i.e., pre-play communication that carries no cost. Cheap talk consists of costless[2], non-binding, non-verifiable messages that may affect the customer's beliefs. While providing the information does not *directly* affect the payoffs, it has an indirect implication through the customer's reaction and the equilibrium outcomes. The information has no impact on the payoffs of the different players per se i.e., the payoffs of both sides depend only on the actions taken by the customers and queueing dynamics. This, in turn, means that if a customer does not follow the recommendation made by the firm, he is not penalized, nor is he rewarded when he follows it. However, as it will be shown, the announcements do have an impact on the service provider's profits and the customers' utility, in equilibrium. This is in agreement with both the cheap talk literature (See Crawford and Sobel (1982)) and the queueing literature with strategic customers (See Naor (1969)), where the information provided to the customer is in the form of full visibility of the queue and does not alter the customer's utility directly; it allows him to make a knowledgeable decision whether to join or not, and thus affects his utility in an indirect manner. Our model echoes the framework developed by Crawford and Sobel (1982) but the specifics of our model are different from theirs in significant aspects. We examine these differences more closely in Section 1.2.

This paper focuses on the *strategic* interaction between the customer and the firm in a setting in which their incentives are *misaligned*, when *unverifiable, costless, and non-binding* information is provided to the customer. In all of the instances described in this paper, the information is always unverifiable and has no contractual bearing (i.e., it is non-binding). This stands in contrast to service-level *guarantees*, such as those made by Dominos Pizza, Ameritrade, and E∗trade to name a few, where the commitment is contractually

---

[1] In our context we use the term unverifiable to describe situations in which the customer cannot verify statements regarding the state of the system.

[2] We assume that the cost associated with conveying the message is negligible. In most practical service organizations, while the provider needs to incur fixed costs, for example, by investing in a more sophisticated IT infrastructure to identify the state of the system, the marginal cost of providing the information to the customer is insignificant. There is a voluminous literature that deals with models in which signaling is not costless, and the mere fact that players are willing to incur a cost provides a signal.

binding; see Allon and Federgruen (2007) for a more detailed discussion of the delay guarantees provided by these firms.

Our main findings are as follows:

(i) We characterize the equilibrium of the game played by the service provider and the customers. The characterization depends both on the queueing dynamics (based on a Markov Decision Process), and the strategic nature of the players. We show that, while the characterization is complex, under pure strategies, an equilibrium can be mapped into a single threshold level. The firm provides two signals to indicate whether the level of congestion is below or above the threshold. The customer makes his decision according to the expected congestion level implied by each signal.

(ii) We show that different types of misalignment may lead to existence or non-existence of pure-strategy equilibria. Further, we characterize a two-signal equilibrium when it exists, and show that any equilibrium with more than two signals is outcome-equivalent to the simpler two-signal one. Our results imply that some commonly used non-quantifiable announcements regarding the congestion level in the system can be reduced to a two-signal equilibrium.

(iii) We show that, even if customers are allowed to randomize between multiple actions (i.e., join or balk), any equilibrium with more than two signals will be outcome-equivalent to the two-signal equilibrium. Furthermore, under any equilibrium where the customer randomizes, it is not possible for the firm to have a signaling rule that includes both a signal that induces the customers to balk with probability 1 and a signal that induces the customers to join with probability 1.

(iv) We prove that there always exists a *most informative equilibrium*, i.e., an equilibrium with the finest partition of the state space revealed to the customer by the service provider. The firm employs *intentional vagueness* in this equilibrium, unless the customers and the firm are perfectly aligned. By perfect alignment, we mean that the self-interested customers agree with the profit maximizing firm on the preferred action for every state of the system. The role of intentional vagueness is to lure customers to join the system in states in which, under full state information, the customer would not join, yet the congestion in the system is not too high. The informative equilibria, either with or without intentional vagueness, are shown to be

equivalent to the common practices of announcing the location in the queue, an estimate of the average waiting time, or a confidence interval of the expected waiting time.

(v) We show that a babbling equilibrium where the customer disregards any information provided by the firm (or the firm does not provide any information) always exists, either in pure or mixed strategies. We then show that the service provider and the customers always prefer other, more informative, equilibria over a babbling equilibrium.

(vi) We also test the robustness of the above results by considering models in which we relax the information structure and those in which customers adaptively update their beliefs about the meaning of messages adaptively. We find that most of the findings continue to hold with appropriate modifications.

Organization of the remainder of the paper: The rest of this section is divided into two parts: the first reviews the antecedent literature and the second briefly describes the Crawford and Sobel (1982) game. Section 2 provides the detailed description of our model. In Section 3, we define our notion of equilibrium and state our main results. In Section 4 we discuss the signaling language that emerges in equilibrium and, in particular, introduce and analyze the notion of most informative equilibria. Section 5 concludes the paper. All the proofs are relegated to the online supplement. Extensions of the model, as well as a numerical study are also discussed in the online supplement.

1.1. Literature Review

Queueing models with strategic customers. The literature on queueing models with strategic customers dates back to Naor (1969), who studied a system in which strategic customers observe the length of the queue prior to making the decision whether to join or balk. There is a (partial) conflict of interest between the self-interested customer and the interests of the social-welfare-maximizing service provider. Naor (1969) shows that pricing can be used to achieve the first-best solution. The follow-up literature extends Naor (1969) along multiple dimensions. One such stream studies models where the firm offers different grades of services (see the recent paper by Afeche (2004) and the referenced therein). Another stream focuses on competition in the presence of congestion-sensitive customers (see Allon and Federgruen (2007) and the references therein). All of the above papers assume that the firms provide long-run averages information,

rather than real-time information, and that the information is credible, and is treated as such by the customers. The above models assume that customers are rational when making the decision of which line to choose or when to join versus balk from a queue. For more detailed discussion we refer the reader to Hassin and Haviv (2003) and the references therein.

Queueing models with delay announcements. Hassin (1986) studies the problem of a price-setting revenue-maximizing service provider that has the option to reveal the queue length to arriving customers, but may choose not to disclose this information. The author shows that it may be socially optimal to force the firm to provide full state information. Armony and Maglaras (2004b) analyzes a service system where arriving customers can decide whether to join, balk, or wait for the provider to call within a guaranteed time. The customers' decisions are based on the equilibrium waiting time. Armony and Maglaras (2004a) extends the above model to allow the service manager to provide the customers an estimate of the delay, based on the real-time state of the system upon their arrival. The authors show that providing information on the estimated delay improves the system performance. Armony et al. (2007) studies the impact of delay announcements on the performance of a many-server queue with customer abandonment.

Dobson and Pinker (2006) develops a stochastic model of a custom production environment with pricing, where customers are heterogeneous in terms of their tolerances for waiting. The authors model intermediate levels of information sharing (with a specific structure) ranging from no-sharing to complete state-dependent lead-time information, and compare the resulting performance from the firms and customers perspectives. Guo and Zipkin (2007) studies a model in which customers are provided with information and make decisions based on their expected waiting times, conditional on the provided information. Three types of information are studied: (i) no information, (ii) queue length information, and (iii) exact waiting time information (in systems in which such information is available). The authors provide examples in which accurate delay information improves the system performance, as well as examples in which it does not. Jouini et al. (2007) is the first paper to consider delay announcements in a multi-class setting with priorities. All of these models assume that the information provided to the customer is truthful and that the customer believes the information, even if he will be better off, in terms of his own utility, disregarding it.

1.2. Classical Cheap Talk Game

In this section, we provide an overview of the cheap talk game introduced in Crawford and Sobel (1982) and compare the model to the one studied in this paper. The classical cheap talk game is played between a *sender* who has some private information and a *receiver* who takes the payoffs-relevant actions. The game proceeds as follows: The Sender observes his state/type, which we shall denote by Q. The Sender then sends a signal (or a message) denoted by $m \in \mathcal{M}$, where $\mathcal{M}$ denotes the set of all signals. The Receiver, who cannot observe the state/type Q, but does know its distribution, processes the signal using Bayes rule and chooses an action y which determines the players payoff. Both the Sender and the Receiver obtain utilities which depend on: (a) the action taken by the Receiver, y; and (b) the state/type Q. Two distinctive features of their model should be emphasized: first, the state/type, while random, is static: once realized it does not change over time. Also, the distribution of this state/type is exogenous and independent of the actions of the players.

A variety of papers study mixed-motive economic interactions involving private information and the impact of cheap talk on the outcomes. In political context cheap talk has been studied in multiple papers including Austen-Smith (1990). Recently, Chen et al. (2008) studies questions regarding equilibrium selection in cheap talk games.

In contrast to Crawford and Sobel (1982), our model has two key features: first, the game is played with multiple receivers (customers) whose actions have externalities on other receivers; and second, the stochasticity of the state/type (i.e., the state of the system) is not exogenously given but is determined endogenously. The impact of edogenization of the state of the world already appears in models of political debate, see Austen-Smith (1990). In our model, these features are motivated by applications in service operations. Indeed, the private information in this model (i.e., the queue length) is driven by the system dynamics, which in turn depend on the equilibrium strategies of both the firm and the customers. In particular, in our model, the customers' actions are both payoff-relevant and system-dynamic-relevant. As we will show, the multiplicity of receivers with externalities as well as the endogenization of the uncertainty impact both the nature of the communication between the firm and the customers as well as the outcome for the

various players. The way these features separate our results from those of Crawford and Sobel (1982) will be discussed in Section EC.1.

2. Model

We consider a service provider modeled as an M/M/1 queue. Customers arrive to the system according to a Poisson process with rate λ. Service times are exponentially distributed with mean $1/\mu$.[3] We assume that $\lambda < \mu$. We assume that all customers are ex-ante symmetric: customers obtain a value R if they are served, and incur a waiting cost that is proportional to the time spent in the system, with a unit waiting cost of c.[4] Thus, the utility function of a customer is:

$$U(y,w) = \begin{cases} R - cw & \text{if } y = \text{``join''}, \\ 0 & \text{if } y = \text{``balk''}, \end{cases} \tag{1}$$

where y is the decision made by the customer and w denotes his sojourn[5] time in the system. Throughout the paper, we shall assume that $R > c/\mu$. This assumption ensures that in the absence of delays, the service is beneficial to the customer, on average. Clearly, if $R < c/\mu$, no customer will join regardless of the system announcements. Throughout the paper we assume that involuntary blocking by the firm is prohibitively expensive.

The firm's profit from a customer served is $v > 0$. The firm incurs a holding cost $h(w)$ per customer who waits w in the system. This holding cost is incurred by the firm due to the customer waiting in the system. This cost constitutes, among other factors, the cost associated with loss of goodwill, the actual cost of holding the customer and in some settings the opportunity cost associated with the customer not being able to generate revenues. Disney's theme parks, for example, incur two costs due to waiting: the opportunity cost of having a customer standing in line without the ability to spend money and the wages of the entertainment staff that is in charge of alleviating the pain of waiting. Thus, the firm's profit from a customer is $v - h(w)$ for a customer who received service from the firm and spent w in the system and is 0 for the customer who decides not to join. Here, h is a non-negative and convex increasing function. We

---

[3] We discuss a model where the firm selects the service rate μ optimally in Appendix EC.3.2.

[4] If the customer's cost of waiting is non-linear, then the main results of the paper continue to hold. However, some of the results in the Appendices hinge on this linearity. We will comment more elaborately on those in the appendices.

[5] For a model with waiting time in the queue, see Allon and Bassamboo (2009a).

assume that $v - \mathbb{E}h(X) > 0$, where X is an exponential random variable with mean $1/\mu$ to avoid cases in which even customers that arrive without any waiting are not profitable. We assume that customers who join the system are served in a First-Come-First-Served manner.

Information Provision and Sequence of Events. The static structure described above is assumed to be common knowledge. The real-time state of the system, corresponding to the number of customers waiting in queue, is the firm's private information. The evolution of this state will depend on the equilibrium strategies of both the provider and the customers. The interaction between the customers and the firm is described as follows: upon a customer's arrival to the system, the firm, that can observe the queue length, makes an announcement to this customer. The customer reacts to that signal by deciding whether to join or not join (balk). In making the decision, the customer uses the information, denoted by I, that he can infer from the firm's announcement regarding the current state of the system and chooses an action that maximizes his expected utility. Therefore the customer will join if and only if $R \geq c\mathbb{E}(w|I)$. The type of signals that are used by the firm, and the way the customers interpret these in making their decisions, are described precisely in the next section.

Note that the customers' and the service provider's incentives are not completely misaligned: both prefer short waiting times, which result in higher utility for the customer and higher profits for the service provider. At the same time, we observe that these incentives are not perfectly aligned and this would lead to the equilibria described in the next section. We refer the reader to Farrell and Rabin (1996) for a discussion of settings in which incentives are perfectly misaligned. We formally define the concept of misalignment for our setting in Section 3.

3. Main Results

3.1. Problem formulation

In practice, one observes different types of messages that are conveyed to the customers. These messages could provide tangible information such as expected wait in the queue or the position of the customer in the queue. However, in some cases, the message may only have intangible information such as the congestion level or volume of calls being high or low. We propose a framework that allows us to study the provision of

unverifiable information using a unified approach that does not assume a specific type of the announcement. The framework accounts for the following key features: a) the state-of-world changes dynamically; b) the customer cannot verify the information provided by the firm; and c) the customers would process any information provided to them by the firm and base their action on it. The structure that we lay out next covers the suggested framework that allows us not to restrict attention to any specific type of announcements. In particular, it allows us to account for announcements of the variety mentioned above.

In this section we formally define the game between the service provider and the customers. The equilibrium concept we employ is one of Markov Perfect Bayesian Nash Equilibrium (MPBNE), which, in this case, is simply a Nash equilibrium in the decision rules that relate agents' actions to their information and to the situation in which they find themselves, while allowing for actions to depend only on payoff-relevant histories.

To formally define the game let $\mathcal{M} = \{m_0, m_1, m_2, \ldots\}$ represent the discrete set of feasible signals that the firm can provide to the customer. For example, the set of feasible messages may contain non-negative integers, rational numbers, concrete messages (those with some tangible information) or non-quantifiable messages (those with only intangible information), among others. We represent the signaling rule by a function $g : \mathbb{Z} \mapsto \mathcal{M}$, where $g(q) = m$ if the firm uses the signal m when the queue length is q.

Recall that customers are indistinguishable and their strategies are ex-ante symmetric, both in their interpretation of the signals and in their actions. Let $y : \mathcal{M} \mapsto \{0, 1\}$ denote the strategy of the customer, where $y(m)$ is the probability that a customer joins when the firm signals m. Consequently, we interpret $y(m) = 1$ as a "join" decision and $y(m) = 0$ as a "balk" decision. We initially restrict ourselves to pure strategies. These provide valuable insights into the structure of information transmission, while keeping the model simple. We will relax this restriction in Section EC.6.

Under a signaling rule g and decision rule y, the queue evolves as a birth and death process on the positive integers where, for any state q, the birth rate is $\lambda y(g(q))$ and the death rate is $\mu \mathbb{I}\{q > 0\}$. Here, we use the notation $\mathbb{I}\{\cdot\}$ to denote the indicator function. As we restrict attention to $\lambda < \mu$, this birth-death process has a unique steady-state distribution. Let $p_q(y, g)$ be the steady-state probability of having q customers in the

system. We let $Q_{y,g}$ be a random variable with the steady state distribution of the number of customers in the system.

The requirements of a Markov Perfect Bayesian Nash equilibrium in our context are explained as follows: Given a signaling rule for the system, customers with an action rule that dictates joining the system when the signal is m will not deviate from this rule if their expected conditional utility from joining the system, given by $\mathbb{E}[R - c(q+1)/\mu | g(q) = m]$, is positive. Given the customer's action rule y, the firm will not deviate from its signaling rule g if it maximizes its steady-state profit, i.e, if g solves an appropriate Markov Decision Process (MDP, see below) with respect to the action rule y. This discussion is formalized in the following definition.

Definition 3.1 (**Markov Perfect Bayesian Nash Equilibrium** $MPBNE$) *We say that the signaling rule $g(\cdot)$ and the action rule $y(\cdot)$ constitute a Markov Perfect Bayesian Nash Equilibrium (MPBNE),[6] if they satisfy the following conditions:*

1. *For each $m \in \mathcal{M}$, we have*

$$y(m) = \begin{cases} 1 & \frac{\sum_{\{q:g(q)=m\}} \left[R - c\frac{q+1}{\mu}\right] p_q(y,g)}{\sum_{\{q:g(q)=m\}} p_q(y,g)} \geq 0, \\ 0 & otherwise. \end{cases}$$

2. *With $f(j) = v - \mathbb{E}[h(W(j+1))]$, where $W(j+1)$ is the time spent by a customer who joins the system with j customers, there exists constants $J_0, J_1, \ldots,$ and γ that solve the following set of equations:*

$$\begin{aligned}
J_0 &= \max_{m \in \mathcal{M}} \left\{ \frac{f(0)y(m) - \gamma}{\lambda} + J_0(1 - y(m)) + J_1 y(m) \right\} \\
&= \frac{f(0)y(g(0)) - \gamma}{\lambda} + J_0(1 - y(g(0))) + J_1 y(g(0)) \\
J_q &= \max_{m \in \mathcal{M}} \left\{ \frac{f(q)y(m) - \gamma}{\lambda + \mu} + \frac{\mu}{\lambda + \mu} J_{q-1} + \frac{\lambda}{\lambda + \mu} (J_q(1 - y(m)) + J_{q+1} y(m)) \right\} \\
&= \left\{ \frac{f(q)y(g(q)) - \gamma}{\lambda + \mu} + \frac{\mu}{\lambda + \mu} J_{q-1} + \frac{\lambda}{\lambda + \mu} (J_q(1 - y(g(q))) + J_{q+1} y(g(q))) \right\} \quad (2)
\end{aligned}$$

---

[6] Note that $p_q(y,g)$ can be thought of as the beliefs of the agents on the state of the system. These beliefs must be consistent with the strategy of the other players.

In the above definition of MPBNE, the first condition uses the Bayesian rule for the customer based on the signaling function g to determine whether to join or balk. The second condition states that the composite function $y \circ g$ solves the firm's MDP which is closely related to the *admission control problem* in the MDP literature. In the optimality equations (EC.8), the constant γ represents the long-run average profit (referred to as the gain in MDP literature) made by the firm under optimal policy, and constants $J_0, J_1, \ldots$ represent the *relative cost or the bias* for states $0, 1, \ldots$. Thus, we require that y, g form an equilibrium only if g is bias optimal. Our setting is very close to Haviv and Puterman (1998) which characterizes the bias optimal policy for an M/M/1 queue, see also Chapter 4 in Bertsekas (2001).

Discussion of the customers' ability and beliefs. The above definition takes the approach that the customer is rational, has full state information about the system parameters, and is capable of forming beliefs about the system performance given a set of strategies. The assumption that the customer has full state information about the arrival rate is relaxed in Section EC.2. In that section, we study a model in which the firm has private information on the arrival rate, and show that all the results of the above model extend to that setting as well. We also relax the assumption that customers are capable of forming beliefs on the system performance given the strategies of all players in the game, i.e., the firm and the customers. This is done in Section EC.3.1 by conducting a numerical study in which customers form their action rules through repeated interactions with the firm. We numerically show that such an iterative mechanism would lead the system to the equilibria characterized in this section. Another implicit assumption made in this definition is that customers make a decision only regarding joining vs. balking, and do not leave the system after joining but before receiving service. This assumption is relaxed in Appendix EC.4. There, we show that as long as the customer experiences linear waiting costs, even if customers are allowed to update their belief about the system and renege the queue, the equilibria characterization remains unchanged.

All these relaxations allow us to test the robustness of the assumptions made in the model above, and show that the resulting characterizations continue to hold when the assumptions are relaxed.

One of the goals of this paper is to identify the conditions under which a firm can credibly communicate unverifiable information. Our litmus test for such credibility will be the existence, or lack thereof, of an equilibrium with influential cheap talk. When such an equilibrium exists, it means that the firm can induce,

by virtue of using at least two distinct messages, two distinct actions. An equilibrium with influential cheap talk is formally defined as follows:

Definition 3.2   *We say that an equilibrium (y, g) has influential cheap talk if there exist two distinct signals m_i, m_j, that are used with positive probability in equilibrium, i.e., $\sum_{\{q:g(q)=m_i\}} p_q(y, g) > 0$ and $\sum_{\{q:g(q)=m_j\}} p_q(y, g) > 0$ such that $y(m_i) \neq y(m_j)$.*

As in most cheap talk games, the equilibrium, even if it exists, is never unique. This stems from the fact that one can relabel the signals or introduce new signals and still have the same outcomes and payoffs for the players.

Definition 3.3 (DOE)   *We say that two MPBNE (y_1, g_1) and (y_2, g_2) are Dynamics- and Outcome- Equivalent (DOE), if $y_1(g_1(q)) = y_2(g_2(q))$ for all $q = 0, 1, \ldots$.*

3.2. Characterizing key equilibria

While Definition 3.1 of the pure strategy MPBNE in the previous section is complete, it is not directly amenable for further analysis. Thus, our first step toward characterizing the equilibria is to show that any pure strategy MPBNE can be described by a threshold level.

Proposition 3.1   *For any pure strategy equilibrium (y, g) with influential cheap talk, there exists $N :=$ $\inf\{q : y(g(q)) = 0\}$ which is finite such that $y(g(q)) = 0$ for all $q \geq N$ and $y(g(q)) = 1$ for all $q < N$.*

The next proposition shows that one can always reduce a pure strategy influential equilibrium to an equilibrium where the firm uses exactly two signals.

Proposition 3.2   *Let the pair (y, g) be a pure strategy MPBNE with influential cheap talk and let $N :=$ $\inf\{q : y(g(q)) = 0\}$. Then the pair $(\widetilde{g}(\cdot), \widetilde{y}(\cdot))$ given by*

$$\widetilde{g}(q) = \begin{cases} m_1 & q \leq N, \\ m_0 & otherwise \end{cases}, \quad \widetilde{y}(m) = \begin{cases} 1 & m = m_1, \\ 0 & otherwise \end{cases} \tag{3}$$

forms a MPBNE with the same firm profit and customer utility. Further, (y, g) and $(\widetilde{y}, \widetilde{g})$ are DOE.

Proposition 3.1 implies that instead of studying the customers' actions and the firm's announcement in each state of the system (i.e., queue length), it suffices to focus on the threshold queue length, below which the customer's action is "join," and above which it is "balk." Note that every equilibrium characterized using Proposition 3.2 requires that the equilibrium has influential cheap talk and thus that the constant $N = \inf\{q : y(g(q)) = 0\}$ is finite. There may exist equilibria where the constant N is infinite. We shall discuss these in Section 3.3.

Definition 3.4   *We say that the threshold q induces a pure strategy MPBNE if the pair $(\widetilde{g}(\cdot), \widetilde{y}(\cdot))$ given by (3) forms an MPBNE.*

Before delving into the analysis of the model and the characterization of the equilibrium, we take a step back and develop intuition into the possible regimes and outcomes. In order to do that, and knowing that we can restrict attention to threshold-based equilibria, we introduce two important threshold levels: the first, q^*, denotes a threshold value above which a customer *will not* join if he has **full state information** of the state of the system, and below which he *will join*. The second threshold level, $\widehat{q}$, is motivated by the service provider's point of view. It denotes the threshold level below which the service provider would like the customers to join, and above which she would like them to balk, if she had **full control** of their actions.

   Full state information. We define q^* to be the threshold above which customers do not obtain positive utility, in expectation, given full queue-length information, that is

$$q^* = \left[\frac{R\mu}{c}\right], \tag{4}$$

where $[\cdot]$ is the bracket function; i.e., q^* is the largest integer not exceeding $R\mu/c$. Note that this threshold pertains to the marginal customer who decides to balk. We will refer to this as the first-best from the customer's perspective, as this maximizes the utility for the individual (self interested) customer. As shown in Naor (1969), this threshold which is based on self-optimization (to use the terminology of Naor (1969)), does not maximize the overall expected utility of the customer population. For each $q \geq 0$ we let $y_{FI}(q)$ be the probability of a customer joining the system when there are q customers in the system and customers have full state information.

Full control. From the service provider's point of view, deciding on a threshold level amounts to fixing the waiting space in a M/M/1/k queueing system. For each value of k, the expected number of customers joining the queue per unit of time equals $\lambda(1 - \rho^k)/(1 - \rho^{k+1})$ where $\rho = \lambda/\mu$. Let $\widehat{q}$ denote the optimal waiting space. Thus, $\widehat{q}$ solves the following *full-control optimization problem*:

$$\widehat{q} = \arg\max_{k \in \mathbb{Z}_+} \lambda \frac{1 - \rho^k}{1 - \rho^{k+1}} \left[v - \mathbb{E}[h(W_k)]\right], \tag{5}$$

where W_k is the steady-state sojourn time of the customers who join the $M/M/1/k$ queue. The following proposition, which relies on Knudsen (1972) shows that such a maximizer exists, and characterizes properties of the objective function in (5).

Proposition 3.3 *The function $\Pi(\cdot)$ defined by*

$$\Pi(k) := \lambda \frac{1 - \rho^k}{1 - \rho^{k+1}} \left[v - \mathbb{E}[h(W_k)]\right], \quad k = 1, 2, \dots$$

is unimodal in k, i.e., there exists $k^* \in \{1, 2, \dots, \infty\}$ such that the function $\Pi(k)$ is strictly increasing for $k < k^*$ and strictly decreasing for $k \geq k^*$.

Based on the optimal threshold $\widehat{q}$ we define $y_{FC}(q)$ for each $q \geq 0$ to be the probability of a customer being admitted to the system when there are q customers in the system and the firm has full control.

Our equilibrium analysis will be based on the level of misalignment between the firm and the customers. To quantify this intuitive concept of misalignment, we make the following definition:

Definition 3.5 *We define ϕ as the* misalignment *between the firm and the customers as follows:*

$$\phi = \sum_{j=0}^{\infty} (y_{FI}(j) - y_{FC}(j)).$$

For our game, the misalignment ϕ equals the difference in thresholds q^* and $\widehat{q}$, i.e., $\phi = q^* - \widehat{q}$. This misalignment is related to the bias in Crawford and Sobel (1982). Using this misalignment measure ϕ, which is based on unilateral optimization under full state information to the customers and the full control of the service provider respectively, we can identify three regions. Each of these regions results in a different conflict of interest, and thus in different equilibria and outcomes for the customers and the firm. We will

initially outline the key equilibrium in each of the three regions and explain the intuition behind them. We provide a formal statement in Proposition 3.4. The three cases are as follows.

I. Complete alignment: $\phi = q^* - \widehat{q} = 0$ **(zero misalignment).** In this region, the interests of the two parties are completely aligned, and thus a pure strategy MPBNE is as follows: The firm gives two signals: i) the first for low congestion, which can be denoted as "Low." This signal is announced if the queue length is below q^*. ii) A second signal denoted by "High," which indicates high congestion, and is given when the queue length exceeds q^*. Thus we have $g(q) =$ "Low" if $q < q^*$ and $g(q) =$ "High" if $q \geq q^*$; the customer joins the queue when he/she receives the signal "Low" and balks otherwise, i.e., $y(\text{"Low"}) =$ "join", $y(\text{"High"}) =$ "balk".

As stated before, this is the key equilibrium in this region; however, this need not be the unique pure strategy MPBNE. As we will show in Section 4, there are multiple equilibria in this model. There, we will, however, be able to show that even as the equilibria become progressively more informative, it is outcome- and dynamics- equivalent to the one described above.

II. Overly patient customers: $\phi = q^* - \widehat{q} > 0$ **(strictly positive misalignment).** In this region, if customers are endowed with full state information, they would like to join the system even when the service provider would like them to balk (if she had full control). Thus, we use the term "overly patient" to emphasize the fact that, in this case, customers are willing to join a more congested system than what the firm would like. Specifically, when the queue length is between $\widehat{q}$ and q^*, the customers would like to join whereas the firm would like them to balk.

We will show that there is no threshold which is immune to defection by both the customers and the firm and consequently that there is no pure strategy MPBNE with influential cheap talk. Indeed, for pure strategy MPBNE to exist the firm should be able to signal "High" and customers who receive "High" should balk. The only threshold immune to profitable deviation by the firm is $\widehat{q}$. Given that under any pure strategy MPBNE, the customers respond to "High" by balking, a profitable deviation for the firm from any other candidate threshold is to announce "High" when the state of the system is equal to and greater than $\widehat{q}$. The customers, however, know that $\widehat{q} < q^*$ so that $\widehat{q}$ cannot induce an equilibrium: an arriving customer that receives the signal that instructs him to "balk," can deviate from the prescribed equilibrium strategy by joining; the customer will then earn positive utility (since the *only* state in which he can receive such a

signal is on the threshold itself, which is, by assumption, below q^*), and thus *detect* that such a deviation is profitable — hence ruling out the possibility of a pure strategy MPBNE with influential cheap talk.

III. Impatient customers: $\phi = q^* - \widehat{q} < 0$ **(strictly negative misalignment).** In this region, the service provider would like the customers to join a more congested system than the one they wish to join. Specifically, when the queue length is between q^* and $\widehat{q}$, the firm would like the customers to join, whereas the customers would like to balk. In order to study this region, we define $F(q)$ to be the customer's expected utility if he finds q customers in the system upon arrival and decides to join the queue; i.e, $F(q) := R - c(q+1)/\mu$. We define for $\ell < k$,

$$G(\ell, k) = \sum_{q=\ell}^{k-1} p_q^k F(q), \tag{6}$$

where $p_q^k := (\rho^q(1-\rho))/(1-\rho^{k+1})$ is the steady state probability of q customers in an $M/M/1/k$ queue. Here, $G(l, k)$ is interpreted as the average utility of a customer joining the $M/M/1$ queue given that the number of customers in the system is between l and k. Then, we have two subcases to consider:

a) $G(0, \widehat{q}) \geq 0$: if the firm announces "Low" when the queue length is below $\widehat{q}$ and "High" otherwise, the customer would like to join when they get the "Low" signal, as their expected utility is positive (since $G(0, \widehat{q}) > 0$). Further, since in equilibrium "High" would be announced only when the queue exactly equals $\widehat{q}$, the customer would balk as they know that $q^* < \widehat{q}$. Alternatively, the customer can experience the negative utility and thus will not follow the prescription of the firm. This is optimal for the firm and also describes our pure strategy MPBNE for this setting. Thus, the firm is capable of achieving its first best profits and operates as if it has full control over the customers' decisions.

b) $G(0, \widehat{q}) < 0$: In this case there is no pure strategy MPBNE with influential cheap talk. For pure strategies equilibrium with influential cheap talk to exist the firm should be able to signal "Low" and customers who receive "Low" should join. As in case II, the only threshold immune to profitable deviation of the firm is $\widehat{q}$. However, the customers know that $\widehat{q} > q^*$, thus threshold $\widehat{q}$ cannot constitute an equilibrium; an arriving customer that receives a signal that instructs him to "join" would receive negative expected utility and thus

can deviate from the prescribed equilibrium strategy by balking and obtaining zero utility. This rules out the possibility of a threshold-induced pure strategy MPBNE.

The intuition is simple: if the expected utility of the customers under an $M/M/1/\widehat{q}$ system, as given by $G(0, \widehat{q})$, is positive, they will have no incentive to deviate. Any deviation here will lead to zero utility for the customers. If, on the other hand, their utility is negative, they would be better off by not joining at all. Consequently, the threshold $\widehat{q}$ cannot induce a pure strategy MPBNE. Further, no other threshold is immune to profitable deviation on the firm's part. Thus, in case III(b) there does not exist a pure strategy MPBNE with influential cheap talk. We emphasize, however, that in case III(a) the customers can be lured to join the system even in states in which they obtain negative expected utility as long as their utility averaged over all states in which they join is positive.

We turn now to the formal statement of the equilibria we have discussed thus far. To this end, we let Π_{FI} and Π_{FC} be the firm's profit under full state information and full control respectively. Let U_{FI} and U_{FC} denote the expected utility of the customers under full state information and full control, respectively. As discussed before, Π_{FC} is the first-best profit for the firm and U_{FI} is the first-best utility for the customer. The next proposition summarizes the above observations and also compares the firm's profit and expected customer utility under the different equilibria.

Proposition 3.4

I. If $q^* = \widehat{q}$, then q^* induces a pure strategy MPBNE. Under this equilibrium the firm's profit equals Π_{FC} and the expected utility of the customers is U_{FI}.

II. If $q^* > \widehat{q}$, there is no finite q that induces a pure strategy MPBNE.

III. If $q^* < \widehat{q}$, then there are two cases:

(a) If $G(0, \widehat{q}) \geq 0$, $\widehat{q}$ induces a pure strategy MPBNE. Under this equilibrium the firm's profit equals Π_{FC} and the expected utility of the customers is U_{FC}.

(b) If $G(0, \widehat{q}) < 0$, there is no finite q that induces a pure strategy MPBNE.

First, it is important to note that all pure strategy equilibria that have influential cheap talk are neologism-proof. Neologism proof merely means that given an equilibrium, the firm (i.e., the sender of the information)

does not have incentives to use words that are not used in equilibrium but may have a "focal meaning" for the customer (i.e., the receiver) and are profit enhancing. Having said that, it is important to recall that there are multiple pure strategy equilibria in the basic game, which can be generated by relabeling the messages the firm provides. Yet, when pure strategy equilibria exist, all equilibria result in the same profit for the firm, the same average utility for all customers and the same system dynamics. That is, all equilibria are DOE (Dynamics and Outcome Equivalent). The main implication of showing that all equilibria are DOE is that there is a unique outcome for the firm and the customers.

Social Optimization. By setting $h(w) = cw$ and $v = R$, (where c is the disutility experienced by customers due to waiting and R is the value obtained by the customer from service) the system manager's problem amounts to maximizing the social welfare in the system. It follows from Naor (1969) that $\widehat{q} \leq q^*$. In this case, the solution of the first-best and full control problem lies in Region I or in Region II. Thus, either the customers are fully aligned and the social planner can announce the true state of the system or there is no pure strategy equilibrium with influential cheap talk. The firms' myopic optimization (based on each customer separately) may seem to be aligned with the customer's utility maximizing problem. However, due to the externalities and the fact that $\widehat{q} \leq q^*$, influential communication between a social planner and customers is impossible in pure strategies.

To summarize our findings so far: we have identified three regions, each with a different equilibrium behavior. We observed that a pure strategy MPBNE with influential cheap talk exists only if the firm's and the customers' incentives are perfectly aligned or if the customers are mildly impatient. Further, we prove that any MPBNE with influential cheap talk is DOE with these equilibria that have only a two-signal language.

Next, we would like to understand if an MPBNE exists with non-influential cheap talk. We shall study the following three types of equilibria: First, in Section 3.3 we show the existence of a *babbling* equilibria, where the customers disregard any information that the firm provides and hence render the cheap talk noninfluential. Next, in Section EC.6, we extend the definition of MPBNE to allow customers to randomize their actions. We characterize the *mixed* strategy MPBNE with *non-influential* cheap talk as well as the one with *influential* cheap talk. Lastly, we examine whether or not firms can provide more detailed information

regarding the state of the system. As we have already seen in this section, such revelation cannot change the outcome or the dynamics in equilibrium (if such revelation leads to an MPBNE with influential cheap talk). To this end, in Section 4 we will define a notion of *most informative* equilibria, that characterizes an equilibria where the information provided to the customer cannot be "refined" by the firm in a credible manner.

3.3. Babbling equilibrium

The equilibria constructed above are based on a signaling rule with two signals. In practice, however, there are many service providers that share no information whatsoever with the customer, whether it is direct information or one that is implicit in the type of recorded music heard while waiting. Are these systems, where no meaningful information is transmitted, in equilibrium? Furthermore, is it possible to have an equilibrium in which the firm does provide information, but due to the lack of its credibility, customers do not follow the firm's recommendation? It turns out that such an equilibrium may indeed exist in our setting. When it does exist, it is referred to as a "babbling" equilibrium (see Farrell and Rabin (1996)) to denote that the information transmitted is uncorrelated with the state of the system, and any information provided is treated by the customers as meaningless. We first provide a formal definition of babbling equilibrium. In the following let $Q_{y,g}$ be a random variable with the steady state distribution of queue length under an MPBNE (y,g)

Definition 3.6   *We say that a pure strategy MPBNE equilibrium (y,g) is a babbling equilibrium if the random variables $g(Q_{y,g})$ and $Q_{y,g}$ are independent and $y(m_i) = y(m_j)$ for all $m_i, m_j \in \mathcal{M}$.*

Definition 3.6 states that the firm provides signals that are uncorrelated with the state, and that the customers are disregarding any information which is provided by the firm, either because the messages provide no information (i.e they are identical) or because the firm lacks credibility. In the setting of Crawford and Sobel (1982), such an equilibrium always exists and it is sometimes the only one. In contrast, in our model, such an equilibrium exists in pure strategies only under the conditions that are stated in Proposition 3.5 below.[7] Using the definition of pure strategy MPBNE, it is clear that for an equilibrium to be a pure strategy

---

[7] Seidmann (1992) provides an example of a setting in which a babbling equilibrium does not exist.

babbling equilibrium it must be the case that for each $m_j \in \mathcal{M}$, $y(m_j)$ must either be 0 or 1. However, it is easy to see that $y(m_j) = 0$ for all $m_j \in \mathcal{M}$ cannot form an equilibrium. Thus, in our model, a "babbling equilibrium" exists in pure strategies if, in the absence of information, all customers join. If indeed all customers join, the resulting queueing system is an $M/M/1$ queue (i.e., with infinite waiting space), in which case the average steady state waiting time is $E[W] = 1/(\mu - \lambda)$ and customers join if and only if $R \geq cE[W]$, i.e, if $R \geq c/(\mu - \lambda)$. To provide rigorous characterization we have the following result.

Proposition 3.5   *There exists a pure strategy babbling equilibrium if and only if $R \geq \frac{c}{\mu - \lambda}$. Further, if $q^* < \widehat{q}$ and $G(0, \widehat{q}) < 0$, i.e., Case III(b) of Proposition 3.4, there does not exist a pure strategy babbling equilibrium.*

Whenever a babbling equilibrium exists, the firm's profit under such an equilibrium is denoted by Π_{NI}, and the customer's utility under such an equilibrium is denoted by U_{NI}. (NI stands for non-influential.) The following proposition shows that even though a babbling equilibrium may exist, the firm's profit under this equilibrium is dominated by the firm's profit under the two-signal equilibria described in Section 3.2. Further, the customer's expected utility is lower under the babbling equilibrium as compared to his utility under the two-signal equilibrium.

Proposition 3.6   *Assume that both a pure strategy babbling equilibrium and an equilibrium with influential cheap talk exist, then $U_{NI} < U_I$ and $\Pi_{NI} \leq \Pi_I$.*

   Proposition 3.6 underscores the value of communication. Even though an equilibrium with non-influential cheap talk does exist, both the service provider and the customers are better off when moving from a babbling equilibrium to an equilibrium with influential cheap talk (if such equilibria exist). This communication does not necessarily maximize the customer's overall expected utility but it does improve it. The rationale behind Proposition 3.6 is as follows: Naor (1969) shows that when customers are self interested and can observe the length of the queue prior to joining, their optimal threshold, q^*, is higher than what the social optimum prescribes but *it is finite*. In our setting, we observe the threshold queue length for the two-signal equilibrium is at least as high as q^*. Further, for the babbling equilibrium, when it exists, the threshold is

infinite. Thus, using information improves the customer's overall expected utility when compared to settings where the service provider is giving no information. Note that this improvement in utility and profit is achieved even though the information is unverifiable.

At this point, we remind the reader that in the region where the customers are overly impatient (region III(b)), there is no pure strategy MPBNE neither influential nor babbling. Without expanding the strategy set for the customers or the firm, it is unclear how the system would behave in this parameter regime. To alleviate this unresolved matter, we next explore the possibility of a MPBNE in which the customers are allowed to randomize.

3.4. Does randomization enable communication?

In this section we summarize the results of Appendix EC.6 where we relax the assumption that the customer, given an announcement, either joins with probability 1 or balks with probability 1, i.e. we allow the customers to randomize. We discuss the impact of allowing the firm to randomize in Appendix EC.5.

We show that in this setting it is never optimal for the firm to use the same signal for two states but use some other signal for a state in between. Thus, a firm will never provide the same signal on two disjoint intervals, and provide a different signal in between.

We have shown in Section 3.3 that babbling equilibria need not exist in pure-strategies. When customers are allowed to randomize, however, such equilibria always exist. We characterize the details of such an equilibria in Appendix EC.6.1.

In Section EC.6.2 we show that in addition to the babbling equilibria, there may exist more *informative* MPBNE in mixed strategies. We start by showing that it suffices to consider equilibria with certain characteristics. Specifically, in the following proposition we show that, if customers randomize, they will do so only within an interval. Furthermore, we show that balking with probability 1 in some states and joining with probability 1 in other states cannot co-exist. Thus, there are only two possible types of two-signal mixed strategy MPBNE in which randomization is used. The two types can be described as follows: The firm announces "High" and "Low" based on the threshold q_{mix}, the customers then react as follows: (a) in the first type of MPBNE which we shall refer to as *Join or Randomize* equilibria, the customers who receive

"Low" join the system and the customers who receive "High" join the system with probability $\theta \in (0, 1)$; (b) in the second type of MPBNE which we shall refer to as *Randomize or Balk* equilibria, the customers who receive "Low" join the system with probability θ, and the customers who receive "High" balk. Both of these equilibria are completely defined by two parameters: the threshold q_{mix} used by the firm for signaling and the randomization parameter θ. We provide a characterization of the pair (q_{mix}, θ) which induces the mixed strategy MPBNE in Appendix EC.6.3.

So far, we characterized and proved the existence (or non-existence) of two-signal communication between the firm and its customers that form a MPBNE. We next turn to the question regarding the existence of additional equilibria.

4. Most Informative Equilibria

In practice, firms frequently use announcements regarding the congestion in the system or the volume of calls that are equivalent to the two-signal equilibrium. In this section we show that other types of equilibria, in which more information is provided to the customers are also possible. While these equilibria may be different in terms of the firm's announcements, they are be equivalent in terms of the customers' actions, their utility, and the firm's profits.

To formally examine this question we introduce the notion of *most informative equilibria*.

Definition 4.1   *Consider an ordered set $\mathbb{S} = \{a_0 = 0, a_1, a_2, a_3, ..., a_J = \infty\}$, where $a_i \in \mathbb{Z}_+$[8]. Consider the signaling rule defined as follows: $g(q) = m_j$ for all $q \in [a_{j-1}, a_j)$, for all $1 \leq j \leq J$. We say that the set $\mathbb{S}$ induces a mixed strategy MPBNE if there exists an action rule $y(n) \in [0, 1]$ for all $n \in \mathbb{N}$, such that $(y(\cdot), g(\cdot))$ forms a mixed strategy MPBNE. Further, if there does not exist an action rule $y(\cdot)$, such that $(y(\cdot), g(\cdot))$ form an MPBNE, we say that the ordered set does not induce a mixed strategy MPBNE.*

Definition 4.2   *We say that an ordered set $\mathbb{S}$ induces the most informative pure strategy MPBNE, if the following holds: (a) $\mathbb{S}$ induces a mixed strategy MPBNE in the sense of Definition 4.1, and (b) every ordered set $\mathbb{S}' \supsetneq \mathbb{S}$ does not induce a mixed strategy MPBNE.*

---

[8] Note that J could possibly be infinite.

For the two types of mixed equilibria defined in the Section EC.6.2, we put:

$$q_{mix}^{JR} = \max\{q : \text{there exists } \theta \text{ such that } (q, \theta) \text{ induces a Join/Randomize type mixed strategy MPBNE.}\}$$

$$q_{mix}^{RB} = \min\{q : \text{there exists } \theta \text{ such that } (q, \theta) \text{ induces a Randomize/Balk type mixed strategy MPBNE.}\}$$

The following result describes the *most informative* pure strategy MPBNE.

Proposition 4.1

***I.** If $q^* = \widehat{q}$, the set $\mathbb{Z}_+$ induces the most informative MPBNE. In this equilibrium the firm's profit equals Π_{FC}, and, the expected utility of the customers is U_{FI}.*

***II.** If $q^* > \widehat{q}$, we have the following cases: (a) If there exists a mixed strategy MPBNE of type Join/Randomize with parameters (q_{mix}, θ), then the set $\{0, 1, \ldots, q_{mix}^{JR}, \infty\}$ induces the most informative equilibrium, and (b) If there exists a mixed strategy MPBNE of type Randomize/Balk with parameters (q_{mix}, θ) then the set $\{0, q_{mix}^{RB}, q_{mix}^{RB} + 1, q_{mix}^{RB} + 2 \ldots, \infty\}$ induces the most informative equilibria. (c) If there does not exist any informative mixed strategy MPBNE then the most informative equilibria is the babbling equilibria.*

***III.** If $q^* < \widehat{q}$, either: (a) $G(0, \widehat{q}) > 0$, in which case the set $\{0, 1, 2, \ldots, \overline{q}, \widehat{q}, \widehat{q} + 1, \ldots, \infty\}$ induces the most informative MPBNE where $\overline{q} = \arg\max\{q : G(q, \widehat{q}) > 0\}$. Under this equilibrium the firm's profit equals Π_{FC}. Further, the expected utility of the customers is U_{FC}. (b) If $G(0, \widehat{q}) \leq 0$, and the mixed strategy MPBNE of type Randomize/Balk has parameters (q_{mix}, θ) then the set $\{0, q_{mix}^{RB}, q_{mix}^{RB} + 1, q_{mix}^{RB} + 2, \ldots, \infty\}$ induces the most informative equilibria.*

One consequence of Proposition 4.1 is that if the firm is not perfectly aligned with the customers, it will always maintain some level of *intentional vagueness*. Recall, that by intentional vagueness we mean that under the most informative equilibrium, the firm provides the same signal on multiple states of the system. This occurs, for example, in case III(a) where the firm is intentionally vague when the state of the system is between $\overline{q}$ and $\widehat{q}$ so that customers are lured to the system. If the firm provides any further information in these states, the firm would forgo some of its profits and this cannot be sustained in equilibrium. It is important to note that, even though the most informative equilibria could have countably infinite signals

(under Randomize/Balk and full alignment), only a finite subset of these are transmitted with strictly positive probability.

In discussing the opportunities for the service provider to lure a customer to take an action against his best interest we need to distinguish between two types of strategies: (i) misrepresentation of the state of the system, and (ii) intentional vagueness. The latter was shown in Proposition 4.1 to be not only feasible, but an actual equilibrium that may arise under certain conditions. The former, however, is not possible. There are several reasons for this impossibility. These are based on the fact that customers are strategic and thus can detect such a lie, in the same way that they can detect a profitable deviation from a strategy profile.

Summary of the Results and Implications. The following figure summarizes our findings thus far. Note that the if the customer and firm have identical impatience the firm can reveal complete information regarding the system state, this corresponds to the 45 degree line shown in Figure 2. Further, if the customers are more patient than the firm then any equilibrium that exists must entail randomization by the customers, and thus the firm cannot achieve its first-best profits. However, if the customers are less patient than the firm, then the firm achieves its first-best profit by using vague announcements unless the difference between the patience is not too large. This is depicted in the region above the 45 degree line in the figure. Thus, there always exists intentional vagueness in the announcement unless the firm and the customers are perfectly aligned.

Announcements and Managing Customer Expectations. The announcements in our model are used to build customer expectations about the congestion in the system. In this sense the announcement plays a role of managing customer expectations regarding the system congestion. It is interesting to note that the message provided to the customer might not exhibit this information literally but is captured by the bayesian rule employed by the customer (Condition 1 in Definition 3.1 of MPBNE). In this section, we explore the extent to which the firm can reveal the information, i.e., how precisely can we tell the customers "what to expect?" while sustaining the equilibrium. The above results indicate that in our setting, since the incentives of the parties are not perfectly aligned, the firm almost always needs to use intentional vagueness when managing expectations regarding the system congestion. In pure strategies, the firm uses intentional vagueness to lure customers into the system, in order to make sure the firm itself does not have incentive to deviate
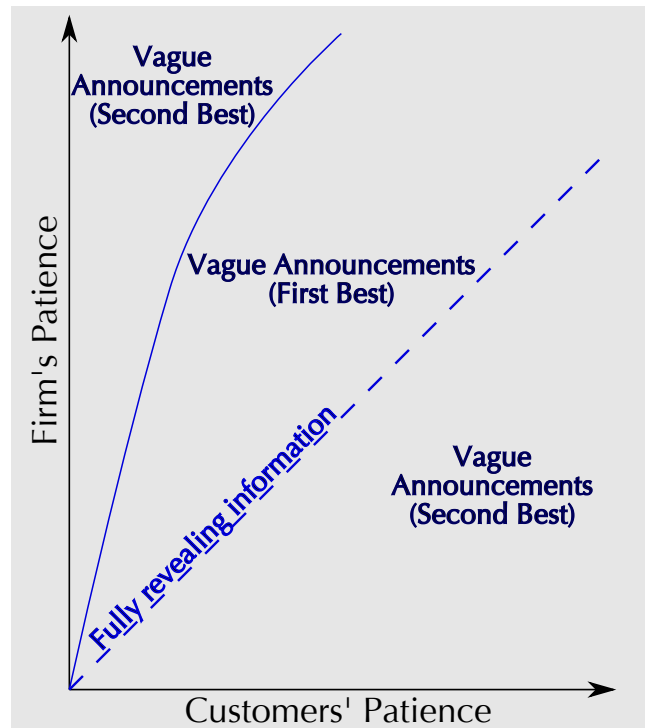
Figure 1 The three regions as defined in Proposition 3.4, based on the customers' patience and the firm's patience.

from this announcement. When allowing for mixed strategies, the firm is using intentional vagueness, even though the firm's profit is below its first-best, in order to sustain an equilibrium - this time making sure that the incentives for the customers are aligned.

5. Conclusions

We study a model of a service system where customers are not only strategic in their actions but also in the way they interpret information. The service provider on her end is strategic in the way she provides information. We show that even though the information is costless and unverifiable, it can improve the outcome for all players. Through a stylized model, we are able to provide a theoretical basis for intentional vagueness on the part of the service provider.[9]

The framework used in the paper echoes the cheap-talk model proposed in Crawford and Sobel (1982). Driven by the application to services, we proposed a model that has two key features: First, the model

---

[9] It is important to note that different firms may have different reasons for using intentional vagueness. For example, real service systems are more complex than simple M/M/1 queues (both operationally and in terms of information availability to both parties) which implies that the service provider may be unsure about the mapping between the number of customers in the system and the resulting waiting that a customer would experience. This uncertainty may be a driver for the intentional vagueness observed in practice.

considers multiple receivers (customers) whose actions have externalities on other receivers. Second, the stochasticity of the type (i.e., state of the system) is not exogenously given but is determined by the equilibrium strategies of the players. We show that these have non-trivial impacts on the structure of equilibria. The framework developed in this paper can be applied to other operations management settings where the customers cannot credibly verify the information provided to them. Allon and Bassamboo (2008) studies a setting where a retailer signals strategic information to the potential buyers.

Throughout this paper, we assumed that the queue is Markovian and that there is a single server. The results in the paper are extendable to the case of multiple servers. Another assumption which can be relaxed is the assumption that $\lambda < \mu$ (i.e., demand is less than capacity). The structure provided in the paper also holds for $\lambda > \mu$. The model in the paper can be suitably extended to allow for multiple customer classes. The MPBNE that arises from this model would have a richer equilibrium language. However, it can be shown that under mild condition, the solution would exhibit a similar threshold-type structure. (See Allon and Bassamboo (2009b) for analysis of such models). In this paper, we assumed that announcements are provided immediately upon the customer's arrival. Allon and Bassamboo (2009a) shows that delaying the information providing may assist the firms in improving its credibility for some parameters, yet it may hurt its credibility as well.

Future models should also account for the fact that customers may not be expected-utility maximizers. Under such modeling assumptions customers may prefer receiving information over no information, and in particular, customers may prefer more accurate announcements. Further, in many settings one may expect that the value obtained from the service and the cost of waiting to be affected by the announcements made by the firm.

References

Afeche, P. 2004. Incentive-compatible revenue management in queueing systems: Optimal strategic idleness and other delaying tactics Working paper, Kellogg School of Management, Northwestern University.

Allon, G., A. Bassamboo. 2008. Buying from the babbling newsvendor: Availability information and cheap talk. *Working paper, Kellogg School of Management, Northwestern University* .

Allon, G., A. Bassamboo. 2009a. The impact of delaying the delay announcements. *Working paper, Kellogg School of Management, Northwestern University* .

Allon, G., A. Bassamboo. 2009b. Managing customer expectations and priorities in service systems. *Working paper, Kellogg School of Management, Northwestern University* .

Allon, G., A. Federgruen. 2007. Competition in service industries. *Operations Research* **55**(1) 37–55.

Armony, M., C. Maglaras. 2004a. Contact center with a call-back option and real-time delay information. *Operations Research* **52**(4) 527–545.

Armony, M., C. Maglaras. 2004b. On customer contact centers with a call-back options: Customer decisions, routing rules and system design. *Operations Research* **52**(4) 271–292.

Armony, M., N. Shimkin, W. Whitt. 2007. The impact of delay announcements in many-server queues with abandonment. *Working Paper, Stern School of Business, NYU, NY* .

Austen-Smith, D. 1990. Information transmission in debate. *American Journal of Political Science* **34**(1) 124–152.

Chen, Y., N. Kartik, J. Sobel. 2008. Selecting cheap-talk equilibria. *Econometrica* **76**(1) 117–136.

Crawford, V. P., J. Sobel. 1982. Strategic information transmission. *Econometrica* **50**(6) 1431–1451.

Dobson, G., J. Pinker. 2006. The value of sharing lead time information. *IIE Transactions* **38**(3) 171–183.

Farrell, J., M. Rabin. 1996. Cheap talk. *The Journal of Economic Perspectives* **10**(3) 103–118.

Guo, P., P. Zipkin. 2007. Analysis and comparison of queues with deifferent levels of delay information. *Management Science* **53**(6) 962–970.

Hassin, R. 1986. Consumer information in markets with random product quality: The case of queues and balking. *Econometrica* **54**(5) 1185–1195.

Hassin, R., M. Haviv. 2003. *To Queue or Not to Queue: Equilibrium Behavior in Queueing Systems*. Kluwer Academic Publishers, Boston, MA.

Haviv, M., M.L. Puterman. 1998. Bias optimality in controlled queueing systems. *Journal of Applied Probability* **35**(1) 136–150.

Haviv, M., Y. Ritov. 2001. Homogeneous customers renege from invisible queues at random times under deteriorating waiting conditions. *Queueing Systems* **38**(4) 495–508.

Jouini, O., Y. Dallery, O.Z. Aksin. 2007. Queueing models for multiclass call centers with real-time anticipated delays. *Working Paper* .

Knudsen, N. C. 1972. Individual and social optimization in a multi-server queue with general cost-benefit structure. *Econometrica* **40**(3) 515–528.

Mandelbaum, A., N. Shimkin. 2000. A model for rational abandonments from invisible queues. *Queueing Systems* **36**(1) 141–173.

Naor, P. 1969. The regulation of queue size by levying tolls. *Econometrica* **37**(1) 15–24.

Ross, S.M., J.G. Shanthikumar, Z. Zhu. 2005. On increasing-failure-rate random variables. *Journal of Applied Probability* **42**(3) 797.

Seidmann, D. 1992. Cheap talk games may have unique, informative equilibrium outcomes. *Games and Economic Behavior* **4**(3) 422–425.

Stidham, S., R.R. Weber. 1989. Monotonic and insensitive optimal policies for control of queues with undiscounted costs. *Operations Research* **37**(4) 611–625.

E-Companion to *"We will be right with you": Managing customer expectations with vague promises and cheap talk*

The Appendices are organized as follows. Appendix EC.1 compares our findings with the classical cheap talk model. Appendinx EC.2 studies a model where we relax the assumption that the arrival rate is known to the customers Appendix EC.3.1 describes a model where a Tâtonnement scheme leads to the equilibrium described in the main paper. Appendix EC.3.2 presents results of simulation study for a model where the firm chooses the capacity level optimally. Appendix EC.4 discusses the setting where the customers can abandon the system after joining. Appendix EC.5 discusses enriching the signaling rules by considering strategies where the firm is allowed to randomize. Appendix EC.6.3 provides the characterization for various types of mixed strategy MPBNE. Lastly, Appendices EC.7, EC.8 and EC.9 contains the proofs of the results in the main paper, the e-companion and the supporting lemmas, respectively.

EC.1. Comparison of our results with classical cheap talk model

In this section, we contrast the key conclusions of our cheap talk model with the classical Crawford and Sobel (1982) model. To do this effectively, we summarize the classical cheap talk game and its results. The static game in Crawford and Sobel is played between a sender who has some private information and a receiver who takes the payoffs-relevant actions. Both the sender and the receiver obtain utilities which depend on: (a) the action taken by the receiver, y; (b) the state/type Q. We denote the utility of the sender by $\pi(y, Q) = -(y - (Q - b))^2$, and the receiver's by $U(y, Q) = -(y - Q)^2$. In this setting, the receiver would take the action Q if the receiver has full state information; and if the sender has full control, his preferred action is $Q - b$. Thus, using our definition, the misalignment in this setting equals b.

Comparison of key results. The following are the key differences between the outcome of our game and the classical cheap talk game:

1. In our model, Proposition 4.1 shows that it is never advisable for the firm to provide full state information, even if it could have committed to fully reveal the state of the system, unless the firm and the customers are fully aligned (i.e., if the misalignment is zero). Hence, there are two reasons for the firm to shade information: first, to lure customers to join in states in which they would otherwise balk had they known the real

state of the system (as long as, of course, the system is not too congested) and, second, in order to sustain an

equilibrium with influential cheap talk, for example by inducing a mixed strategies equilibrium. While the

latter is analogous to the benefit of shading information in Crawford and Sobel (1982), the former is novel

to our operational setting. From the above, it may seem that there is a tension between the firm and the

customer, i.e., when the firm tries to improve its profit by moving from a non-informative to an informative

equilibrium by luring the customers, the customers may lower their overall utility. However, we find that

there are situations where the firm can create influential language by shading information and also improve

not only itself but also the customer's utility.

2. In our model, the existence (or non-existence) of the equilibrium depends not only on the magnitude

of the misalignment, but also on the direction. While the existence of an equilibrium in the classical cheap

talk model depends only on the magnitude $|b|$. In our case, the sign (and not just the magnitude) of the mis-

alignment $q^* - \widehat{q}$ is crucial in determining whether an equilibrium exists or not. This is due to the different

levels of detectability associated with the different actions - by joining the system when "recommended" to

balk, the customer can potentially detect the true state of the system; however, by balking when "asked" to

join, the customer does not gain any information.

EC.2. Relaxing the Information Structure

The paper thus far assumes that customers have perfect information about the parameters of the service

system. In this section, we relax this assumption and test the robustness of our results. We consider a model

in which the firm has perfect information of the arrival rate, yet customers know only the distribution of

the arrival rate. For simplicity, we assume that there are only two possible states-of-the-world, ω_1 and ω_2 so

that the arrival rate is λ_{ω_1} w.p. p_{ω_1} or λ_{ω_2} w.p. $p_{\omega_2} = 1 - p_{\omega_1}$. Further, we define the set $\Omega = \{\omega_1, \omega_2\}$. This

models a service provider whose call volume varies from day to day. Most call centers can distinguish very

quickly between the two states. However, customers calling on any specific day do not know the exact state

of the world.

We begin by defining the notion of MPBNE for this setting. To this end, we denote the firm's signaling

rule by $g : \mathbb{Z} \times \Omega \mapsto \mathcal{M}$, where $g(q, \omega)$ is the message that the firm announces when the number of customers

in the system is q and that the state of the world is ω, where $\omega \in \Omega$. Thus, the firm's signal depends on both the state-of-the-system q and the state-of-the-world $\omega \in \Omega$. As before, y denotes the customer's strategy which maps the message to the probability of joining the system. Let $p_q(g, y)$ be the steady-state probability that the number of customers in the system is q when the firm uses the signaling rule g and the customers use the strategy y. (Note that $p_q(g, y)$ depends on p_{ω_1} and p_{ω_2}.)

Definition EC.2.1 *(**Markov Perfect Bayesian Nash Equilibrium** MPBNE) We say that the signaling rule $g(\cdot, \cdot)$ and the action rule $y(\cdot)$ constitute a Markov Perfect Bayesian Nash Equilibrium (MPBNE), if they satisfy the following conditions:*

1. *For each $m \in \mathcal{M}$,*

$$y(m) = \begin{cases} 1 & \frac{\sum_{\{q:g(q,\omega)=m, \omega \in \Omega\}} \left[R - c \frac{q+1}{\mu} \right] p_q(y,g)}{\sum_{\{q:g(q,\omega)=m, \omega \in \Omega\}} p_q(y,g)} \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

2. *With $f(j) = v - \mathbb{E}[h(W(j+1))]$, where $W(j+1)$ is the time spent by a customer who joins the system with j customers, there exist constants $J_0^\omega, J_1^\omega, \ldots,$ and γ^ω for $\omega \in \Omega$ that solve the following set of equations:*

$$\begin{aligned} J_0^\omega &= \max_{m \in \mathcal{M}} \left\{ \frac{f(0)y(m) - \gamma^\omega}{\lambda_\omega} + J_0^\omega(1 - y(m)) + J_1^\omega y(m) \right\} \\ &= \frac{f(0)y(g(0,\omega)) - \gamma^\omega}{\lambda_\omega} + J_0^\omega(1 - y(g(0,\omega))) + J_1^\omega y(g(0,\omega)) \\ J_q^\omega &= \max_{m \in \mathcal{M}} \left\{ \frac{f(q)y(m) - \gamma^\omega}{\lambda_\omega + \mu} + \frac{\mu}{\lambda_\omega + \mu} J_{q-1}^\omega + \frac{\lambda_\omega}{\lambda_\omega + \mu}(J_q^\omega(1 - y(m)) + J_{q+1}y(m)) \right\} \\ &= \left\{ \frac{f(q)y(g(q,\omega)) - \gamma^\omega}{\lambda_\omega + \mu} + \frac{\mu}{\lambda_\omega + \mu} J_{q-1}^\omega + \frac{\lambda_\omega}{\lambda_\omega + \mu}(J_q^\omega(1 - y(g(q,\omega))) + J_{q+1}^\omega y(g(q,\omega))) \right\} \text{(EC.1)} \end{aligned}$$

As in the previous section, we begin by proving that any equilibrium with influential cheap talk is DOE with another equilibrium that has only two-signals.

Proposition EC.2.1 *Let the pair (y, g) be a pure strategy MPBNE with influential cheap talk. Then there exists two thresholds: $\bar{q}_{\omega_1}$ and $\bar{q}_{\omega_2}$, so that the firm's signaling rule and the customers' strategy are defined as follows: for any $q \in \mathbb{Z}$ and $\omega \in \Omega$,*

$$\tilde{g}(q, \omega) = \begin{cases} m_1 & q \leq \bar{q}_\omega, \\ m_0 & \text{otherwise.} \end{cases}, \quad \tilde{y}(m) = \begin{cases} 1 & m = m_1, \\ 0 & \text{otherwise.} \end{cases} \tag{EC.2}$$

such that $(\widetilde{y}, \widetilde{g})$ *is DOE to* (y, g).

Thus, it suffices to study only two-signal equilibria. To this end, we define the following.

Definition EC.2.2   *We say that the thresholds* $(\overline{q}_{\omega_1}, \overline{q}_{\omega_2})$ *induce a pure strategy MPBNE if the pair* $(\widetilde{g}(\cdot, \cdot), \widetilde{y}(\cdot))$ *given by* (EC.2) *forms an MPBNE.*

Using Proposition 3.3 for each arrival rate λ_{ω_1} and λ_{ω_2} along with the fact that the firm has perfect knowledge about the arrival rate, we obtain that the full control solution in this case can be described in terms of two thresholds: $\hat{q}_{\omega_1}$ and $\hat{q}_{\omega_2}$. Under this full control, the firm would admit customers if the number of customers in the system is less than $\hat{q}_{\omega_i}$ when the arrival rate is λ_{ω_i} for $i = 1$ and 2. We now study the customers response when the firm tries to implement its full control solution. In this case, the firm would give a message geared to induce customers to join and a message that induces customers to balk. Thus, the firm would use the same message to signal different levels of congestion for different states-of-the-world. The customer, when presented with a message needs to figure out (using bayesian rules) the state of the system, which will require him to update his prior on the state of the world (given the message).

We will next characterize the equilibrium that arises in such a game. To that end, we begin by discussing the updating mechanism. Let $\hat{p}_\omega$ denote the conditional probability of being in state-of-the-world $\omega \in \Omega$, given that the firm uses a message that tries to induce the customers to balk. To compute this probability, we define $p_{q,\omega}^k := (\rho_\omega^q (1 - \rho_\omega))/(1 - \rho_\omega^{k+1})$ to be the steady state probability of q customers in an $M/M/1/k$ queue with traffic intensity $\rho_\omega = \lambda_\omega/\mu$ and $\omega \in \Omega$. Note that this definition is linked to the function p_q^k defined earlier. The key difference is that here we need to account for the different arrival rates. We thus have

$$\hat{p}_{\omega_1} = \frac{p_{\omega_1} p_{\hat{q}_{\omega_1}, \omega_1}^{\hat{q}_{\omega_1}}}{p_{\omega_1} p_{\hat{q}_{\omega_1}, \omega_1}^{\hat{q}_{\omega_1}} + p_{\omega_2} p_{\hat{q}_{\omega_2}, \omega_2}^{\hat{q}_{\omega_2}}}, \quad \hat{p}_{\omega_2} = 1 - \hat{p}_{\omega_1}. \tag{EC.3}$$

Combining these estimates to obtain the state-of-the-system, the customer would conclude that the average number of customers in the system is

$$\widehat{q} = p_{\omega_1} \hat{q}_{\omega_1} + p_{\omega_2} \hat{q}_{\omega_2}.$$

In the same spirit, we define $\tilde{p}_\omega$ to be the conditional probability of being in state-of-the-world $\omega \in \Omega$, given that the firms uses a message that tries to induce the customer to join. Using Bayes rule, we obtain:

$$\tilde{p}_{\omega_1} = \frac{p_{\omega_1}(1 - p_{\hat{q}_{\omega_1},\omega_1}^{\hat{q}_{\omega_1}})}{p_{\omega_1}(1 - p_{\hat{q}_{\omega_1},\omega_1}^{\hat{q}_{\omega_1}}) + p_{\omega_2}(1 - p_{\hat{q}_{\omega_2},\omega_2}^{\hat{q}_{\omega_2}})}, \tilde{p}_{\omega_2} = 1 - \tilde{p}_{\omega_1}. \tag{EC.4}$$

In order to state the formal result we also need to extend the definition of the function $G(\cdot)$, defined in equation (6). Recall that $F(q)$ is the customer's expected utility if he finds q customers in the system upon arrival and decides to join the queue; i.e, $F(q) := R - c(q+1)/\mu$. We then define for $\ell < k_\omega$,

$$G_\omega(\ell, k_\omega) = \sum_{q=\ell}^{k_\omega - 1} p_{q,\omega}^{k_\omega} F(q). \tag{EC.5}$$

The next result formally states the conditions for existence of an equilibrium with influential cheap talk. Specifically, it shows that the findings in Proposition 3.4 hold even when there is asymmetry with respect to the information about the state-of-the-world (in this case, the arrival rate). In stating the proposition, we use q^* as defined in equation (4).

Proposition EC.2.2

I. *If $q^* = \hat{q}$, then thresholds $(\hat{q}_{\omega_1}, \hat{q}_{\omega_2})$ induce a pure strategy MPBNE.*

II. *If $q^* > \hat{q}$, there is no influential pure strategy MPBNE.*

III. *If $q^* < \hat{q}$, then:*

(a) *If $\tilde{p}_{\omega_1} G_{\omega_1}(0, \hat{q}_{\omega_1}) + \tilde{p}_{\omega_2} G_{\omega_2}(0, \hat{q}_{\omega_2}) \geq 0$, then the thresholds $(\hat{q}_{\omega_1}, \hat{q}_{\omega_2})$ induce a pure strategy MPBNE.*

(b) *If $\tilde{p}_{\omega_1} G_{\omega_1}(0, \hat{q}_{\omega_1}) + \tilde{p}_{\omega_2} G_{\omega_2}(0, \hat{q}_{\omega_2}) < 0$, there is no influential pure strategy MPBNE.*

The above proposition shows that even with an additional layer of asymmetry of information, the equilibrium structure and the conditions for the existence of equilibrium remain essentially the same as in the base case. One of the key conclusions from the base model is that there is (almost) always some level of vagueness needed to ensure the existence of an equilibrium with influential cheap talk. We shall next show that the additional layer of information asymmetry only adds to the extent of vagueness in equilibrium.

Appealing to Proposition 3.4 for each state-of-the-world separately, we observe that it will be possible

for the firm to reveal the true state of the world and still be able to sustain an equilibrium with influential

cheap talk, if and only if

$$Rc/\mu < \min\{\hat{q}_{\omega_1}, \hat{q}_{\omega_2}\}, \ G_{\omega_1}(0, \widehat{q}_{\omega_1}) \geq 0 \text{ and } G_{\omega_2}(0, \widehat{q}_{\omega_2}) \geq 0. \tag{EC.6}$$

Note that if (EC.6) is satisfied then a two-signal equilbrium (where the firm does not disclose the state-

of-the-world) can also be sustained. However, one can also observe that there might be settings where an

equilibrium characterized in Proposition EC.2.2 exists but the conditions in (EC.6) are violated.

That is, when these conditions are violated, an additional level of vagueness (this time regarding the state

of the world) is needed to guarantee the existence of an equilibrium with influential cheap talk. Moreover,

there is almost always a region in which the firm cannot reveal the state-of-the-world and still sustain an

equilibrium. In other words, we observe that the firm may have incentives to suppress information about the

state of the world and to be intentionally vague about the state of the system.

The results are best explained using Figure 2. In the figure, we depict the range of values that $R/c\mu$ can

take for the firm to sustain an equilibrium. Here $\tilde{q}_{\omega_i}$ is the expected number of customers in the system

when the firm implements its full control and the state of the world is $\omega_i \in \Omega$. The figure captures three

possible configurations: one in which the customer perfectly knows the arrival rate to be λ_{ω_1} (the vertical

axis), second, in which the customer perfectly knows the arrival rate to be λ_{ω_2} (the horizontal axis), and a

third, in which the customer does not know the exact state of the world (the 45 degree line). Note that the

first two are sub-games for the setting where the arrival rate is known to the customer.

When the customer knows that the arrival rate is λ_{ω_1}, an equilibrium with influential cheap talk can be

sustained as long as $q^* \in [\tilde{q}_{\omega_1}, \hat{q}_{\omega_1}]$. Analogously, when the customer knows that the arrival rate is λ_{ω_2}, an

equilibrium with influential cheap talk can be sustained as long as $q^* \in [\tilde{q}_{\omega_2}, \hat{q}_{\omega_2}]$. The main result above

states that when the arrival rate is not known to the customers with certainty, there exists an equilibrium

with influential cheap talk, provided that $q^* \in [\tilde{q}, \hat{q}]$. As we observe in Figure 2, for the value of q^* between

the red dots, the firm can sustain an equilibrium with influential cheap talk in all three configurations. Thus,

an equilibrium with influential cheap talk in the game with imperfect information can be sustained whether

or not the firm reveals the state-of-the-world. For the values of q^* along the blue line, an equilibrium cannot be sustained if the customers have information about the state-of-the-world or the firm tries to reveal it. Thus, we obtain the result that suppression of information may be needed to sustain an equilibrium.
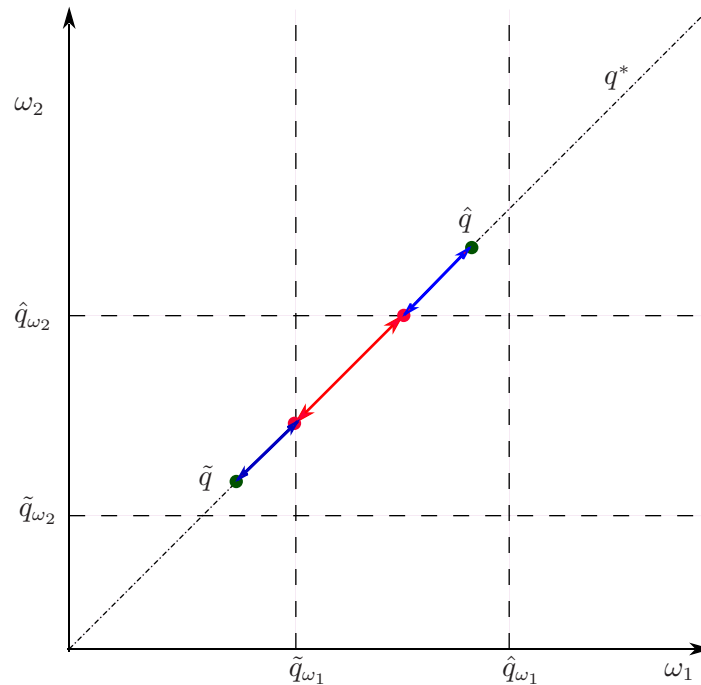

Figure EC.1   Information Asymmetry: We observe that in some settings (if the q^* lies on the blue arrows) the firm prefers suppressing the state-of-the-world information.

Discussion. In our model a customer does not need to know the threshold used by the firm or the actual arrival rate. Customers are rational and apply Bayes rule along with the steady state probability of the resulting queuing system, to obtain the "mapping" between the signal they receive from the firm and the decision they take whether to join or balk. In that sense, the customer does not need to know the threshold of the firm (i.e. the strategy). Customers only need to form beliefs (that are consistent on the equilibrium path) on the system being in a specific state, in terms of the number of people in the system, given a message and their prior belief about the arrival rate to the system.

We observe that all the findings for the pure strategies extend to this relaxed model with information asymmetry. Further, one can show that when allowing for mixed strategies, similar results are observed. Specifically, any equilibrium has a DOE equilibrium with two signals. Moreover, the structure of the mixed

strategies equilibrium can be divided into the same Join/Randomize type and Randomize/Balk type equilibrium. It is also worth noting that all these findings can easily be extended to settings where there is a countable number of states of the world.

To summarize, we note that the base model captures some level of information asymmetry, which drives the main result regarding the need for intentional vagueness. In this section we observe that the existence of additional levels of information asymmetry can only lead to increased vagueness in equilibrium.

EC.3. Numerical Study

EC.3.1. Convergence to equilibrium

In this section we conduct a numerical study to illustrate how the different pure strategy equilibria characterized in the main paper are to be reached via a Tâtonnement scheme in which customers adaptively, yet myopically, learn how to map the signal of the firm to their action so as to maximize their utilities. This study sheds light on the robustness of our findings, in particular, when customers are not necessarily capable of forming beliefs about the queueing dynamics given the messages provided by the firm, and the strategies of all customers.

We consider a multi-period setting of our cheap talk game. In each period, the customers use their past experience to decide how to respond to various messages given by the firm. Customers start with arbitrary beliefs regarding the meaning of the messages and the signaling rules that the firm uses. The firm, knowing how the customers respond to these signals in each period, uses these signals to maximize profit. Thus, in any period the strategy of the firm can be computed using an MDP (as in Condition 2 of Definition EC.2.1). The updating mechanism used by the customers is as follows: in period n the probability of joining after receiving a signal m_i is given by:

$$\alpha_n(m_i) = (1 - \alpha_{n-1}(m_i))\alpha_{n-1}(m_i) + \alpha_{n-1}(m_i)(\mathbb{I}\{U^{n-1}(m_i) > 0\}), \qquad \text{(EC.7)}$$

where $U^{n-1}(m_i)$ denotes the expected utility obtained by the customers when the firm announces m_i in period $n-1$. Also, we assume that the length of each period is sufficiently long so that the system reaches its steady-state. For the purpose of the above described updating mechanism, we shall use the expected

steady-state utilities obtained by the customers. The updating rule outlined in (EC.7) assumes that only those observing the state of the system are capable of updating their beliefs.

While other more effective ways to teach the customers may exist, we use a simple, myopic learning scheme. The goal of this numerical study is merely to illustrate how the different equilibria (or lack there of) in this paper may be reached using a simple iterative process even when the customer has no a-priori knowledge of the meaning of the messages, the equilibrium strategies, or even the properties of the queueing system.

We consider three sets of parameters, each with non-zero misalignment. These sets correspond to regions II, IIIa and IIIb of Section 3.2. In the first, we use the following parameters: The arrival rate of the customers is $\lambda = 1$ customer per unit of time. The service rate is $\mu = 2$ customers per unit of time. The firm earns a value of $v = 10$ from each customer served, and incurs an holding cost of 1 per customer per unit of time, i.e., $h(w) = w$. A customer, on the other hand, obtains a reward of $R = 5$ from service and incurs a disutility of $c = 1$ per unit of time from waiting. At, time $n = 0$, the customers begin with subjective beliefs that any signal should result in joining with probability $1/2$ and balking with probability $1/2$. At time $n = 0$, the firm begins with the following strategy: it signals m_1 whenever the queue length is below 25, and signals m_2 otherwise. (Note that under the customers' initial subjective belief, the firm is indifferent among all thresholds.) For these parameters, we compute the probabilities of joining for signals m_1 and m_2 for the customer using (EC.7) and the optimal response for the firm in periods $1, ...20$ using an appropriate MDP. The customers update their beliefs that the firm uses signal m_i when it is beneficial for them to join. (As one can observe, when the firm uses signal m_1 for the low states of the system, the customers' probability of joining based on their belief converge to 1, while for signal m_2 the probability of joining converges to zero.)

Figure EC.2 depicts the probability of joining for the signals m_1 and m_2 in each period $1, \ldots, 20$. The parameters in this case fall under region IIIa, in which a two-signal informative equilibrium exists. In this case, the iterative process converges to an informative equilibrium. In this equilibrium, when the firm announces m_1, all customers join, and when it announces m_2, all customers balk. Note that in this experiment, the literal meaning of the signals (or whether they even have a meaning) played no role in the
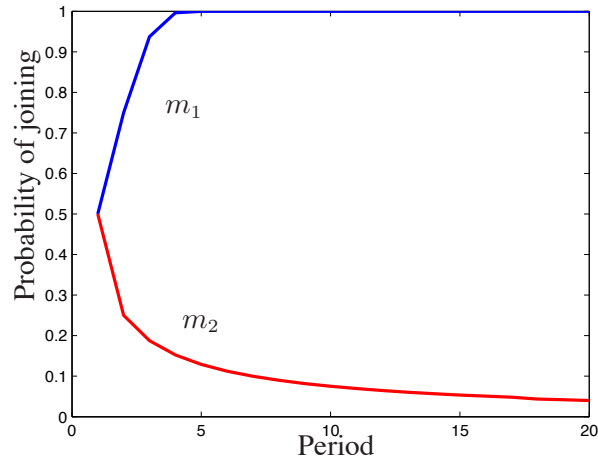
Figure EC.2   Tatonnement scheme for Region IIIa, the case of informative equilibrium. The probability of joining given the

message m_1 converges to 1, and for message m_2 converges to 0.
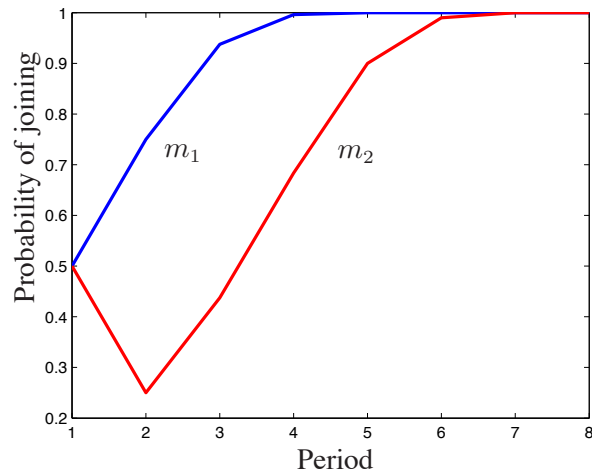
equilibrium.

Figure EC.3   Tatonnement scheme for Region II, the case of non-informative equilibrium. The probability of joining for mes-

sage m_1 and message m_2 converge to 1.

In the next example we modify only the reward obtained by the customers to $R = 10$. As described in

Section 3.2, when the parameters for the customer and firm are equals, i.e., $v = R$ and $h = c$, the solution

$(q^*, \widehat{q})$ lies in region I or II. In this specific setting, the solution falls in region II, so there is no equilibrium

with influential cheap talk. We again plot the the probability of joining for the signals m_1 and m_2 in each

period $1, \ldots, 8$. We observe that the firm cannot create any credibility with regard to the signals and, eventually customers simply join the system regardless of the message used by the firm. As demonstrated in Proposition 3.5, the only non-informative pure strategies equilibrium is one in which all customers join the system.
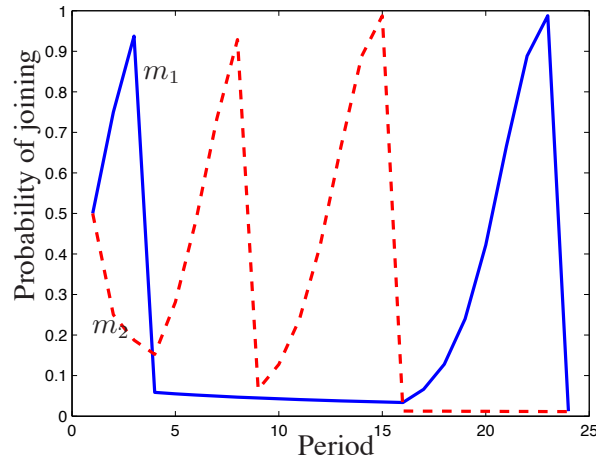

Figure EC.4   Tatonnement scheme for Region IIIb, the case of no equilibrium. The probability of joining for both message m_1 and message m_2 oscillates.

In the last example we again modify the gain obtained by the customers to $R = 0.9$. The solutions to the full control and full state information problems $(q^*, \widehat{q})$ lie in Region IIIb. As shown in Propositions 3.4 and 3.5, there is neither a pure strategies equilibrium with influential cheap talk nor a babbling equilibrium. As one can observe, the iterative process does not converge. In this case the firm does not establish credibility with respect to either signals. We observe that whenever the probability of joining becomes high the firm, in an attempt to maximize profit myopically, uses a threshold close to $\widehat{q}$. Given that the system lies in Region IIIb, however, the expected utility of the customer is negative and the customers do not want to join the system. However, if very few customers join the system, then the congestion is very low. So everyone wants to join to increase their utilities. This leads to the oscillatory behavior in figure EC.4.

EC.3.2. Optimal Capacity Selection

In this section, we conduct a numerical study and extend our model to include the problem of service-rate selection for the firm. Thus, the firm problem now has two stages: the firm first selects a service rate; and

second, it decides what and how real-time information is provided to its customers. Note that in the latter decision he also accounts for the fact that the customers are strategic. Thus, by choosing the service rate, the firm also chooses its ability to influence customer behavior announcements. We assume that the capacity cost is linear in the service rate with parameter K. Thus the cost of selecting a service rate μ is $K\mu$.

For the study, we assume that the customers arrive according to a Poisson process with rate $\lambda = 50$ per min. The firm obtains a value $v = 100$ for each customer it serves. The firm experiences a linear cost of holding with $h(w) = w$, i.e., the cost of holding a customer for one min is 1. The cost of one unit of capacity is $K = 4$. The customer obtains a value $R = 60$ from the service. To illustrate that the equilibria characterized in Section 3 may arise in the continuation game, we consider three settings which differ in the cost of waiting for the customer.

Influential equilibrium with intentional vagueness. In the first study, we assume that the cost of waiting for the customer is $c = 100$. In this case, we computed the firm's optimal capacity selection by a line search taking into account the resulting equilibrium for the continuation game. We obtain that the firm will choose $\mu^* = 75$, and will be able to influence customer behavior. However, as $q^* = 45$ (for this service rate) whereas the $\hat{q} = 50$, thus the firm must use intentional vagueness to sustain this equilbrium.

Influential equilibrium with no vagueness (Perfect alignment). For the next example, we assume that the cost of waiting for the customer is $c = 90$. In this case, we computed the firm's optimal capacity selection by a line search taking into account the resulting equilibrium for the continuation game. We obtain that the firm will choose $\mu^* = 75$, and will be able to influence customer behavior. However, as $q^* = 50$ (for this service rate) whereas the $\hat{q} = 50$, thus the firm may sustain this equilibrium even without resorting to intentional vagueness.

Non-influential equilibrium. For the last example, we assume that the cost of waiting for the customer is $c = 20$. In this case, we computed the firm's optimal capacity selection by a line search taking into account the resulting equilibrium for the continuation game. For this case, we find that for all service rates, the emerging equilibrium language of the continuation game is non-influential. Further, we obtain that the firm will choose $\mu^* = 54$, and will not able to influence customer behavior. Specifically, for the optimal μ^* we have that $q^* = 162$ whereas $\hat{q} = 16$.

EC.4. Abandonment-Proofness

In many service settings, customers can make a decision not only regarding joining vs. balking, but also about leaving the system after joining but before receiving service. So far, we focused on the first two decisions, while disallowing customer abandonment. In this section we show that if the customers are allowed to update their belief about the system and renege the queue, the equilibria characterization remains unchanged. Mandelbaum and Shimkin (2000) and Haviv and Ritov (2001) shows that if the customers have no information about the system (and the cost of waiting is linear and the system is Markovian) then it is in the best interest of the customers that they do not abandon from the system once they join. Thus, our babbling equilibrium (mixed or pure) are abandonment-proof, i.e., even when the customers have an option to abandon the system after joining they will not exercise this option. The next result shows that the equilibria identified in Sections 3 and EC.6 are also abandonment-proof.

Proposition EC.4.1 *In the equilibria identified in Propositions 3.4 and EC.6.3, a rational customer will not abandon even if allowed to and the firm will not deviate from its signaling rule. In this sense, these equilibria are abandonment-proof.*

The above result states that a rational customer who updates his belief on the state of the system *after* joining the system, would not abandon. We show that even if customers solve an optimal stopping time problem, since the hazard rate of the waiting time distribution is increasing, customers will not abandon once they join the system. While the customer might realize the fact that he was lured to join in a state he otherwise would not join, he is in a better position compared to the one in which he decided to join. Since it is optimal for him to "let bygones, be bygones," and regard the time elapsed as sunk cost, it will be better for him to stay in the system. Moreover, when the customers can leave the system after joining, before the commencement of their service, the full control solution outlined in Section 3 will not change. That is, the firm would not be interested in allowing a customer to join and then leave the system before he gets served. Thus, it will not provide signals that will allow for such a behavior. In Crawford and Sobel (1982), the receiver of the information cannot verify the state of the world until the game is completed and payoffs are received. However, in many service systems the customer may be able to update his beliefs on

the information by joining the system. Our result above shows that even when the customer has a recourse action (to abandon the system) after updating his beliefs, the structure of the equilibria is based on the first time instance of the interaction.

Hence, in this setting rational abandonments will not arise endogenously. It is important to note that this result hinges on the linearity of the cost of waiting by the customer. Other more complex settings, such as the one in which the valuation varies over time or the waiting cost is not linear (see Haviv and Ritov (2001)) or one in which the customers feel that they have been left out of the system without being informed (see Mandelbaum and Shimkin (2000)), can lead to rational abandonments. In our setting these complex systems would entail that the MDP in the definition of MPBNE suffers from the curse of dimensionality. Thus, when embedded in a game, it will be hard to compute the arising equilibria.

EC.5. Allowing the firm to randomize

In this section, we explore how randomization by the firm over the signals it provides to the customers impacts the possibility of influential cheap talk. To this end, we need to enrich the definition of mixed strategy MPBNE. Let $g(q, m)$ be the probability that the firm signals m to an arriving customer when there are q customers in the system. Condition 2 in the definition of mixed strategies (Definition EC.6.1) is thus changed as follows: With $f(j) = v - \mathbb{E}[h(W(j+1))]$, where $W(j+1)$ is the time spent by a customer who joins the system with j customers, there exists constants $J_0, J_1, \ldots$, and γ that solve the following set of equations:

$$
\begin{aligned}
J_0 &= \max_{m \in \mathcal{M}} \left\{ \frac{f(0)y(m) - \gamma}{\lambda} + J_0(1 - y(m)) + J_1 y(m) \right\} \\
&= \frac{f(0)y(m') - \gamma}{\lambda} + J_0(1 - y(m')) + J_1 y(m'), \quad \text{for all } m' \text{ with } g(0, m') > 0 \\
J_q &= \max_{m \in \mathcal{M}} \left\{ \frac{f(q)y(m) - \gamma}{\lambda + \mu} + \frac{\mu}{\lambda + \mu} J_{q-1} + \frac{\lambda}{\lambda + \mu}(J_q(1 - y(m)) + J_{q+1} y(m)) \right\} \\
&= \left\{ \frac{f(q)y(m') - \gamma}{\lambda + \mu} + \frac{\mu}{\lambda + \mu} J_{q-1} + \frac{\lambda}{\lambda + \mu}(J_q(1 - y(m')) + J_{q+1} y(m')) \right\}, \quad \text{(EC.8)}
\end{aligned}
$$

for all m' with $g(q, m') > 0$. The above condition states that the firm would signal any message with positive probability if and only if signaling that message results in a joining probability which solves the above MDP.

We will next show that even if equipped with the option of randomizing, the equilibrium language will remain a two-signal one (subject to DOE).

Proposition EC.5.1   *Given any MPBNE (y, g) in which the firm randomizes, there exists a DOE MPBNE such that the firm uses only two signals. Further, this DOE MPBNE can be constructed such that the firm randomizes on at most one state.*

The above proposition shows that when allowing the firm to randomize, the resulting equilibrium is equivalent to one in which the firm provides only two-signals. This stems from the fact that there is at most one state of the system in which the firm might be indifferent among various actions taken by the customers. However, on this state the firm will not randomize over messages that result in more than two actions, i.e., the resulting actions will have positive probability over $\max_{m \in \mathcal{M}} y(m)$ and $\min_{m \in \mathcal{M}} y(m)$. If the firm signals any other message with positive probability in this state, the customers would learn the state (as there is no vagueness left in the signal about the state of the system).

EC.6. Accounting for mixed strategies on the customer's end

In this section we relax the assumption that the customer, given an announcement, either joins with probability 1 or balks with probability 1. We next allow the customers to randomize.

Towards the definition of MPBNE with this relaxation, as before we use $g(q) \in \mathcal{M}$ to denote the signal that the firm uses when the queue length is q, for $q \in \mathbb{Z}_+$. The firm's policy is then given by the function $g(\cdot)$. The customers' possible randomization is modeled by allowing $y(m)$ for $m \in \mathcal{M}$ to take any value in $[0, 1]$. A value that is strictly in $(0, 1)$ implies that the customer randomizes. That is, $y(m) = p$ implies that the customers will join with probability p whenever given the announcement m, and will balk otherwise. Thus, the action space of the customers is now continuous and uncountable.

Given strategies g and y for the firm and customers respectively, we let $p_q(g, y)$ be the steady-state probability of the number of customers in the system being equal to q, in a queuing system with service rate μ and state-dependent arrivals $\lambda(q) = \lambda y(g(q))$ in state q. The notion of Mixed-Strategies MPBNE is then defined as follows:

Definition EC.6.1　(**Mixed-Strategies Markov Perfect Bayesian Nash Equilibrium**) *We say that the signaling rule $g(\cdot)$ and the action rule $y(\cdot)$ constitute a Mixed-Strategies MPBNE, if $g(\cdot)$ and $y(\cdot)$ satisfy Condition 2 in Definition 3.1 in addition to*

$$y(m) = \arg\max_{\eta \in [0,1]} \frac{\sum_{\{q:g(q)=m\}}\left[R - c\frac{q+1}{\mu}\right]p_q(y,q)}{\sum_{\{q:g(q)=m\}}p_q(y,g)}\eta \quad \textit{for all } m \in \mathcal{M}.$$

In contrast to condition 1 in Definition 3.1, here we allow the customers to randomize if they are indifferent between joining and balking the system. The next result characterizes signaling strategies that can form a Mixed strategies MPBNE.

Proposition EC.6.1　*Let (y,g) be a Mixed-Strategies MPBNE. Then, for any two states $q_1 \neq q_2$, if $g(q_1) = g(q_2)$, then we have $g(q) = g(q_1)$ for all $q \in [q_1, q_2]$.*

Proposition EC.6.1 implies that it is never optimal for the firm to use the same signal for two states but use some other signal for a state in between. The importance of this result is in showing that a firm will never provide the same signal on two disjoint intervals, and provide a different signal in between.

EC.6.1. Babbling equilibrium

We have shown in Section 3.3 that babbling equilibria need not exist in pure-strategies. When customers are allowed to randomize, however, such equilibria always exist. Moreover, if $R \geq c/(\mu - \lambda)$, the resulting equilibrium is identical to the one characterized in Section 3.3. In the complementary case, when $R < c/(\mu - \lambda)$, there is no pure strategy MPBNE. In contrast, when customers can randomize among joining and balking, they can always form a mixed strategy MPBNE as follows: they choose a probability of joining θ that satisfies $R = c/(\mu - \theta\lambda)$. As $R < c/(\mu - \lambda)$, θ is guaranteed to lie in the interval $(0,1)$. Under this equilibrium, the arrival process is thinned by the customer randomization such that an arriving customer is indifferent between joining and balking. In particular, the customers do not have any profitable deviation. Due to the monotonicity of the function $c/(R\mu - \theta\lambda)$ in θ, there is no other randomization strategy that can form a babbling equilibrium in this setting. This is summarized in the following result.

Proposition EC.6.2

1. *If $R \geq c/(\mu - \lambda)$, the unique mixed strategies babbling equilibrium is one in which all customers join the system.*

2. *If $R < c/(\mu - \lambda)$, the unique mixed strategies babbling equilibrium is one in which an arriving customer joins the system with probability $\theta = \frac{R\mu - c}{R\lambda}$ and balks otherwise. The customer's expected utility under this equilibrium is zero.*

Thus, the randomization by customers alleviates the unpredictability of the customer behavior. This suggests in particular, that in absence of any meaningful language between the customer and the firm, the latter can always resort to silence. The customers, however, are always weakly better off in any other equilibria.

Influential cheap talk can be viewed as a mechanism to coordinate the incentives of the service provider and the customers when transmitting information which cannot be verified. If only babbling equilibrium exists, it might suggest that the non-credibility is hampering any possibility of coordination whatsoever between the players. This is exactly the issue we explore in the next section in which we examine whether there is a possibility of improvement in the coordination between the service provider and its customers.

EC.6.2. Mixed strategy informative equilibrium

In this section we show that in addition to the babbling equilibria, there may exist more *informative* MPBNE in mixed strategies. We start by showing that it suffices to consider equilibria with certain characteristics. Specifically, in the following proposition we show that, if customers randomize, they will do so only within an interval. Furthermore, we show that balking with probability 1 in some states and joining with probability 1 in other states cannot co-exist. We discuss this structural observation following the statement of the proposition.

Proposition EC.6.3

Consider a mixed strategy MPBNE (y, g) and suppose that there is a signal m_i such that $y(m_i) \in (0, 1)$ and m_i is transmitted with positive probability in equilibrium, i.e., $\sum_{\{q:g(q)=m_i\}} p_q(y, g) > 0$. Then, there exist q_1 and q_2 such that

$$y(g(q)) = \begin{cases} 1 & q < q_1, \\ y(m_i) & q \in [q_1, q_2], \\ 0 & otherwise. \end{cases} \tag{EC.9}$$

Furthermore, q_1 and q_2 are such that $q^* \in [q_1, q_2]$ and either $q_1 = 0$ or $q_2 = \infty$ or both[10].

Proposition EC.6.3 implies that if the customers randomize in equilibrium, then there cannot exist signals m_j and m_k such that $y(m_j) = 0$ and $y(m_k) = 1$ where both of them are used with strictly positive probability. Thus, there are only two possible types of two-signal mixed strategy MPBNE in which randomization is used. The two types can be described as follows: The firm announces "High" and "Low" based on the threshold q_{mix}, the customers then react as follows: (a) in the first type of MPBNE which we shall refer to as *Join or Randomize* equilibria, the customers who receive "Low" join the system and the customers who receive "High" join the system with probability $\theta \in (0, 1)$; (b) in the second type of MPBNE which we shall refer to as *Randomize or Balk* equilibria, the customers who receive "Low" join the system with probability θ, and the customers who receive "High" balk. Both of these equilibria are completely defined by two parameters: the threshold q_{mix} used by the firm for signaling and the randomization parameter θ. We provide a characterization of the pair (q_{mix}, θ) which induces the mixed strategy MPBNE in the next subsection (Appendix EC.6.3.)

In Proposition 3.4, we identified two cases, namely Case II and Case III(b) where a pure strategy equilibrium with influential cheap talk does not exist. Using the characterization in Appendix EC.6.3, one can verify that for the parameters in case III(b), there always exists a *Randomize or Balk* mixed strategy MPBNE (up to integrality issues). For case II, however, either type of mixed strategy MPBNE may exist. It is also possible that neither exist but a mixed strategy babbling equilibrium does exist. Note that the firm already achieves its first-best profits in Cases I and III(a), thus the firm prefers the pure strategy MPBNE over any other possible mixed strategy MPBNE.

So far, we characterized and proved the existence (or non-existence) of two-signal communication between the firm and its customers that form a MPBNE. We next turn to the question regarding the existence of additional equilibria.

---

[10] If both $q_1 = 0$ and $q_2 = \infty$ then this MPBNE is a babbling equilibrium.

EC.6.3. Characterization of mixed strategy MPBNE.

In Proposition EC.6.3 we established that there are two possible types of mixed strategies MPBNE that we named *Join or Randomize* and *Randomize or Balk*. In this section we explicitly characterize this possible equilibria as a function of the system parameters. For each of the two possible equilibria, we identify the equations that are satisfied by the threshold q_{mix} and the randomization probability θ. We start with the *Randomize or Balk* equilibrium.

Randomize or Balk

Let (q_{mix}, θ) induce a mixed strategy MPBNE of *Randomize or Balk* type, i.e., the firm announces "Low" if the queue length is less than q_{mix} and "High" otherwise. Given a "Low" signal, the customer joins the system with probability θ and balks otherwise. The customer always balks if he receives the signal "High." Let $p_q^{\theta,k}$ be the steady-state probability that the queue length equals q in an $M/M/1/k$ queuing system with arrival rate $\theta\lambda$ and service rate μ. For this equilibrium to be immune to any profitable deviation by a customer, (q_{mix}, θ) must satisfy the following condition:

$$R - c \sum_{q=0}^{q_{mix}-1} p_q^{\theta,q_{mix}} \frac{q+1}{\mu} = 0. \tag{EC.10}$$

This condition is a consequence of the observation that when a customer randomizes, he must be indifferent between joining and balking. Hence, his expected utility over the randomization interval must be zero. Otherwise, there would be a profitable deviation by balking with probability one or joining with probability one. Hence, we must have that $R = cE[W]$ which leads to (EC.10) through

$$E[W] = \sum_{q=0}^{q_{mix}-1} p_q^{\theta,q_{mix}} \frac{q+1}{\mu}.$$

Note that for any $\theta \in (0,1)$, the solution q_{mix} of the above condition is greater or equal to q^*. Hence, if the customer receives the signal "High", he would obtain a strictly negative expected utility by joining.

We now consider the firm's decision. Towards this end, we let $\Pi^\theta(k)$ be the analogue of $\Pi(k)$ (as defined in Proposition 3.3) but with arrival rate $\theta\lambda$ rather than λ. It is clear that for the *Randomize or Balk* equilibrium to be immune to any profitable deviation by the firm we must have that (q_{mix}, θ) satisfies

$$q_{mix} \in \arg\max_{q \geq 0} \Pi^\theta(q). \tag{EC.11}$$

Equation (EC.10) and (EC.11) together characterize the *Randomize or Balk* equilibrium when it exists. We end this subsection by stating that *Randomize or Balk* equilibria cannot co-exist with a **pure-strategy** babbling equilibria.

Proposition EC.6.4   *For any system parameters, for which Randmize/Balk MPBNE exists, there cannot be a pure strategy babbling equilibrium.*

We now turn to *Join or Randomize* equilibria.

Join or Randomize

Let (q_{mix}, θ) induce a mixed strategy MPBNE i.e, the firm announces "Low" if the queue length is less than q_{mix} and "High" otherwise. The customer joins with probability 1 if he receives the signal "Low". If, on the other hand, he receives the signal "High", then he joins with probability θ and balks otherwise.

Towards the characterization of the equilibrium in this case we note first that, given the signal "High", the average expected waiting time of a customer that decides to join is given by

$$\left[\frac{q_{mix}}{\mu} + \frac{1}{\mu - \theta\lambda} \right].$$

For the equilibrium to be immune to any profitable deviation by the customers, then that (q_{mix}, θ) must satisfy the following condition:

$$R - c \left[\frac{q_{mix}}{\mu} + \frac{1}{\mu - \theta\lambda} \right] = 0, \tag{EC.12}$$

so that customers will be indifferent between joining and balking. Indeed, condition (EC.12) ensures that the customer's expected utility by joining is zero, which is the same as his expected utility if he decides to balk. Note that for any $\theta \in (0, 1)$ the solution q_{mix} of the above condition is less than q^*; thus, if the customer receives the signal "High", he would obtain a strictly positive expected utility by joining.

Finally, for this equilibrium to be immune to any profitable deviation by the firm, we must have that, given θ, q_{mix} is the optimal "switching-point" for the firm, where the switching is between an arrival rate of λ to one of $\theta\lambda$. In other words,

$$q_{mix} \in \arg\max_{q \geq 0} \widetilde{\Pi}^\theta(q), \tag{EC.13}$$

where

$$\widetilde{\Pi}^{\theta}(k) := \sum_{q=0}^{\infty} \lambda(q) \tilde{p}_q^{\theta,k} \left[v - \mathbb{E}[h(W(q+1))] \right].$$

Here, $\tilde{p}_q^{\theta,k}$ is the steady-state distribution of a state-dependent $M/M/1$ queue with arrival rate $\lambda(q) = \lambda$ for all $q < k$ and arrival rate $\lambda(q) = \theta\lambda$ for all $q \geq k$.

EC.7. Proofs of Results in the Main Paper

Proof of Proposition 3.1 For any pure strategy influential equilibria, there exists two states q_1 and q_2 such that $y(g(q_1)) = 1$ and $y(g(q_2)) = 0$. Thus there exists a finite N as defined in the proposition. The rest of the proof is omitted as the more general result is proven in Proposition EC.6.1. ■

Proof of Proposition 3.2: Consider any MPBNE (y, g). Define $N = \inf\{q : y(g(q)) = 0\}$ as well as $(\widetilde{y}(\cdot), \widetilde{g}(\cdot))$ using (3) and $\overline{q} = N - 1$. We next show that the pair $(\widetilde{y}(\cdot), \widetilde{g}(\cdot))$ induces the same MPBNE as $(y(\cdot), g(\cdot))$, in terms of the customers' utility and the firm's profit. To this end, assume there exists $m_2 \neq m_3$ such that $y(m_2) = y(m_3) = 1$. Then, by definition of MPBNE we have that

$$\frac{\sum_{\{q:g(q)=m_2\}} \left[R - c\frac{q+1}{\mu}\right] p_q(y, g)}{\sum_{\{q:g(q)=m_2\}} p_q(y, g)} \geq 0, \tag{EC.14}$$

$$\frac{\sum_{\{q:g(q)=m_3\}} \left[R - c\frac{q+1}{\mu}\right] p_q(y, g)}{\sum_{\{q:g(q)=m_3\}} p_q(y, g)} \geq 0. \tag{EC.15}$$

In particular, by the laws of conditional expectation, we have that

$$\frac{\sum_{\{q:g(q)\in\{m_2,m_3\}\}} \left[R - c\frac{q+1}{\mu}\right] p_q(y, g)}{\sum_{\{q:g(q)\in\{m_2,m_3\}\}} p_q(y, g)} \geq \min_{i=2,3} \frac{\sum_{\{q:g(q)=m_i\}} \left[R - c\frac{q+1}{\mu}\right] p_q(y, g)}{\sum_{\{q:g(q)=m_i\}} p_q(y, g)} \geq 0. \tag{EC.16}$$

Consequently, the pair $(\widehat{g}(\cdot), \widehat{y}(\cdot))$ defined by setting $\widehat{g}(q) = m_4, \ \forall q : g(q) \in \{m_2, m_3\}$ and $\widehat{g}(q) = g(q)$ otherwise, and letting $\widehat{y}(m_4) = 1$ and $\widehat{y}(m) = y(m)$ otherwise, is itself a MPBNE.

A similar reasoning applies if there exists $m_4 \neq m_5$ such that $y(m_4) = y(m_5) = 0$. By sequentially applying this reasoning, and noting the fact that q obtains values in a countable set, we have that the pair $(\widetilde{g}(\cdot), \widetilde{y}(\cdot))$ is a MPBNE. Finally, note that $\widetilde{y}(\widetilde{g}(q)) = y(g(q))$ for all $q = 0, 1, \ldots$. Consequently, by (EC.8) the firm's profit and the queue dynamics (and steady-state distribution) remain unchanged which implies also that the customer's utility remains unchanged. This completes the proof. ■

Proof of Proposition 3.3: We may re-write $\Pi(k)$ as follows:

$$\Pi(k) = \lambda \sum_{j=0}^{k-1} p_j^k \left[v - \mathbb{E}[h(W(j+1))] \right], \ k = 1, 2, \ldots,$$

where p_j^k is the steady-state distribution of having j customers in system at the $M/M/1/k$ queue, $W(j+1)$ is the sojourn time for a customer entering when there are j customers in the system. Clearly, $W(j)$ is a Gamma distributed random variable with parameters $j+1$ and μ. We now define for $j \geq 0$, $\alpha_j := v - \mathbb{E}[h(W(j+1))]$ and write

$$\Pi(k) = \lambda \sum_{j=1}^{k-1} p_j^k \alpha_j.$$

Having written the profit function $\Pi(k)$ in this form, we have positioned ourselves in the framework of Knudsen (1972); see equation 4.1 there. The fact that α_j is a decreasing sequence follows from Proposition 1 in Knudsen (1972).

Finally, by Theorem 1 in Knudsen (1972), and using the fact that α_j is decreasing sequence, we have that the function $\Pi(\cdot)$ is unimodal. This completes the proof. ∎

Proof of Proposition 3.4:

I. Since the firm's and the customer's incentives are completely aligned, it is easy to see that the pair $(\widetilde{g}(\cdot), \widetilde{y}(\cdot))$, defined as in Proposition 3.2 with $\bar{q} = q^*$ forms a pure strategy MPBNE.

II. If $q^* > \widehat{q}$, we can observe that any strategy whose corresponding induced threshold is not equal to $\widehat{q}$ does not satisfy the second condition in the definition of pure strategy MPBNE. This is due to the unimodality of the profit function $\Pi(\cdot)$. Further, the threshold level $\widehat{q}$ does not satisfy the first condition in the Definition EC.2.1. To see this, assume that it does induce a pure strategy MPBNE. Then, in the corresponding pure strategy MPBNE we have two signals: m_1 standing for "Low" and m_2 standing for "High" and in order to satisfy condition 1 of the pure strategy MPBNE, we need to have that $y(m_1) = 1$ and $y(m_2) = 0$. But, as $\widehat{q} < q^*$, we have that

$$\frac{\sum_{\{q:g(q)=m_2\}} \left[R - c\frac{q+1}{\mu} \right] p_q^{\widehat{q}}}{\sum_{\{q:g(q)=m_2\}} p_q^{\widehat{q}}} = \left[R - c\frac{\widehat{q}+1}{\mu} \right] p_{\widehat{q}}^{\widehat{q}} \geq 0,$$

where the first equality follows from the fact that if the threshold $\widehat{q}$ induces a pure strategy MPBNE then N in Definition EC.2.1 is $\widehat{q}$. Consequently, we would have that $y(m_2) = 1$ and the threshold $\widehat{q}$ can not induce a pure strategy MPBNE.

III. **(a)** If $q^* < \widehat{q}$ and $G(0, \widehat{q}) \geq 0$, then $F(\widehat{q}) \leq 0$, and thus the proposed pure strategy MPBNE satisfies the first condition in the definition of pure strategy MPBNE. Also, using the optimality of $\widehat{q}$, we get that it satisfies the second condition in the definition of pure strategy MPBNE.

(b) Suppose that $q^* < \widehat{q}$ and $G(0, \widehat{q}) < 0$. Let $\widetilde{q} := \inf\{q \in \mathbb{Z}_+ : G(0, q) \leq 0\}$. Now, any threshold strategy with a threshold that is greater than $\widetilde{q}$ cannot satisfy the first condition in the definition of pure strategy MPBNE (Definition EC.2.1). Note that for any $q \in \{0, \widehat{q}\}$ the MDP in condition 2 of the definition of MPBNE is identical to the MDP for the full control problem. Further, we have that the full control problem is identical to maximizing the function $\Pi(\cdot)$. Thus, using the monotonicity of $\Pi(\cdot)$ on $[0, \widehat{q}]$ (see Proposition 3.3) implies that any strategy obtained from a threshold that is strictly smaller than $\widehat{q}$ cannot satisfy the second condition in the definition of pure strategy MPBNE. Lastly, we can see that the threshold strategy with threshold $\widetilde{q}$ does not satisfy the second condition in the definition of pure strategy MPBNE conditions. This completes the proof.

■

Proof of Proposition 3.5: The first part of the result follows directly from the discussion above. We now turn to region III(b) and show that the babbling equilibrium does not exist as a customer always has a profitable deviation. Recall that $G(0, k)$ is the customer's utility from joining a $M/M/1/k$ queue. Similar to Proposition 4.2, it can be shown that $G(0, k)$ is monotone decreasing in k. Hence, the customer's utility under the babbling equilibrium is equal to $\lim_{k \to \infty} G(0, k)$ and bounded above by $G(0, \widehat{q})$, which is negative. This completes the proof. ■

Proof of Proposition 3.6: We first focus on regions I and III(a). Recall that by the definition of $\widehat{q}$ and the unimodality of the firm's profit function $\Pi(\cdot)$ (see Proposition 3.3), we have that the firm's profit is decreasing over the region $[\widehat{q}, \infty)$. Further, appealing to Proposition 3.4 for regions I and III(a), we have that the threshold that induces the equilibrium is higher than $\widehat{q}$. Consequently, using the fact that $\pi^{ni} = \lim_{k \to \infty} \Pi(k)$, we have that the firm's profit in any of these regions is greater than or equal to the babbling equilibrium. A similar argument is applied to the customer's expected utility. Let the overall customers' expected utility be maximized at a point q^{SO}. Using Section 4 from Naor (1969), we have that the expected

overall utility of the customers is unimodal in q and $q^{SO} < q^*$. Also, we have that the threshold that induces

the equilibrium in Proposition 3.4 is higher than q^*. This completes the proof for regions I and III (a).   ■

Proof of Proposition 4.1:

I. Note that we cannot partition the set $\mathbb{Z}_+$ further. The proof then follows directly as the set $\mathbb{Z}_+$ induces

the pure strategy MPBNE described in Proposition 3.4 (I).

II.(a-c) The proof follows from the definition of q_{mix}^{JR} and q_{mix}^{RB} and noting that if the the identified set is

partitioned further, it will not a MPBNE using Proposition EC.6.3.

III.   a) Clearly, this set induces the pure strategy MPBNE described in Proposition 3.4 for case III(a). It

suffices to show that we can not partition the set $\mathbb{S}$ such that the resulting induces a mixed or pure strategy

MPBNE.

We focus first on pure strategies. To reach a contradiction, assume that there exists a set $\mathbb{S}' \supsetneq \mathbb{S}$ that

induces a pure strategy MPBNE. Then, there is q' with $\bar{q} < q' < \hat{q}$, such that the customers receive a different

signal on $[\bar{q}, q')$ then on $[q', \hat{q})$. But, by the definition of $\bar{q}$ and condition 1 of the pure strategy MPBNE

(see Definition EC.2.1), the customers would not join given the signal on $[q', \hat{q})$. Moreover, since $\Pi(\cdot)$ is

strictly monotone increasing for $q \leq \hat{q}$ (see Proposition 3.3) and since there exists, by Proposition 3.4, a

pure strategy MPBNE induced by threshold $\hat{q}$, the firm would be better off deviating.

Lastly, we shall focus on mixed strategies. To arrive at contradiction, assume that there exists a set $\mathbb{S}' \supsetneq \mathbb{S}$

that induces a mixed strategy MPBNE. Using proposition EC.6.3, there could only be two types of equilib-

ria. It is easy to note that this mixed strategy equilibria must be of type Randomize/Balk. Let the parameter

for this mixed strategy MPBNE be (q_{mix}, θ) where $0 < \theta < 1$. Given that $\mathbb{S}' \supsetneq \mathbb{S}$, it is easy to see that $q_{mix} <$

$\hat{q}$. Further, it must be the case that $\bar{q} := \arg\max\{q : G(q, \hat{q}) > 0\} = 0$. Define $\tilde{q} = \arg\max\{q : G(0, \tilde{q}) > 0\}$.

Note that $\tilde{q} \geq \hat{q}$. From Appendix EC.6.3, we further know that $R - c\sum_{q=0}^{q_{mix}-1} p_q^{\theta, q_{mix}} \frac{q+1}{\mu} = 0$. Thus, we

have that $q_{mix} \geq \tilde{q} \geq \hat{q}$ leading to a contradiction.

Thus, the set $\{0, 1, 2, \ldots, \bar{q}, \hat{q}, \hat{q}+1, \ldots\}$ induces the most informative pure strategy MPBNE.

b) The argument is very similar to part II.

    ■

EC.8. Proof of the results in the e-companion

Proof of Proposition EC.2.1: Consider any MPBNE (y, g). Define $q_\omega = \inf\{q : y(g(q, \omega)) = 0\}$ as well as $(\widetilde{y}(\cdot), \widetilde{g}(\cdot))$ using (EC.2). We next show that the pair $(\widetilde{y}(\cdot), \widetilde{g}(\cdot))$ induces the same MPBNE as $(y(\cdot), g(\cdot))$, in terms of the customers' utility and the firm's profit. To this end, assume there exists $m_2 \neq m_3$ such that $y(m_2) = y(m_3) = 1$. Then, by the definition of MPBNE we have that

$$\frac{\sum_{\{q:g(q)=m_2\}} \left[R - c\frac{q+1}{\mu}\right] p_q(y, g)}{\sum_{\{q:g(q)=m_2\}} p_q(y, g)} \geq 0, \tag{EC.17}$$

$$\frac{\sum_{\{q:g(q)=m_3\}} \left[R - c\frac{q+1}{\mu}\right] p_q(y, g)}{\sum_{\{q:g(q)=m_3\}} p_q(y, g)} \geq 0. \tag{EC.18}$$

In particular, by the laws of conditional expectation, we have that

$$\frac{\sum_{\{q:g(q)\in\{m_2,m_3\}\}} \left[R - c\frac{q+1}{\mu}\right] p_q(y, g)}{\sum_{\{q:g(q)\in\{m_2,m_3\}\}} p_q(y, g)} \geq \min_{i=2,3} \frac{\sum_{\{q:g(q)=m_i\}} \left[R - c\frac{q+1}{\mu}\right] p_q(y, g)}{\sum_{\{q:g(q)=m_i\}} p_q(y, g)} \geq 0. \tag{EC.19}$$

Consequently, the pair $(\widehat{g}(\cdot), \widehat{y}(\cdot))$ defined by setting $\widehat{g}(q, \omega) = m_4$, $\forall q, \omega : g(q, \omega) \in \{m_2, m_3\}$ and $\widehat{g}(q, \omega) = g(q, \omega)$ otherwise, and letting $\widehat{y}(m_4) = 1$ and $\widehat{y}(m) = y(m)$ otherwise, is itself a MPBNE.

A similar reasoning applies if there exists $m_4 \neq m_5$ such that $y(m_4) = y(m_5) = 0$. By sequentially applying this reasoning, and noting the fact that q obtains values in a countable set, we have that the pair $(\widetilde{g}(\cdot, \cdot), \widetilde{y}(\cdot))$ is a MPBNE. Finally, note that $\widetilde{y}(\widetilde{g}(q, \omega)) = y(g(q, \omega))$ for all $q = 0, 1, \ldots$. Consequently, by (EC.8) the firm's profit and the queue dynamics (and steady-state distribution) remain unchanged which implies also that the customer's utility remains unchanged. This completes the proof.   ∎

Proof of Proposition EC.4.1. We need the following two lemmas:

Lemma EC.8.1 *In an $M/M/1/N$ system, the waiting time of the customer has an increasing hazard rate.*

Lemma EC.8.2 *Consider an $M/M/1$ system where the arrival rate is λ for the states $1, \ldots, \widehat{q}$ and λ' for states $\widehat{q} + 1, \ldots$, where $0 < \lambda' < \lambda$. We have*

1. The random variable $W(Q)$ conditioned on the event $\{Q \leq \widehat{q}\}$ has increasing hazard rate.

2. The random variable $W(Q)$ conditioned on the event $\{Q > \widehat{q}\}$ has increasing hazard rate.

The proof of the above results are relegated to Appendix EC.9 The system described in Lemma EC.8.1 and EC.8.2 corresponds to the equilibirum described in Propositions 3.4 and EC.6.3. The lemmas imply that conditioned on joining and the information received by the customer, the waiting time in the queue has an increasing hazard rate. Using the arguments similar to the proof of Proposition 2(i) in Mandelbaum and Shimkin (2000), we obtain that the customers will never abandon the system if they join the system. This shows that the equilibria described earlier are abandonment-proof from the customer's perspective. Given that the customers do not abandon and the firm had no profitable deviation from its signaling rule, it follows that the firm will not have any profitable deviation when the customers are allowed to deviate. This completes the proof.   ∎

Proof of Proposition EC.5.1 Fix a MPBNE (y, g). Note that J_q is not impacted by the fact that the firm is allowed to randomize. Let $\check{q}^{g,y} = \inf\{q : J_{q+1} - J_q + f(q) \leq 0\}$. Using the arguments in proof of Proposition EC.6.1, we obtain the following: the set $\{q : J_{q+1} - J_q + f(q) = 0\}$ is either empty or equals $\{\check{q}^{g,y}\}$. We shall consider two scenarios: Case I: The set $\{q : J_{q+1} - J_q + f(q) = 0\}$ is empty. Then using the new condition from MDP, it must be the case that whenever $g(q, m) > 0$ for some $q \in \{0, 1, \ldots\}$ and $m \in \mathcal{M}$ then

$$y(m) = \max_{m \in \mathcal{M}} y(m) \text{ if } q < \check{q}^{g,y}, \quad y(m) = \min_{m \in \mathcal{M}} y(m) \text{ if } q \geq \check{q}^{g,y}.$$

Let m' be a signal such that $y(m') = \max_{m \in \mathcal{M}} y(m)$ and m'' be a signal such that $y(m') = \max_{m \in \mathcal{M}} y(m)$. Consider the following signaling rule: $\tilde{g}(q, m) = 1$, if either $q < \check{q}^{g,y}$ and $m = m'$, or if $q \geq \check{q}^{g,y}$ and $m = m''$. It is easy to see that $(\tilde{g}, y)$ forms a MPBNE where the firm only uses two signals on the equilibrium path. Further, this MPBNE is DOE with respect to the original (y, g) MPBNE. This completes the proof. Case II: The set $\{q : J_{q+1} - J_q + f(q) = 0\}$ is a $\{\check{q}^{g,y}\}$. Then using the new condition from MDP, it must be the case that whenever $g(q, m) > 0$ for some $q \in \{0, 1, \ldots\}$ and $m \in \mathcal{M}$ then

$$y(m) = \max_{m \in \mathcal{M}} y(m) \text{ if } q < \check{q}^{g,y}, \quad y(m) = \min_{m \in \mathcal{M}} y(m) \text{ if } q > \check{q}^{g,y}.$$

Let m' be a signal such that $y(m') = \max_{m \in \mathcal{M}} y(m)$ and m'' be a signal such that $y(m') = \max_{m \in \mathcal{M}} y(m)$.

Consider the following signaling rule: $\tilde{g}(q, m) = 1$, if either $q < \check{q}^{g,y}$ and $m = m'$, or if $q > \check{q}^{g,y}$ and $m = m''$. For $q = \check{q}^{g,y}$, we choose $\tilde{g}(q, m')$ that solves

$$\tilde{g}(q, m')y(m') + (1 - \tilde{g}(q, m'))y(m'') = \sum_{m \in \mathcal{M}} y(m)g(q, m),$$

and $\tilde{g}(q, m'') = 1 - \tilde{g}(q, m')$. One can show that $(y, \tilde{g})$ is a MPBNE. It is easy to see that it is in fact DOE with respect to (y, g) and uses only two signals m' and m'' in equilibrium. This completes the proof. ∎

Proof of Proposition EC.6.1: Focusing on the MDP (condition 2) in Definition EC.6.1 of Mixed-Strategies MPBNE, and using the arguments similar to those used in Section 2 of Stidham and Weber (1989) we have that the function J_q is strictly convex. Let $\check{q}^{g,y} = \inf\{q : J_{q+1} - J_q + f(q) \leq 0\}$. Note that due to the strict convexity of J_q and monotonicity of f, the set $\{q : J_{q+1} - J_q + f(q) = 0\}$ is either empty or equals $\{\check{q}^{g,y}\}$. Note that the condition 2, requires that $g(q) \in \arg\max_{m \in \mathcal{M}} y(m)(J_{q+1} - J_q + f(q))$. Thus, for (y, g) to be equilibrium, it must be the case that:

$$y(g(q)) = \sup_{m \in \mathcal{M}} y(m) \text{ for all } q < \check{q}^{g,y}, \quad y(g(q)) = \inf_{m \in \mathcal{M}} y(m) \text{ for all } q > \check{q}^{g,y}.$$

This completes the proof. ∎

Proof of Proposition EC.6.2

1. It is easy to see that if everyone joins then every customer obtains a strictly positive utility. Thus, this is indeed an equilibrium. Further, towards contradiction assume that there exist a $\theta < 1$ such that if all customers randomize and join with probability θ then it forms an equilibrium. However, given that with arrival rate λ all customers have strictly positive utility thus with arrival rate $\theta\lambda$ the customers who will join the system will have strictly positive utility as well. Thus, an arriving customer do not have any incentive of balking or randomizing. Hence we cannot sustain any $\theta < 1$ to form an equilibrium.

2. If every customer joins with probability $\theta = \frac{R\mu - c}{R\lambda}$ and balks otherwise, then the waiting time in the system is such that an arriving customer who decides to join the system will obtain zero utility. (This is based on the fact that the steady state waiting time in such a system is R/c.) Further, the balking customers also obtain zero utility. Hence, the proposed strategy does not have any profitable deviation and thus forms an equilibrium. To observe that this equilibrium is unique note that if the joining probability is greater than

(less than) the proposed θ then every customer who joins obtains strictly negative (positive) utility. Hence, an equilibrium cannot be sustained with any other randomization for the customer. This completes the proof.

∎

Proof of Proposition EC.6.3:

1. Fix a mixed strategy MPBNE with $y(m_i) \in (0, 1)$. Further, without loss of generality we assume $y(m_2) \in (0, 1)$. Let $p_q(y, g)$ be the induced steady state of the probability that there are q customers in the system under the mixed strategy MPBNE (y, g). Using the definition of the mixed strategy MPBNE. There exists q_1 and q_2 such that the firm's signals m_2 if the queue length is between q_1 and q_2. Further, since the customers are randomizing it must be the case that the customers are indifferent between joining and not joining. Thus, we have that

$$\frac{\sum_{q=q_1}^{q_2} \left[R - c\frac{q+1}{\mu} \right] p_q}{\sum_{q=q_1}^{q_2} p_q} = 0,$$

which implies that $q^* \in [q_1, q_2]$. Further, the firm signals m_i where $i \neq 2$, for some interval $[q_a, q_b]$, using the definition of mixed MPBNE it must be the case that $q_a > q^*$ or $q_b < q^*$. Thus, for any signal m_i, the expected utility of the customers would be either strictly positive or strictly negative depending on whether $q_b < q^*$ or $q_a > q^*$, respectively. Thus, for all other signals the customers will join with probability one or balk with probability one. This completes the proof.

2. By Proposition EC.6.1 it suffices to consider signals that are announced on intervals. That immediately implies that any equilibrium can be reduced to an equilibrium with three signals. Further, two of these signals must correspond to pure Join and pure Balk. Using the MDP approach, we know that for a given customer strategy there exists a solution which is threshold type. There are two cases to be considered here:

a) There is no state on which the firm is indifferent: In this case the firm would try to obtain its first-best and use only two signals due to a bang-bang solution for the MDP.

b) If the firm is indifferent in action on one or more state: Using the arguments in Proof of Proposition EC.6.1 it must be the case that the firm is indifferent at exactly one state. Let $\tilde{q}$ denote this state. Thus, the only equilibrium with three signal language possible in this case is if the firm signals Join on lower states (strictly less than $\tilde{q}$) and Balk on higher states (strictly greater than $\tilde{q}$), and randomize on $\tilde{q}$. Note that

the customer would then randomize when they receive the signal randomize only if $R - c(\tilde{q}+1)/\mu = 0$. For this to be an equilibrium, it must be the case that the parameters of the region lie in Region I. Thus, there cannot be a three signal equilibrium.

∎

Proof of Proposition EC.6.4. In Proposition 3.5 we proved that a pure-strategy babbling equilibrium exists if and only if $R \geq c/(\mu - \lambda)$. We recall, however, that in the *Randomize or Balk* equilibrium with threshold and randomization constant (q_{mix}, θ), the customers are indifferent between joining and balking and, in particular, the pair (q_{mix}, θ) satisfies (EC.10). Consequently, if the buffer size increases to infinity while keeping the arrival rate $\theta\lambda$ the utility will become negative and the customers will not join the system. In turn, we must have that

$$R < \frac{c}{\mu - \theta\lambda} < \frac{c}{\mu - \lambda},$$

which contradicts the condition for pure-strategy babbling equilibrium. ∎

EC.9. Proof of Auxiliary Results

Proof of Lemma EC.8.1: In an M/M/1/N system using the steady state analysis and the PASTA property, the number of customers in the system when a customer arrives and joins the system has a geometric distribution conditioned on it being less than N. Thus, we have that the waiting time in the system for a customer has the same distribution as $\sum_{j=1}^{Q+1} X_j$, where X_j are exponential with mean $1/\mu$ and Q has the same distribution as the number in system of $M/M/1/N$ conditioned on the event that the number of customers is less than N. We also know that the $Q+1$ has an increasing hazard rate. Using Theorem 7.1 in Ross et al. (2005), we have that the waiting time will also have an increasing hazard rate. This completes the proof. ∎

Proof of Lemma EC.8.2:

1. The proof follows exactly as in Lemma EC.8.1.

2. The proof follows by noting that the number of customers in the system is $\hat{q} + Q$, where Q is a geometric random variable. Using Theorem 7.1 of Ross et al. (2005), we obtain the result in a similar manner to Lemma EC.8.1. This completes the proof.

∎