*Adv. Appl. Prob.* **52,** 319–347 (2020)
*doi:*10.1017/apr.2019.64
© *Applied Probability Trust* 2020

# THE POTENTIALLY NEGATIVE EFFECTS OF COOPERATION IN SERVICE SYSTEMS

HAKJIN CHUNG,* *Korea Advanced Institute of Science and Technology*
HYUN-SOO AHN,** *University of Michigan*
RHONDA RIGHTER,*** *University of California, Berkeley*

**Abstract**

The 'Price of Anarchy' states that the performance of multi-agent service systems degrades with the agents' selfishness (anarchy). We investigate a service model in which both customers and the firm are strategic. We find that, for a Stackelberg game in which the server invests in capacity before customers decide whether or not to join, there can be a 'Benefit of Anarchy', that is, customers acting selfishly can have a greater overall utility than customers who are coordinated to maximize their overall utility. We also show that customer anarchy can be socially beneficial, resulting in a 'Social Benefit of Anarchy'. We show that such phenomena are rather general and can arise in multiple settings (e.g. in both profit-maximizing and welfare-maximizing firms, in both capacity-setting and price-setting firms, and in both observable and unobservable queues). However, the underlying mechanism leading to the Benefit of Anarchy can differ significantly from one setting to another.

*Keywords:* Queueing; Price of Anarchy; balking; individual versus social welfare

2010 Mathematics Subject Classification: Primary 90B22
Secondary 60K25

## 1. Introduction

In multi-agent service systems performance degrades with agents' selfishness. The social inefficiencies due to selfishness are often captured by the Price of Anarchy (PoA: see e.g. Roughgarden [18], Gilboa-Freedman *et al.* [6], and Zhang *et al.* [22]). For example, in routeing problems, the PoA is defined as the ratio of the socially optimal welfare (utility) over the welfare (utility) under a selfish routeing equilibrium, which measures the inefficiency of non-cooperation and non-coordination.

We study a simple queueing model where the firm is a leader in the Stackelberg game and anticipates customer behavior when choosing the capacity or price. We compare the system performance for several systems by changing the customer's objective (selfish or collective), the firm's objective (profit maximization or social), and the system set-ups (capacity setting versus price setting; observable queue and unobservable queue). Our main goal is to see whether there can be a Benefit of Anarchy (BoA) when the firm and customers act strategically in a Stackelberg game.

Received 14 July 2018; revision received 25 October 2019.
* Postal address: KAIST College of Business, Seoul, 02455, Republic of Korea. Email address: hakjin.chung@kaist.ac.kr
** Postal address: Ross School of Business, University of Michigan, Ann Arbor, MI 48109, USA. Email address: hsahn@umich.edu
*** Postal address: Department of Industrial Engineering and Operations Research, University of California, Berkeley, CA 94720, USA. Email address: rrighter@ieor.berkeley.edu



https://doi.org/10.1017/apr.2019.64 Published online by Cambridge University Press

There have been many papers that analyze customers' strategic behavior in queueing systems. The first paper in this area is the seminal paper by Naor [17], who studied customers' joining decisions after observing the queue length. Naor showed that selfish customers do not consider their negative externalities, so they over-join, resulting in suboptimal customer utilities. He also showed that a congestion toll can induce selfish customers to choose socially optimal decisions. Naor's work has been generalized to multi-server systems (e.g. Knudsen [12] and Yechiali [21]), generalized arrival processes (e.g. Yechiali [20]), state-dependent pricing (e.g. Borgs *et al.* [1]), and variable service rate (e.g. Lippman and Stidham [13]). The case in which customers do not observe the queue, and choose a joining probability, was studied by Edelson and Hildebrand [5], Littlechild [14], and Hassin [8]. Hassin and Haviv [10] and Hassin [9] have provided a comprehensive review of earlier work.

While these papers focused on customer behavior and the impact of prices, relatively few papers have studied how the firm's capacity impacts the customers' decision to join the queue. Grassmann [7] developed a numerical method to find the welfare-maximizing service rate assuming the customers' behavior is given regardless of the service rate. De Vany [3], Mendelson [16], Dewan and Mendelson [4], Stidham [19], and Chen and Frank [2] studied the short-run pricing and the long-run capacity decisions together. However, these works on the capacity problem assumed an unobservable queue and did not consider the case where customers' decisions are state-dependent, while De Vany [3] considered a state-dependent decision, but the optimization problem is not explicitly solved. Our base model considers the case where the server chooses the service rate when customers decide whether or not to join after observing the queue length.

The literature on the Price of Anarchy (Roughgarden [18] and Gilboa-Freedman *et al.* [6]) focused on the differences in outcomes when there are selfish (individually optimizing) customers versus when they are coordinated (collectively optimizing) in routeing systems. The PoA is defined as the ratio of the collective and selfish utilities (the average utility of customers who are coordinating to maximize their utility over the average utility of individual customers who selfishly maximize their own utility), and it must be greater than 1 for a fixed service rate. These works assumed that the service capacity is an infrastructure instead of a decision. The main point of departure is that we view the server as a decision maker with its own objective and able to choose the service rate or price. Our results show that when the firm strategically chooses the service rate as a leader of a Stackelberg game, selfish customers, who are more likely to join than collective customers, may induce the firm to set a higher service rate than collective customers, resulting in PoA $< 1$. In such cases, selfishness is not necessarily harmful, thus anarchy can be beneficial (BoA).

In order to determine whether the Benefit of Anarchy is a general phenomenon, we study several variants of our model, in which the firm chooses the price rather than the capacity, and in which the queue may be unobservable to customers rather than observable. When the server chooses the price (and the capacity is fixed) and the queue is observable, building on the work of Yechiali [20], we show that again there may be a Benefit of Anarchy when the firm is selfish (profit-maximizing). Of course, when the social (welfare-maximizing) firm chooses the price, anarchy is always costly regardless of whether the queue is observable. When the selfish firm sets either the price or the capacity, we show that there exists no Benefit of Anarchy when the queue is unobservable. Indeed, in this case overall customer utility is zero when they act selfishly, while the utility is positive when they act collectively. On the other hand, and in contrast to the price-setting social firm, when the social firm sets the capacity and the queue is unobservable, we identify cases when a BoA arises. Finally, we show that for both a

https://doi.org/10.1017/apr.2019.64 Published online by Cambridge University Press

profit-maximizing and welfare-maximizing firm that sets the capacity, customer anarchy can be socially beneficial.

The remainder of this paper is organized as follows. Section 2 outlines our model. We study the customers' individual and collective problems in Section 3. In Section 4 we consider the firm's capacity problem as a Stackelberg leader. With an observable queue, the customer uses a discrete (threshold) strategy, and thus, as a result, profit and welfare are discontinuous in the service rate. We derive the mathematical properties of the profit (welfare function) and use them to compute the optimal policy efficiently. In Section 5 we compare the equilibrium outcomes for collective and selfish customers. The extensions of our base model are studied in Sections 6 and 7. All proofs are provided in the appendices.

Our major findings, when the server sets the capacity in an observable queue, are as follows.

- When a profit-maximizing firm sets the capacity, a *Benefit of Anarchy* (BoA) can arise if the firm under-serves collective customers relative to selfish customers. When a BoA with the selfish firm arises, the BoA is also socially beneficial as it generates higher social welfare (Observation 1). When this is the case, we say there is a *Social Benefit of Anarchy* (SBoA).
- As one parameter changes while the others are held fixed, there are alternating intervals in which anarchy is beneficial and in which anarchy is harmful (Observation 2).
- When a welfare-maximizing firm sets the capacity, a BoA can arise. We find that anarchy is socially beneficial at a sufficiently high price (Proposition 3).

For the models in which the server sets the price instead of the capacity, or the queue is unobservable, we find the following.

- When the queue is observable, a BoA arises if the profit-maximizing firm sets a higher price for collective customers than for selfish customers (Observation 3). As with the capacity case, when there is a BoA, it is socially beneficial.
- When the queue is unobservable and the firm maximizes its profit, anarchy is always costly for customers regardless of whether the firm sets the capacity or the price (Propositions 5 and 7).
- When the queue is unobservable and the firm sets the capacity to maximize social welfare, a BoA arises if and only if the firm chooses to serve selfish customers but does not find it optimal to serve collective customers. When the firm serves collective customers at an equilibrium, anarchy is always costly for the customers (Proposition 6).
- When the queue is unobservable and the firm sets the capacity, we identify the condition under which there is an SBoA. When the selfish firm sets the price and the queue is unobservable, there is always an SBoA (Proposition 7).

## 2. Model description

In our basic model we consider a system where a firm determines the capacity (service rate) for serving delay-sensitive customers who decide whether or not to join after seeing the queue length. We model this with an observable M/M/1 queue with arrival rate $\lambda$ and service rate $\mu$. We define the traffic intensity as $\rho = \lambda/\mu$. Because customers balk if delay is too excessive,

https://doi.org/10.1017/apr.2019.64 Published online by Cambridge University Press

we have no stability issues even when $\rho > 1$. Upon arrival, the customer then decides whether or not to join the queue based on her specific objective (i.e. selfish or collective utility, which will be discussed later in detail). We assume that all customers are risk-neutral and identical in their willingness-to-pay and disutility for waiting. Let $R$ represent the customer surplus from a service, and $p$ the fee for the service. If she joins the system, the customer incurs a disutility for waiting which we model as a linear holding cost with the cost per unit time, $h$. If she leaves without being served, there is no gain or loss, yielding zero utility. Anticipating the customer's action, the firm sets the service rate to maximize its objective (profit or social welfare) as a leader in a Stackelberg game. We consider the case where the firm is the first mover because the firm makes infrastructure decisions that have long-term ramifications, while customers (i.e. followers) can simply change their decisions upon arrival.

## 3. The customers' problem

We first describe the customers' decisions given the service rate, $\mu$, and price, $p$, fixed. Following Naor [17], a customer's optimal strategy can be specified as a threshold strategy. That is, there exists a threshold $n$ such that a customer joins the system if and only if she observes fewer than $n$ customers upon arrival. To measure the impact of customers' selfishness, we consider two different types of customers: (1) *selfish customers* who make decisions that maximize their own individual utility, and (2) *collective customers* who want to maximize the average customer utility.

### 3.1. Selfish customers

For the selfish customers' joining strategy, we invoke Naor's result. The expected utility that an arriving customer draws from joining when there are $i$ customers in the system is

$$U(i) = R - p - h\frac{(i+1)}{\mu}.$$

Since not joining gives her zero utility, the selfish customer joins the system as long as her expected utility is non-negative. We assume that a customer joins the queue if she is indifferent. Then selfish customers join the queue if and only if the number of customers present is strictly smaller than the *selfish threshold*, $n_S(\mu)$, which is given by

$$n_S(\mu) = \max\left\{n \in \mathbb{N}_0 \mid R - p - h\frac{(n-1)+1}{\mu} \geq 0\right\} = \left\lfloor \frac{R-p}{h}\mu \right\rfloor,$$

where $\lfloor \cdot \rfloor$ indicates the largest integer not exceeding the value inside the bracket.

### 3.2. Collective customers

A collective customer considers not only her own utility but also the utility of the other customers. In this case, customers are coordinated and follow the same decision rule that maximizes the average utility of all customers. The average utility depends on how many customers receive service and how long they wait. Since the individual utility from joining decreases in the queue length, it can be shown that collective customers' optimal strategy is also characterized by a threshold.

When the threshold is $n$, the resulting system is an M/M/1/$n$ queue in which the maximum number of customers in the system is bounded by $n$. The average customers' utility is

$$\mathcal{U}(\mu, n) = (R-p)\lambda_e(\mu, n) - hL(\mu, n), \tag{1}$$

https://doi.org/10.1017/apr.2019.64 Published online by Cambridge University Press

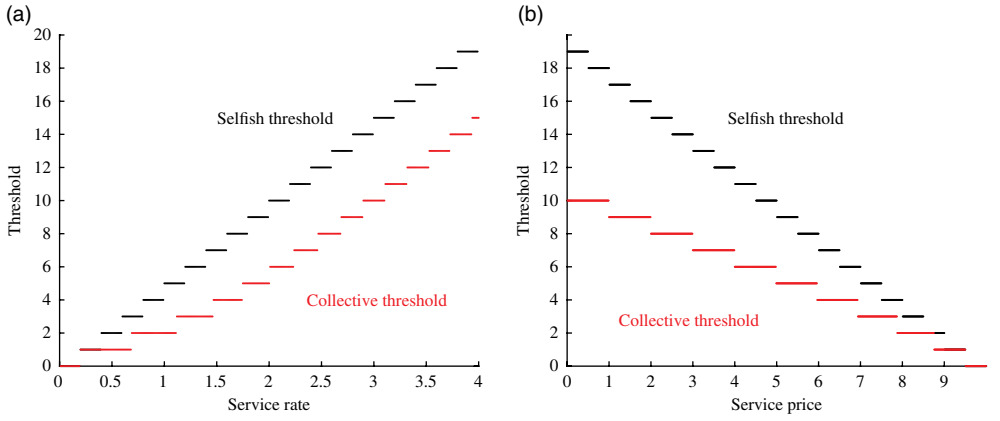



FIGURE 1: Threshold strategy of selfish and collective customers, respectively, as a function of $\mu$ when $p = 5$ (a) and a function of $p$ when $\mu = 2$ (b); other parameters are given by $\lambda = 1$, $R = 10$, and $h = 1$.

where

$$\lambda_e(\mu, n) = \lambda \cdot \frac{1-\rho^n}{1-\rho^{n+1}}$$

is the effective arrival rate (for those who join), and

$$L(\mu, n) = \frac{\rho}{1-\rho} - \frac{(n+1)\rho^{n+1}}{1-\rho^{n+1}}$$

is the average number of customers in the system. The optimal threshold for collective customers (which we call the *collective threshold*), $n_C(\mu)$, maximizes (1) and balances the throughput and the waiting cost.

Naor [17] derived the welfare-maximizing threshold when the welfare is given by $R\lambda_e(\mu, n) - hL(\mu, n)$. Then, by modifying Naor [17], the collective threshold is (see Figure 1)

$$n_C(\mu) = \max\left\{n \in \mathbb{N}_0 \mid \frac{n(1-\rho) - \rho(1-\rho^n)}{\mu(1-\rho)^2} \le \frac{R-p}{h}\right\}. \tag{2}$$

### 3.3. Price of Anarchy for a given service rate

Prior to analyzing the equilibrium service rate and pricing, we compare the thresholds for the two customer types.

**Proposition 1.** *For any service rate $\mu > 0$ and price $p \ge 0$, the collective threshold is not greater than the selfish threshold, i.e. $n_C(\mu) \le n_S(\mu)$. That is, collective customers achieve a higher average utility than selfish customers by not joining the queue in congested states where selfish customers would join.*

This result is immediate from those of Naor [17] and Yechiali [20]. As selfish customers do not consider their negative externalities, the extra congestion due to selfish customers lowers overall utility for everyone who joins the queue. Hence, the Price of Anarchy (PoA), defined as the ratio of $\mathcal{U}(\mu, n_C(\mu))$ (collective customers' utility) to $\mathcal{U}(\mu, n_S(\mu))$ (selfish customers' utility), is always greater than or equal to 1 for a fixed $\mu$, that is,

$$\frac{\mathcal{U}(\mu, n_C(\mu))}{\mathcal{U}(\mu, n_S(\mu))} \ge 1.$$

https://doi.org/10.1017/apr.2019.64 Published online by Cambridge University Press

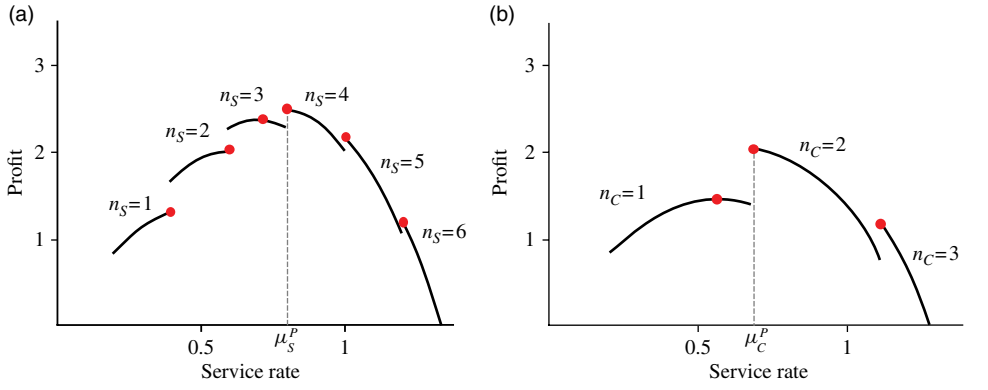



FIGURE 2: Profit as a function of $\mu$ with selfish customers (a) and collective customers (b); $\lambda = 1, R = 10$, $p = 5$, $h = 1$, and $c(\mu) = 2\mu^3$.

However, upon a close examination, the result does not extend to the *social welfare*. Note that, for the same service rate, the revenue increases if more customers join the queue. Thus Proposition 1 implies that the firm earns a smaller profit with collective customers than with selfish customers for the same service rate. Because of this, the social welfare (the sum of customer utility and firm's profit), can be smaller with collective customers although the average utility of customer is higher.

## 4. The firm's capacity problem

Now we consider the firm's capacity decision. Knowing the customer behavior and the price, the firm chooses the service rate to maximize its objective. The *selfish firm* makes a decision that maximizes its profit, and the *social firm* maximizes social welfare, the sum of customer utility, and its profit. To operate the system at $\mu$, the operating cost $c(\mu)$ is incurred; we assume $c(\mu)$ is convex increasing in $\mu$. We define the firm's optimal service rate as $\mu_\alpha^\beta$, where $\alpha = C$ or $S$ for collective and selfish customers respectively, and $\beta = P$ or $W$ for the profit-maximizing or the welfare-maximizing firm.

### 4.1. The selfish firm

From the fact that the optimal strategy of customers is characterized by a threshold, the firm's average profit when the service rate is $\mu$ is given by

$$\Pi(\mu, n_\alpha(\mu)) = p\lambda_e(\mu, n_\alpha(\mu)) - c(\mu) \quad \text{for } \alpha \in \{S, C\}.$$

Note that both the effective arrival rate, $\lambda_e(\mu, n_\alpha(\mu))$, and the operating cost, $c(\mu)$, are increasing in $\mu$. Thus the optimal service rate for the profit-maximizing firm, $\mu_\alpha^P$, should balance the trade-off between the revenue and the operating cost. However, as the customer threshold is a step function in $\mu$ (Figure 1), the profit function is not smooth in $\mu$, as illustrated in Figure 2.

Let $\underline{\mu}_\alpha(n)$ denote the smallest $\mu$ that induces the threshold of $\alpha$-type customers to be $n$. These define *threshold intervals* of $\mu$, i.e. $n_\alpha(\mu) = n$ for any $\mu \in [\underline{\mu}_\alpha(n), \underline{\mu}_\alpha(n+1))$. The following lemma summarizes the properties of the profit function.

https://doi.org/10.1017/apr.2019.64 Published online by Cambridge University Press

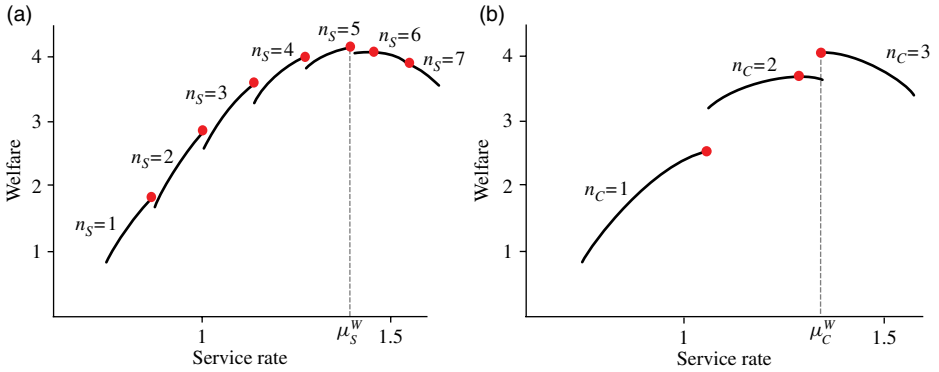



FIGURE 3: Welfare as a function of $\mu$ with selfish customers (a) and collective customers (b); $\lambda = 1$, $R = 10$, $p = 5$, $h = 1$, and $c(\mu) = 2\mu^2$

**Lemma 1.** *For each customer type $\alpha \in \{S, C\}$:*

(i) $\Pi(\mu, n_\alpha(\mu))$ *is concave in $\mu$ within each threshold interval.*

(ii) $\Pi(\mu, n_\alpha(\mu))$ *is discontinuous and jumps upwards when the threshold changes.*

The first concavity result enables us to find the local maximizer for each threshold. However, the second result (it is discontinuous) complicates the capacity problem. When the threshold changes from $n-1$ to $n$ at $\mu = \underline{\mu}_\alpha(n)$, the effective arrival rate, and hence the profit function, jumps upwards. In order to solve for the profit-maximizing service rate, one may need to exhaustively search for the optimal $\mu$ for each threshold interval and then compare them. Figure 2 illustrates the difficulty of finding the optimal service rate. We further develop an efficient method for determining $\mu_\alpha^P$, $\alpha \in \{C, S\}$, by utilizing the upper envelope function; the details are given in Appendix A.

### 4.2. The social firm

We now consider a social firm, for example a government agency, that wants to maximize social welfare with its choice of service rate (see Figure 3). The average social welfare is given by

$$\mathcal{W}(\mu, n_\alpha(\mu)) = R\lambda_e(\mu, n_\alpha(\mu)) - hL(\mu, n_\alpha(\mu)) - c(\mu) \quad \text{for } \alpha \in \{S, C\}.$$

Now, in addition to the discontinuous effective arrival rate that we had in the profit function, the waiting cost is also discontinuous, making the analysis even more difficult. The following lemma characterizes the properties of the welfare function.

**Lemma 2.** *Given customer type $\alpha \in \{S, C\}$:*

(i) *Within each threshold interval, the marginal welfare is greater than the marginal profit in $\mu$, i.e.* $\mathrm{d}\mathcal{W}(\mu, n_\alpha(\mu))/\mathrm{d}\mu > \mathrm{d}\Pi(\mu, n_\alpha(\mu))/\mathrm{d}\mu$ *for* $\underline{\mu}_\alpha(n) \le \mu < \underline{\mu}_\alpha(n+1)$.

(ii) *For $\mu > \lambda$, $\mathcal{W}(\mu, n_\alpha(\mu))$ is concave in $\mu$ within each threshold interval.*

(iii) *The social welfare with selfish customers, $\mathcal{W}(\mu, n_S(\mu))$, is discontinuous and jumps either upwards or downwards at $\underline{\mu}_S(n)$ for all $n$. Specifically, there exist $\mu_L$ and $\mu_H$*

https://doi.org/10.1017/apr.2019.64 Published online by Cambridge University Press

*such that the welfare jumps downwards at* $\underline{\mu}_S(n) \in (\mu_L, \mu_H)$ *and upwards at* $\underline{\mu}_S(n) \notin (\mu_L, \mu_H)$.

(iv) *The social welfare with collective customers,* $\mathcal{W}(\mu, n_C(\mu))$, *is discontinuous and jumps upwards at* $\underline{\mu}_S(n)$ *for all n.*

Lemma 2(i) states that the marginal welfare is greater than the marginal profit within each threshold interval because the customer utility is always increasing in $\mu$ within the threshold interval. Lemma 2(ii) states the concavity of welfare within a threshold interval, which is analogous to the concavity result in the profit function (Lemma 1(i)). However, the concavity can be guaranteed only when $\mu > \lambda$. This is because, if $\mu < \lambda$, the average number of customers can be concave in $\mu$, which in turn implies that the welfare function can be convex. This non-concavity result makes the determination of the optimal service rate difficult.

Lemma 2(iii, iv) highlights how the welfare functions for selfish and collective customers differ. From Lemma 1(i) we know that the profit function is discontinuous and jumps upwards because the throughput (and revenue) increases. The welfare is also discontinuous, but the welfare can move upwards or downwards at a threshold for selfish customers. This is because the selfish customer's average utility may jump upwards or downwards as the threshold increases, depending on whether the additional revenue (from having more customers with a higher threshold) is greater than the increase in the average waiting cost with a higher threshold. In contrast, collective customers increase their threshold if and only if it increases their utility. Therefore, because the profit function also jumps upwards at $\underline{\mu}_C(n)$ for all $n$, the welfare jumps upwards when the threshold increases.

Due to the irregular discontinuities of the welfare function, the only way to find the socially optimal service rate is to exhaustively search for the local optimal $\mu$ for each threshold interval and then compare them. If the service rate is too high, the welfare is negative. Thus we need to search the threshold intervals for $\{\mu | c(\mu) < \lambda R\}$, and compare the local maximizers of the threshold intervals to find the socially optimal service rate, $\mu_\alpha^W$, for $\alpha \in \{S, C\}$.

## 5. The effect of selfishness

We now compare the equilibrium outcomes and examine the effects of customers' selfishness (which we also call *anarchy*) and the firm's selfishness. In Section 3 we saw that collective customers draw a higher average utility than selfish customers for given $\mu$, but in Section 4 we showed that the firm can choose a significantly different service rate depending on whether customers are selfish or collective. We first examine whether selfish customers draw a higher average utility (*Benefit of Anarchy*) and result in a higher welfare for the system (*Social Benefit of Anarchy*).

### 5.1. On customers' selfishness with the selfish firm

**Observation 1.** For the selfish (profit-maximizing) firm:

(i) The utility of selfish customers can be larger than that of collective customers if $\mu_S^P > \mu_C^P$. In other words there can be a Benefit of Anarchy when the firm under-serves collective customers compared to selfish customers.

(ii) If there is a Benefit of Anarchy, the welfare with selfish customers is higher than that with collective customers, that is, a Social Benefit of Anarchy also exists.

https://doi.org/10.1017/apr.2019.64 Published online by Cambridge University Press

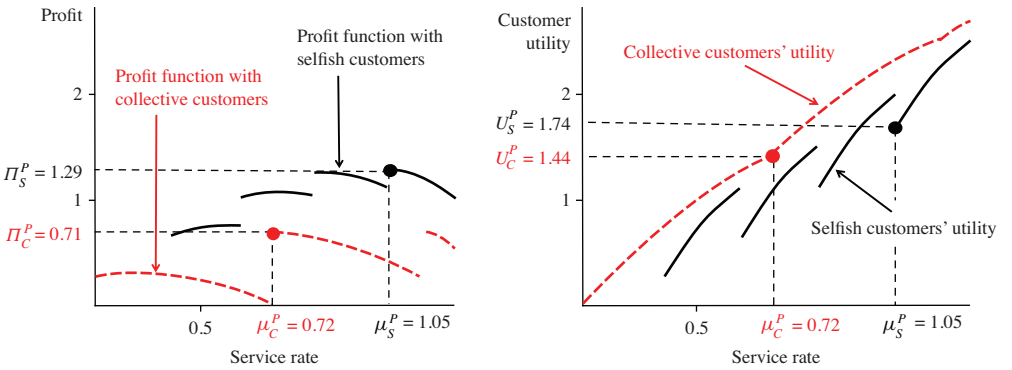



FIGURE 4: Benefit of Anarchy when $\lambda = 1$, $R = 10$, $p = 5.2$, $h = 1$, and $c(\mu) = 3\mu$.

Observation 1 provides a necessary condition for BoA and SBoA. Recall that collective customers draw a higher utility by joining less than selfish customers. From the profit perspective, this means smaller revenue will be earned from collective customers. As a response, the selfish firm chooses a different service rate for collective customers. When the firm chooses a lower service rate for collective customers, the slow service can lower the utility of collective customers below that of selfish customers, resulting in PoA $< 1$ as seen in Figure 4 (PoA 0.826). The selfish firm sets a lower service rate for collective customers than for selfish customers (Figure 4(a)). Consequently, the utility of collective customers is lower (Figure 4(b)): in this case collective customers earn 17.4% less overall utility than selfish customers. This contrasts with the result in the PoA literature (Roughgarden [18], Haviv and Roughgarden [11], and Gilboa-Freedman *et al.* [6]), where the PoA, the ratio of collective customers' utility to selfish customers' utility, must be greater than or equal to 1 when all parameters are fixed.

Observation 1(ii) further states that a BoA is also beneficial to society, that is, it increases the welfare. Note that the selfish firm always earns a higher profit from selfish customers than from collective customers: it always benefits from anarchy. Hence, when the utility of selfish customers is higher, the welfare is also higher with selfish customers, as seen in Figure 4.

The notion of a Benefit of Anarchy is intriguing, but it is difficult to give an analytic characterization of exactly when it happens. As shown in Section 4, the profit function is discontinuous, and we must search for a local maximizer of each threshold and compare them to find the optimal service rate. Indeed, as Figure 5 shows, the regions in which a BoA exists are disjoint in the problem parameters. Thus we have Observation 2, which means simple structural results are not possible.

**Observation 2.** The regions (in the parameter space) where a BoA or SBoA occurs cannot be characterized by simple thresholds on the parameters $h$, $\bar{c}$ (for $c(\mu) = \bar{c}\mu^2$), $R$, and $p$.

### 5.2. On customers' selfishness with the social firm

Now we consider the social firm. We first study how the social firm differs from the selfish firm given the same type of customers. This will help us understand the social efficiency loss when public service becomes privatized.

From the fact that the social firm generates greater welfare while the selfish firm earns a higher profit, we know that customer utility is always higher with the social firm. The following proposition shows how the social firm's service rate differs from that of the selfish firm.

https://doi.org/10.1017/apr.2019.64 Published online by Cambridge University Press

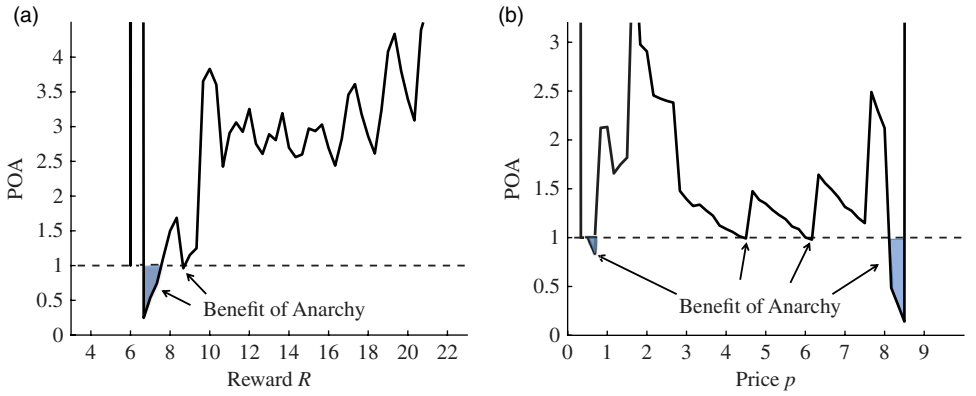



FIGURE 5: Irregular patterns of PoA; $\lambda = 1$ and $h = 1$ for (a) $p = 3$, $c(\mu) = 3\mu^2$, (b) $R = 10$, $c(\mu) = 2\mu^2$.

**Proposition 2.**

(i) *With collective customers, the social firm's service rate is at least as large as that of the selfish firm, i.e.* $\mu_C^P \leq \mu_C^W$. *Thus the social firm always serves more collective customers than the selfish firm.*

(ii) *With selfish customers, the social firm's service rate can be either larger or smaller than that of the selfish firm.*

Proposition 2(i) comes from the fact that collective customers draw a higher utility as $\mu$ increases (Figure 4). Thus the social firm sets a higher service rate than the selfish firm (which underinvests). As a result, the social firm improves the customer utility and therefore the welfare.

Surprisingly, Proposition 2(ii) shows that the result does not extend to selfish customers. To see why, note that the utility of selfish customers does not always increase in the service rate (Figure 4). However, congestion increases as the threshold increases. Thus, with selfish customers, the social firm may choose a lower service rate than the selfish firm just to reduce congestion.

Our next result summarizes when the customers' selfishness is beneficial for the social firm.

**Proposition 3.** *For the social (welfare-maximizing) firm:*

(i) *There can be a Benefit of Anarchy if* $\mu_S^W > \mu_C^W$.

(ii) *There is a Social Benefit of Anarchy when the price is sufficiently high.*

Although Proposition 3(i) is similar to Observation 1(i), the mechanism is quite different. For the selfish firm, the BoA occurs when collective customers are under-served. When customers are selfish and incur excessive waiting costs, the social firm may choose a higher service rate than the profit-maximizing firm or the social firm serving collective customers. In such a case, a BoA can arise by serving more.

Proposition 3(ii) states that the customers' selfishness is beneficial to society when the price is high. Suppose that the price is zero (or sufficiently low) so the firm earns little or no revenue. In such a case, the collective customers' behavior, which maximizes their overall utility,

https://doi.org/10.1017/apr.2019.64 Published online by Cambridge University Press

becomes socially optimal, maximizing the welfare. When the price is zero, the social firm and collective customers together achieve the first-best welfare. As price increases, the action of collective customers (i.e. joining less than selfish customers) can decrease the throughput below the socially optimal level, resulting in a decrease in welfare. In the case of selfish customers, a high price also makes them join less, but they still join more than collective customers. Therefore, when the price is sufficiently large, the social firm can achieve a higher welfare with selfish customers (through increased revenue) than with collective customers.

## 6. Pricing model with an observable queue

In this section we consider a service system where the firm chooses the price $p$ at time 0, assuming $\mu$ is given. We define the firm's optimal price, depending on its objective and the customer types, as $p_\alpha^\beta$, where $\alpha \in \{C, S\}$ and $\beta \in \{P, W\}$, as before. We start with the easier case when the firm is social. In contrast to the social capacity-setting firm, the social price-setting firm is able to maximize social welfare.

**Proposition 4.** *For the price-setting social firm:*

(i) *The social firm sets a non-zero congestion toll for selfish customers,* $p_S^W > 0$, *and it sets* $p_C^W = 0$ *for collective customers. Thus there exists no BoA.*

(ii) *The social firm is always able to maximize the welfare.*

Naor [17] showed that a congestion toll can control customer selfishness and induces socially optimal behavior. Now consider the collective customers who maximize their utility. As we discussed with Proposition 3, the social firm no longer needs to charge a congestion toll (i.e. the zero price is socially optimal). Since collective customers draw a higher utility than selfish customers for a given price, the utility of collective customers is always higher than that of selfish customers. Hence anarchy is always costly to customers when the social firm sets the price.

**Observation 3.** For the price-setting selfish firm:

(i) There can be a BoA when the firm sets a higher price for collective customers than for selfish customers, i.e. $p_S^P < p_C^P$.

(ii) When there is a BoA, there is also an SBoA.

Contrary to the social firm's case, there can be a Benefit of Anarchy when the firm is selfish. From the profit perspective, collective customers join less and are always less preferred than selfish customers. To compensate this, the profit-maximizing firm can set a very high price for collective customers. If the price is sufficiently high, the utility of collective customers can be lower than that of selfish customers, resulting in the BoA. This is analogous to the case where the firm sets a significantly lower service rate for collective customers. In addition, if the utility of selfish customers is higher than that of collective customers, the welfare is also higher with selfish customers since the selfish firm always earns a higher profit with selfish customers. Hence a BoA is socially beneficial. Figure 6 presents an example where PoA $= 0.758$, so the collective customer utility is 24% lower and the anarchy is a win–win for both the firm and the customers.

https://doi.org/10.1017/apr.2019.64 Published online by Cambridge University Press

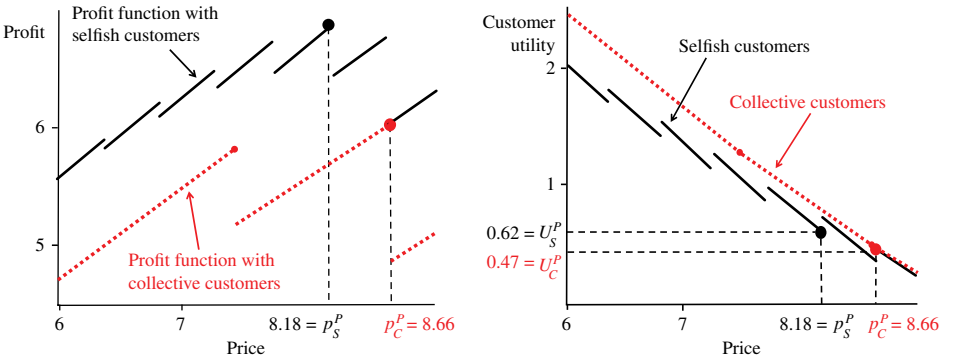



FIGURE 6: (Social) Benefit of Anarchy when $\lambda = 1$, $R = 10$, $h = 0.5$, and $\mu = 1.1$.

## 7. Unobservable queue

In this section we summarize and expand on the results of Maxey *et al.* [15] for the case where customers cannot observe the queue length. Hassin [8] showed that, for the unobservable queue, selfish customers employ a mixed strategy such that they join with probability $q \in [0, 1]$. Using this result, we will focus on whether there is a (Social) Benefit of Anarchy when the firm chooses a service rate (capacity) or price. As before, we define $\mu_\alpha^\beta$ and $p_\alpha^\beta$ for the optimal capacity and price depending on the customer and firm type, $\alpha \in \{C, S\}$ and $\beta \in \{P, W\}$.

### 7.1. Capacity model

Here we consider the capacity model, where the firm chooses a service rate given a fixed price. We first examine how the customer's strategy depends on the service rate. Let $q_\alpha(\mu)$ denote the joining probability of customer type $\alpha \in \{S, C\}$ when the service rate is $\mu$, so the effective arrival rate is $\lambda \cdot q_\alpha(\mu)$. Then the overall customer utility is

$$U_\alpha(\mu) := \lambda q_\alpha(\mu)\left(R - p - \frac{h}{\mu - \lambda q_\alpha(\mu)}\right), \tag{3}$$

as long as $\mu > \lambda q_\alpha(\mu)$. Note that the expected individual utility of joining is $R - p - h/(\mu - \lambda q)$ if all other customers join with probability $q$. Hassin [8] showed that the optimal joining probability of selfish customers, $q_S(\mu)$, is $q$ such that $R - p - h/(\mu - \lambda q) = 0$ unless $q$ is bounded at 0 or 1. On the other hand, collective customers in our model choose $q_C(\mu)$ to maximize (3). The following lemma and Figure 7 summarize the customers' optimal strategy. In preparation, we define $b := (R - p)/h$.

**Lemma 3.**

(i) *For a given price p, selfish customers are more likely to join the queue than collective customers for any $\mu$, i.e. $q_S(\mu) \geq q_C(\mu)$ for any $\mu > 0$.*

(ii) *For each type $\alpha \in \{S, C\}$, the joining probability $q_\alpha(\mu)$ is monotonically increasing in $\mu$. Specifically, there exists a pair of thresholds, $\mu_\alpha^0$ and $\mu_\alpha^1$, such that $q_\alpha(\mu) = 0$ for $\mu \leq \mu_\alpha^0$ and $q_\alpha(\mu) = 1$ for $\mu \geq \mu_\alpha^1$, where $\mu_S^0 = \mu_C^0 = 1/b$, $\mu_S^1 = 1/b + \lambda$, and $\mu_C^1 = \lambda + (1 + \sqrt{1 + 4b\lambda})/2b$.*

https://doi.org/10.1017/apr.2019.64 Published online by Cambridge University Press

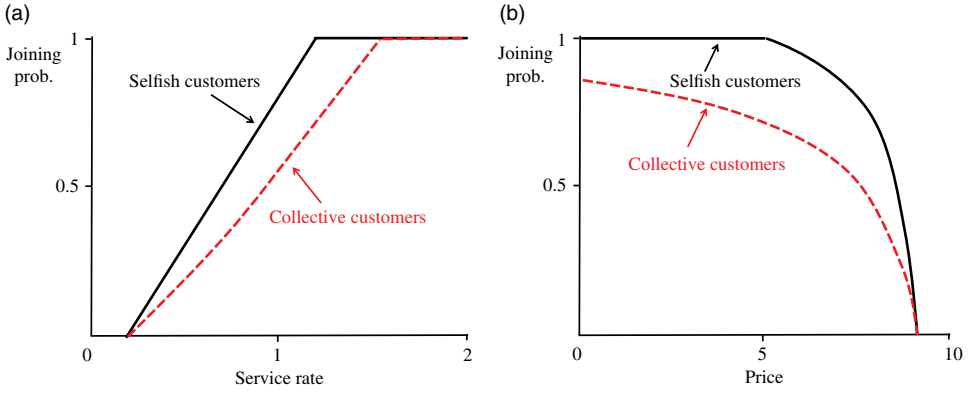



FIGURE 7: Joining probability of selfish and collective customers, respectively, as a function of $\mu$ when $p = 5$ (a) and a function of $p$ when $\mu = 1.2$ (b); other parameters are given by $\lambda = 1$, $R = 10$, and $h = 1$.

*For intermediate values of* $\mu$ ($\mu_\alpha^0 < \mu < \mu_\alpha^1$), *the joining probabilities are given by*

$$q_S(\mu) = \frac{\mu - 1/b}{\lambda} \quad \text{and} \quad q_C(\mu) = \frac{\mu - \sqrt{\mu/b}}{\lambda},$$

*respectively.*

Now consider the selfish firm's problem of choosing the profit-maximizing service rate $\mu_\alpha^P$. For simplicity, we assume a linear operating cost, i.e. $c(\mu) = \bar{c} \cdot \mu$ for some $\bar{c} > 0$. Then the firm's profit is given by $\Pi_\alpha(\mu) := \lambda p q_\alpha(\mu) - \bar{c}\mu$, where we assume $p > \bar{c}$; otherwise the firm chooses $\mu_\alpha^P = 0$ for all values of $R$.

**Lemma 4.** *For a given price* $p > \bar{c}$, *the selfish firm chooses the smallest service rate at which all customers join as long as the resultant profit is positive. Specifically, for each type* $\alpha \in \{S, C\}$, $\mu_\alpha^P = \mu_\alpha^1$ *if* $\lambda p/\bar{c} \geq \mu_\alpha^1$, *or* $\mu_\alpha^P = 0$ *otherwise.*

To see why, consider the three different ranges of $\mu$. For any $\mu \leq \mu_\alpha^0$, no customers join the queue, so the selfish firm will not operate ($\mu = 0$). For any $\mu \geq \mu_\alpha^1$, all customers join, and the firm will choose the smallest service rate, $\mu_\alpha^1$. Lastly, for $\mu_\alpha^0 < \mu < \mu_\alpha^1$, it can be shown that the profit function is linear (with selfish customers) or convex (with collective customers). Combining these results, the selfish firm chooses the smallest service rate at which all customers join ($\mu_\alpha^P = \mu_\alpha^1$) if the resultant profit is positive; otherwise, it simply closes the system ($\mu_\alpha^P = 0$).

**Proposition 5.** *For the selfish firm choosing the service rate:*

(i) *There is no Benefit of Anarchy and PoA* $= \infty$.

(ii) *There exists a Social Benefit of Anarchy if and only if* $\mu_S^1 \leq \lambda p/\bar{c} < \mu_C^1$.

Note that the utility of selfish customers is 0 as the selfish firm never chooses a service rate higher than needed to induce all customers to join, $\mu_S^P \leq \mu_S^1$ (Lemma 4), and selfish customers choose $q_S(\mu)$ so that they are indifferent between joining and not joining. On the other hand, collective customers achieve positive utility as long as the firm operates. Therefore no Benefit

https://doi.org/10.1017/apr.2019.64 Published online by Cambridge University Press

of Anarchy exists and PoA $= \infty$. However, a Social Benefit of Anarchy arises when the price to cost ratio falls between $\mu_S^1$ and $\mu_C^1$. In this case, the firm does not serve any collective customers ($\mu_C^P = 0$) while the firm still serves all selfish customers ($\mu_S^P = \mu_S^1$).

Now consider the social firm. The social welfare function is given by

$$\mathcal{W}_\alpha(\mu) := \lambda q_\alpha(\mu)\left(R - \frac{h}{\mu - \lambda q_\alpha(\mu)}\right) - \bar{c}\mu,$$

where $\mu > \lambda q_\alpha(\mu)$, for otherwise the resultant queue is non-stable and customers would not join.

**Lemma 5.** *For a given price $p$ and customer type $\alpha \in \{S, C\}$, the social firm chooses $\mu_\alpha^W = \max\{\bar{\mu}, \mu_\alpha^1\}$ as long as the resultant welfare is positive. Specifically, $\mu_\alpha^W = \bar{\mu} := \lambda + \sqrt{h\lambda/\bar{c}}$ if $\bar{c} \le h\lambda/(\mu_\alpha^1 - \lambda)^2$ and $\mathcal{W}_\alpha(\bar{\mu}) > 0$, $\mu_\alpha^W = \mu_\alpha^1$ if $\bar{c} > h\lambda/(\mu_\alpha^1 - \lambda)^2$ and $\mathcal{W}_\alpha(\mu_\alpha^1) > 0$, and $\mu_\alpha^W = \mathcal{W}_\alpha(\mu_\alpha^W) = 0$ otherwise.*

As in the selfish firm's case, for $\bar{\mu} \le \mu_\alpha^1$, the social firm's optimal service rate is either $\mu_\alpha^1$ or 0. The difference arises when $\bar{\mu} > \mu_\alpha^1$, in which case all customers join. While there is no reason for the selfish firm to serve faster than $\mu_\alpha^1$ since no more revenue can be collected, the social firm may do so to increase customer utility. In particular, the welfare function assuming $q = 1$ can be shown to be concave where the first-order condition is satisfied at $\bar{\mu} := \lambda + \sqrt{h\lambda/\bar{c}}$. Combining these results, the social firm's optimal service rate can be one of three, $\mu_\alpha^W = \bar{\mu}$, $\mu_\alpha^1$, or 0.

**Proposition 6.** *For the social firm choosing the service rate:*

(i) *If and only if $\mu_S^W = \bar{\mu} > \mu_C^W = 0$, PoA $= 0$ and the BoA exists. If $\mu_C^W > 0$, then $\mu_C^W > \mu_S^W$ and there cannot be any Benefit of Anarchy.*

(ii) *There exists a Social Benefit of Anarchy if and only if $\bar{\mu} < \mu_C^1$, i.e. $\bar{c} > 4b^2h\lambda/(1 + \sqrt{1 + 4b\lambda})^2$.*

Proposition 6(i) states that a BoA arises when the social firm only serves selfish customers. As long as the social firm serves both types of customers, it chooses a higher service rate for collective customers, so their utility is always higher than that of selfish customers. However, when the cost of service is so high that the social firm cannot achieve positive welfare by serving collective customers, the social firm will choose $\mu_C^W = 0$. In this case, if selfish customers draw positive utility from a service rate that can be provided at reasonable cost ($\mu_S^W = \bar{\mu} > \mu_S^1$), we have PoA $= 0$.

Proposition 6(ii) states that when the operating cost is high, the demand coordination of collective customers precludes the social firm from achieving the first-best welfare, which would be obtained for $\mu = \bar{\mu}$ and $q = 1$. When the cost is sufficiently high, $\bar{\mu}$ is low enough that not all collective customers join when $\mu = \bar{\mu}$. In this case, the collective behavior of customers harms the equilibrium welfare, so the anarchy is socially beneficial. Of course, the social firm might do better at maximizing social welfare by adjusting the price, which is considered next.

### 7.2. Pricing model

Now we consider the pricing model in which the firm chooses price $p$ given a fixed service rate. Following the results of Hassin [8] and Edelson and Hilderbrand [5], customers adopt a

https://doi.org/10.1017/apr.2019.64 Published online by Cambridge University Press

strategy where they join with probability $q_\alpha(p)$ for $\alpha \in \{S, C\}$, which is now a function of $p$. The expected customer utility is

$$U_\alpha(p) = \lambda q_\alpha(p)\left(R - p - \frac{h}{\mu - \lambda q_\alpha(p)}\right).$$

The optimal joining strategy and its dependence on the price are summarized in the following corollary of Lemma 3.

**Corollary 1.** *For each type $\alpha \in \{S, C\}$, the joining probability $q_\alpha(p)$ is monotonically decreasing in $p$. Specifically, there exist a pair of thresholds, $p_\alpha^1$ and $p_\alpha^0$, such that $q_\alpha(p) = 1$ for $p \le p_\alpha^1$ and $q_\alpha(p) = 0$ for $p \ge p_\alpha^0$, where*

$$p_S^0 = p_C^0 := R - \frac{h}{\mu}, \quad p_S^1 := \left[R - \frac{h}{\mu - \lambda}\right]^+ I\{\mu > \lambda\} \quad \text{and} \quad p_C^1 := \left[R - \frac{h\mu}{(\mu - \lambda)^2}\right]^+ I\{\mu > \lambda\}.$$

*For intermediate values of $p$ ($p_\alpha^1 < p < p_\alpha^0$), the joining probabilities are given by*

$$q_S(p) = \frac{1}{\lambda}\left(\mu - \frac{h}{R - p}\right) \quad \text{and} \quad q_C(p) = \frac{1}{\lambda}\left(\mu - \sqrt{\frac{h\mu}{R - p}}\right),$$

*respectively.*

We now consider the selfish firm that chooses $p_\alpha^P$ to maximize its profit, $\Pi_\alpha(p) = \lambda p q_\alpha(p) - \bar{c}\mu$. We assume $R > h/\mu$; otherwise $q_\alpha(p) = 0$ for any $p \ge 0$.

**Lemma 6.** *For given service rate $\mu$ and customer type $\alpha \in \{S, C\}$, the selfish firm chooses the price $p_\alpha^P = \max\{\bar{p}_\alpha, p_\alpha^1\}$, where $\bar{p}_S = R - \sqrt{hR/\mu}$ and $\bar{p}_C$ is the solution of the first-order condition, $h(2R - \bar{p}_C)^2 = 4\mu(R - \bar{p}_C)^3$.*

For $p_\alpha^1 < p < p_\alpha^0$, where customers join with probability $q_\alpha(p) \in (0, 1)$, the profit function is concave in $p$. Let $\bar{p}_\alpha$ denote the price that satisfies the first-order condition without assuming $q_\alpha(p)$ must be between [0,1]. If $\bar{p}_\alpha < p_\alpha^1$, then $p_\alpha^1$ must be optimal, because everyone joins at $p_\alpha^1$ and the firm gets a higher price. On the other hand, if the stationary point is greater than $p_\alpha^1$, the optimal price is $\bar{p}_\alpha$ because it satisfies both the first-order condition and the constraint that $q(p)$ can be no more than 1.

**Proposition 7.** *For the selfish firm setting the price:*

(i) *There is no Benefit of Anarchy and* $\text{PoA} = \infty$.

(ii) *There is always a Social Benefit of Anarchy, i.e. $\mathcal{W}_S^P \ge \mathcal{W}_C^P$.*

Note that selfish customers can achieve a positive utility only if $p < p_\alpha^1$. Since the selfish firm's price is always greater than $p_\alpha^1$, the utility of selfish customers is always zero. Hence, as Edelson and Hilderbrand [5] observed, the social welfare is entirely from the firm. Therefore, when the selfish firm sets the price, anarchy is always costly for customers while it is always beneficial for social welfare.

Now consider the social firm's problem. The social welfare is given by

$$\mathcal{W}_\alpha(p) := \lambda q_\alpha(p)\left(R - \frac{h}{\mu - \lambda q_\alpha(p)}\right) - \bar{c}\mu.$$

https://doi.org/10.1017/apr.2019.64 Published online by Cambridge University Press

We assume $\mu > \lambda q_\alpha(p)$; otherwise the resultant queue is not stable.

**Lemma 7.** *For given service rate $\mu$, the social firm chooses the price $p_S^W = R - \sqrt{hR/\mu}$ for selfish customers and $p_C^W = 0$ for collective customers. In both cases, the first-best welfare is achieved.*

As in the case of the observable queue, the social firm sets the price to zero for collective customers. On the other hand, when customers are selfish, Edelson and Hilderbrand [5] and Hassin [8] showed that the social firm sets a non-zero congestion toll. To see why, $p$ does not appear in the welfare function. Let $q^W$ denote the socially optimal joining probability and let us write the welfare as a function of $q$,

$$\mathcal{W}(q) := \lambda q\left(R - \frac{h}{\mu - \lambda q}\right) - \bar{c}\mu.$$

This can be shown to be strictly concave with a unique optimal $q^W$. Therefore the socially optimal price should induce $q^W$ from customers, which is $p_S^P = R - \sqrt{hR/\mu}$ for selfish customers and $p_C^W = 0$ for collective customers.

**Proposition 8.** *For the social firm setting the price:*

(i) *Anarchy is always costly for customers.*

(ii) *With the social firm, the first-best welfare is always achieved.*

Combining Propositions 4 and 8, we see that when the social firm chooses the price, the first-best welfare is achieved by setting a non-zero congestion toll for selfish customers and zero price for collective customers, regardless of whether customers can observe the queue length.

## 8. Conclusion

In this paper we show that anarchy can be beneficial to customers and to the overall social welfare. We summarize our findings in Table 1. As the table shows, the fact that anarchy can be beneficial is a general phenomenon not limited to the service-rate-setting queue. Except for a few special cases, the BoA can occur in a profit-maximizing or social firm or in an observable or unobservable queue. On the other hand, we find that anarchy can be socially beneficial in all settings.

Our results can be further applied to the cases when the firm sets both the price and the capacity simultaneously. Note that, when the queue is observable, anarchy can be beneficial only when the firm is selfish. In this case it can be shown than the social welfare is also higher with selfish customers. However, in any other cases, there is no Benefit of Anarchy. For instance, the social firm sets zero price for collective customers, so anarchy is always costly, regardless of whether the queue is observable or unobservable. Under the selfish firm with unobservable queue, the utility of selfish customers is always zero, so no Benefit of Anarchy exists in such a case.

## Appendix A. The upper-envelope method to find the profit-maximizing service rate

We solve the problem in Section 4.1 (the profit-maximizing service rate) more efficiently by utilizing the upper envelope function. For this, we first relax the constraint that the thresholds

https://doi.org/10.1017/apr.2019.64 Published online by Cambridge University Press

TABLE 1: Summary of findings.

| | Model | Firm type | Benefit of Anarchy | Social Benefit of Anarchy |
|---|---|---|---|---|
| observable queue | capacity | selfish | necessary condition: $\mu_S^P > \mu_C^P$ | sufficient condition: BoA exists |
| | control | social | necessary condition: $\mu_S^W > \mu_C^W$ | sufficient condition: $p$ sufficiently high |
| | price control | selfish | necessary condition: $p_S^P < p_C^P$ | sufficient condition: BoA exists |
| | | social | never ($p_C^W = 0$) | $\mathcal{W}_S^W = \mathcal{W}_C^W$ |
| unobservable queue | capacity | selfish | never ($\mathcal{U}_S^P = 0$) | iff condition: $\mu_S^1 \le \lambda p/\bar{c} < \mu_C^1$ |
| | control | social | iff condition: $\mu_S^W > \mu_C^W = 0$ | iff condition: $\bar{\mu} < \mu_C^1$ |
| | price control | selfish | never ($\mathcal{U}_S^P = 0$) | always |
| | | social | never ($p_C^W = 0$) | $\mathcal{W}_S^W = \mathcal{W}_C^W$ |

be integers and define relaxed thresholds as follows:

$$\tilde{n}_S(\mu) = b\mu \quad \text{and} \quad \tilde{n}_C(\mu) = \left\{ n \mid \frac{n(1-\rho) - \rho(1-\rho^n)}{\mu(1-\rho)^2} = b \right\}, \quad \text{where } b = \frac{R-p}{h}.$$

Then the profit function with the relaxed threshold (which we call the *relaxed profit function*) is given by

$$\Pi(\mu, \tilde{n}_\alpha(\mu)) = p \cdot \lambda \frac{1-\rho^{\tilde{n}_\alpha(\mu)}}{1-\rho^{\tilde{n}_\alpha(\mu)+1}} - c(\mu) \quad \text{for } \alpha \in \{S, C\}.$$

See Figure 8. The relaxed profit function is amenable to analysis. The following lemma gives its properties and an efficient means of determining the optimal service rate.

**Lemma 8.** *For each customer type $\alpha \in \{S, C\}$:*

(i) *$\Pi(\mu, \tilde{n}_\alpha(\mu))$ is differentiable and is an upper envelope of the profit function, that is, $\Pi(\mu, \tilde{n}_\alpha(\mu)) \ge \Pi(\mu, n_\alpha(\mu))$, where equality holds if and only if $\mu = \underline{\mu}_\alpha(n)$ for all $n$.*

(ii) *The optimal service rate, $\mu_\alpha^P$, exists in an interval that contains a maximizer of the relaxed profit function, that is, $\mu_\alpha^P \in [\underline{\mu}_\alpha(n_\alpha^0),\ \underline{\mu}_\alpha(n_\alpha^0 + 1)]$ for $n_\alpha^0 = n_\alpha(\tilde{\mu})$, where $\tilde{\mu}$ satisfies the first-order condition of $\Pi(\mu, \tilde{n}_\alpha(\mu))$.*

By using the relaxed profit function, we can specify intervals that contain the profit-maximizing service rate via first-order analysis. Then Lemma 8(ii) implies that we only need to search the intervals, $[\underline{\mu}_\alpha(n_\alpha^0),\ \underline{\mu}_\alpha(n_\alpha^0 + 1)]$, that contain stationary points of the envelope function. For each stationary point, the corresponding interval can be divided into two parts: the interval $[\underline{\mu}_\alpha(n_\alpha^0),\ \underline{\mu}_\alpha(n_\alpha^0 + 1))$ in which the threshold is $n_\alpha^0$, and the end point of $\underline{\mu}_\alpha(n_\alpha^0 + 1)$ with threshold $n_\alpha^0 + 1$. For the first sub-interval, the original profit function is concave, thus again we can find the maximum by first-order analysis.

https://doi.org/10.1017/apr.2019.64 Published online by Cambridge University Press

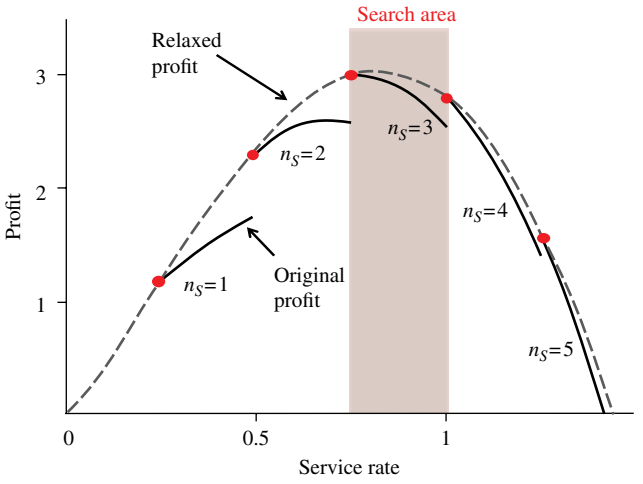



FIGURE 8: The original (solid line) and the relaxed (dashed line) profit functions when $\lambda = 1$, $R = 10$, $p = 6$, $h = 1$, and $c(\mu) = 2\mu^3$.

### A.1. Profit maximization with selfish customers

When customers are individually selfish, we can further show that the relaxed profit function is concave under a minor condition.

**Lemma 9.** *If customers are individually selfish and if the sufficient condition*

$$c\left(\frac{2}{b}\right) - c\left(\frac{1}{b}\right) \leq \frac{4bp}{(b+2)(b^2+2b+4)} \tag{4}$$

*is satisfied, then the equilibrium threshold is greater than* 1 *and the envelope function has a unique optimal solution, which is contained in the interval that also contains the optimal service rate for the original problem.*

From Lemma 9 we need to search only one interval, identified by the envelope method, if condition (4) holds. This is because the relaxed profit function is concave for values of $\mu$ such that the induced threshold is greater than 1, and condition (4) implies that the operating cost is not too big relative to the revenue, thus yielding an equilibrium threshold that is greater than 1.

The violation of condition (4) implies that the equilibrium threshold can be 1, which is a trivial case where no customer will join the system if a server is busy. Even when condition (4) is not satisfied, the envelope method is still efficient because the relaxed profit function is still unimodal for $\mu \geq 2/b$. In this case we need to search at most two intervals, $[1/b, 2/b]$ and $[n_\alpha^0/b, (n_\alpha^0+1)/b]$, for the maximizer of the relaxed profit function.

### A.2. Profit maximization with collective customers

When customers are collective, the concavity of the relaxed profit function cannot be shown analytically. One difficulty comes from the fact that a closed form for the threshold does not exist (see (2)) and we are not able to write the firm's profit explicitly as a function of $\mu$. Second, the collective threshold follows a convex pattern (see Figure 1), which, in turn, can make the relaxed profit function convex. Therefore, analytically there may exist multiple stationary points of the relaxed profit function, so, in principle, we need to examine multiple intervals to

https://doi.org/10.1017/apr.2019.64 Published online by Cambridge University Press

find the profit-maximizer. However, in most of our numerical experiments, the relaxed profit function is unimodal.

## Appendix B. Proofs

*Proof of Lemma 1 (profit function properties).* (i) The profit function for given $n$ is

$$\Pi(\mu, n) = p\lambda_e(\mu, n) - c(\mu),$$

where the effective arrival rate ($\lambda_e$) is given by

$$\lambda_e(\mu, n) = \begin{cases} \lambda \cdot \dfrac{1-\rho^n}{1-\rho^{n+1}} & \mu \neq \lambda, \\ \lambda \cdot \dfrac{n}{n+1} & \mu = \lambda. \end{cases}$$

For any $\mu \neq \lambda$, the effective arrival rate is increasing in $\mu$:

$$\begin{aligned} \frac{\partial \lambda_e(\mu, n)}{\partial \mu} &= (-\rho^2)\left(-\frac{\rho^{n-1}}{(1-\rho^{n+1})^2}\right)(\rho-1)\left(\rho\frac{\rho^n-1}{\rho-1} - n\right) \\ &= (-\rho^2)\left(-\frac{\rho^{n-1}}{(1-\rho^{n+1})^2}\right)(\rho-1)((\rho-1) + (\rho^2-1) + \cdots + (\rho^n-1)) \\ &\geq 0. \end{aligned}$$

The effective arrival rate is concave in $\mu$ for given $n$:

$$\begin{aligned} \frac{\partial^2 \lambda_e(\mu, n)}{\partial \mu^2} &= -\frac{(n+1)\rho^n\{n(1-\rho)(1+\rho^{n+1}) - 2\rho(1-\rho^n)\}}{\mu^2(1-\rho^{n+1})^3} \\ &= -\frac{(n+1)\rho^n(1-\rho)\{n(1+\rho^{n+1}) - 2(\rho+\rho^2+\cdots+\rho^n)\}}{\mu^2(1-\rho^{n+1})^3} \\ &\leq 0. \end{aligned} \tag{5}$$

Given the convex operating cost, we complete the proof.

(ii) The effective arrival rate is increasing in $n$ for given $\mu$:

$$\lambda_e(\mu, n+1) - \lambda_e(\mu, n) = \lambda \cdot \frac{1-\rho^{n+1}}{1-\rho^{n+2}} - \lambda \cdot \frac{1-\rho^n}{1-\rho^{n+1}} > 0.$$

This implies that the profit function jumps upwards when the threshold increases. □

*Proof of Lemma 2 (welfare function properties).* (i) This is trivially true from the fact that the difference between the welfare and the profit is the customer utility, which is increasing in $\mu$ within each threshold interval.

(ii) Given $n$ fixed, the effective arrival rate is concave from (5). Thus it is sufficient to show that the average number of customers, $L(\mu, n)$, is convex in $\mu$ for $\mu \geq \lambda$. That is, the second-order derivative with respect to $\mu$ is non-negative for $\mu > \lambda$ (i.e. $\rho < 1$):

$$\frac{\partial^2 L(\mu, n)}{\partial \mu^2} = \frac{2\rho}{\mu^2(1-\rho)^3} - \frac{(n+1)^2\rho^{n+1}((n+2) + n\rho^{n+1})}{\mu^2(1-\rho^{n+1})^3} \geq 0. \tag{6}$$

https://doi.org/10.1017/apr.2019.64 Published online by Cambridge University Press

We will first show that (6) is increasing in $n$ (for the moment we allow $n$ to be continuous) for any $\mu > \lambda$. That is, for any $n \geq 1$ and $\mu > \lambda$,

$$\begin{aligned}\frac{\partial^3 L(\mu, n)}{\partial \mu^2 \partial n} &= -\frac{(n+1)\rho^{n+1}}{\mu^2(1-\rho^{n+1})^4}[(n(n+1)\rho^{2n+2} + 4(n+1)^2\rho^{n+1} + (n+1)(n+2))\log\rho \\ &\quad - ((3n+1)\rho^{2n+2} + 4\rho^{n+1} - (3n+5))] \\ &\geq 0. \end{aligned} \tag{7}$$

This is equivalent to

$$\log\rho - \frac{(3n+1)\rho^{2n+2} + 4\rho^{n+1} - (3n+5)}{n(n+1)\rho^{2n+2} + 4(n+1)^2\rho^{n+1} + (n+1)(n+2)} \leq 0.$$

Letting $x = \rho^{n+1}$, this is equivalent to

$$k(x, n) := \log x - \frac{(3n+1)x^2 + 4x - (3n+5)}{nx^2 + 4(n+1)x + (n+2)} \leq 0.$$

The derivative of $k(x,n)$ with respect to $x$ is positive for any $x < 1$, that is,

$$\frac{(n^2(1-x) + 4n + 4)(1-x)^3}{x(nx^2 + 4(n+1)x + (n+2))^2} > 0,$$

and $k(1, n) = 0$. Thus (7) is true, and (6) is increasing in $n$ for any $\mu > \lambda$. Then, from

$$\frac{\partial^2 L(\mu, n)}{\partial \mu^2} \geq \frac{\partial^2 L(\mu, 1)}{\partial \mu^2} = \frac{2\rho}{\mu^2(1-\rho)^3} - \frac{4\rho^2(3+\rho^2)}{\mu^2(1-\rho^2)^3} = \frac{2\rho}{\mu^2(1+\rho)^3} > 0,$$

we complete the proof.

(iii) We first define the socially optimal threshold $n_W(\mu)$ that maximizes the social welfare:

$$n_W(\mu) = \max\left\{n \in \mathbb{N}_0 \mid \frac{n(1-\rho) - \rho(1-\rho^n)}{\mu(1-\rho)^2} \leq b_0\right\}, \quad \text{where } b_0 := \frac{R}{h}.$$

In this proof we relax the integer constraint on the thresholds, and use the relaxed thresholds, $\tilde{n}_S(\mu)$ and $\tilde{n}_W(\mu)$. We want to show that there exists an interval $(\mu_L,\ \mu_H)$ such that the relaxed selfish threshold is larger than the relaxed social threshold if and only if the service rate falls in this interval, i.e. $\tilde{n}_S(\mu) \geq \tilde{n}_W(\mu)$ if and only if $\mu \in (\mu_L,\ \mu_H)$. This is equivalent to the following:

$$\frac{\tilde{n}_S(\mu)(1-\rho) - \rho(1-\rho^{\tilde{n}_S(\mu)})}{\mu(1-\rho)^2} \geq b_0 \quad \text{if and only if } \mu \in (\mu_L,\ \mu_H). \tag{8}$$

After plugging in $\tilde{n}_S(\mu) = b\mu$ for $b := (R-p)/h$, (8) for $\rho \neq 1$ is equivalent to

$$g(\mu) := \lambda(b_0 - b)\left(\frac{1}{\rho} - 1\right)\left(\frac{1}{\rho} - \frac{b_0}{b_0 - b}\right) + (1 - \rho^{b\mu}) \leq 0 \quad \text{if and only if } \mu \in (\mu_L,\ \mu_H). \tag{9}$$

The first derivative of $g(\mu)$ with respect to $\mu$ is

$$g'(\mu) = (b_0 - b)\left(\frac{2}{\rho} - \frac{2b_0 - b}{b_0 - b}\right) + b\rho^{b\mu}(1 - \log\rho).$$

https://doi.org/10.1017/apr.2019.64 Published online by Cambridge University Press

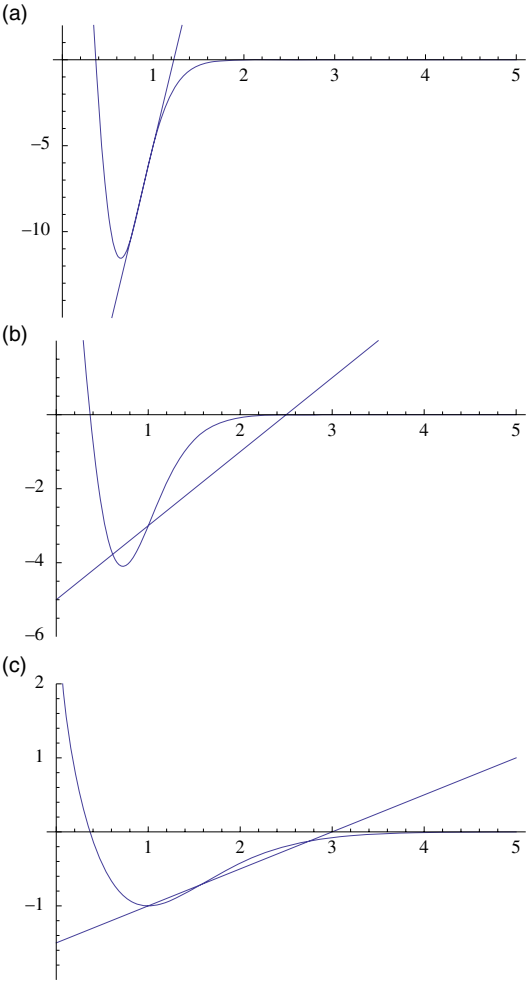



FIGURE 9: Illustration of the left- and right-hand sides of (10), where $\mu = \lambda = 1$ is the largest solution ((a) $b = 5$, $b_0 = 16$), the middle solution ((b) $b = 3$, $b_0 = 4$), or the smallest solution ((c) $b = 1$, $b_0 = 1.25$).

Note that $g(0) > 0$, $\lim_{\mu\to\infty} g(\mu) > 0$, $g(\lambda) = 0$, and $g'(\lambda) = 0$. Thus (9) will hold if we can show that $g'(\mu)$ has at most three zeros including one at $\mu = \lambda$. Hence $g'(\mu) = 0$ is equivalent to

$$(b_0 - b)\left(\frac{2}{\rho} - \frac{2b_0 - b}{b_0 - b}\right) = -b\rho^{b\mu}(1 - \log\rho). \tag{10}$$

We make a number of observations about (10) (see Figure 9):

(a) The left-hand side is an increasing linear function in $\mu$ because $b_0 \geq b$.

(b) The right-hand side is positive only if $\mu < \lambda/e$ and negative otherwise.

(c) The first derivative of the right-hand side with respect to $\mu$ is $(b/\mu)\rho^{b\mu}(b\mu(1 - \log\rho)^2 - 1)$. From (c), we know that there exists a unique $\mu_a$ such that $b\mu(1 -$

https://doi.org/10.1017/apr.2019.64 Published online by Cambridge University Press

$\log \rho)^2 = 1$, and the right-hand side decreases in $\mu$ if $\mu < \mu_a$, and increases to 0 if $\mu > \mu_a$.

(d) The second derivative of the right-hand side with respect to $\mu$ is

$$b\mu^{-2}\rho^{b\mu}(1 + 3b\mu(1 - \log \rho) - b^2\mu^2(1 - \log \rho)^3).$$

There exists $\mu_b$ such that the right-hand side is convex if $\lambda/e < \mu < \mu_b$ and concave if $\mu > \mu_b$.

(e) We have $\mu_b > \mu_a$ because the function is convex around the minimum at $\mu_a$. Then the right-hand side is convex decreasing if $\lambda/e < \mu < \mu_a$, convex increasing if $\mu_a < \mu < \mu_b$, and concave increasing if $\mu > \mu_b$. Depending on the values of $\mu_a$, $\mu_b$, and 1, we have the following three cases.

(i) $\mu_a \le \mu_b \le \lambda$. If the slope of the left-hand side is greater than that of the right-hand side, we can have at most three solutions to (10) where all the non-trivial solutions are smaller than $\lambda$: $\mu_1 \le \mu_2 \le \mu_3 = \lambda$. If the slope of the left-hand side is smaller, we also have three solutions: $\mu_1 \le \mu_2 = 1 \le \mu_3$.

(ii) If $\mu_a \le \lambda \le \mu_b$, we have three solutions to (10): $\mu_1 \le \mu_2 = \lambda \le \mu_3$.

(iii) If $\lambda \le \mu_a \le \mu_b$. we can have at most three solutions to (10): $\mu_1 = \lambda \le \mu_2 \le \mu_3$.

We have shown that $g'(\mu)$ has at most three zeros, which implies that $g(\mu)$ has at most three stationary points. If there are two different intervals of $\mu$ within which $g(\mu) \le 0$, this is only possible if $g(\mu_1) < 0$, $g(\mu_2) > 0$, and $g(\mu_3) < 0$ (i.e. $g(\mu)$ has a $W$ shape). Given that one of the three stationary points is $\mu = \lambda$ and $g(\lambda) = g'(\lambda) = 0$, it follows that one of the $g(\mu_1)$, $g(\mu_2)$, and $g(\mu_3)$ equals zero. Thus there exists a unique interval $(\mu_L, \mu_H)$ such that (9) holds for $\mu \neq \lambda$, and $\tilde{n}_S(\mu) \ge \tilde{n}_W(\mu)$ if and only if $\mu \neq \lambda$ is within that interval. This further implies that, after applying the integer constraint on the threshold, there exists an interval within which $n_S(\mu) \ge n_W(\mu)$, otherwise $n_S(\mu) \le n_W(\mu)$.

For $\mu \notin (\mu_L(p), \mu_H(p))$, in which the selfish threshold is smaller than or equal to the social threshold, the threshold increase makes the throughput closer to the socially optimal level, so it increases the welfare. The argument for $\mu \in (\mu_L(p), \mu_H(p))$ is similar.

(iv) We only need to show that the customers' utility is continuous and increasing in $\mu$. If the service rate changes, the collective threshold also changes if and only if it improves the resultant customer utility. Therefore the average utility is continuous in $\mu$, and trivially, the collective customers' utility increases in $\mu$. Thus the discontinuity of the welfare function when the threshold changes only comes from the profit part, which increases in $n$. □

*Proof of Proposition 3 (Benefit of Anarchy when the social firm sets capacity).* (ii) We first show that the equilibrium welfare with collective customers and the social firm decreases in $p$. From Section 3 we know that $p = 0$ makes collective customers choose the socially optimal threshold for any $\mu$. Let $n_0(\mu)$ denote the socially optimal threshold for given $\mu$. Fix $\mu$ and consider increasing the price. Because the threshold is integer-valued, the price can increase without affecting the threshold. However, increasing the price enough makes the collective threshold smaller, in turn making the gap from $n_0(\mu)$ larger. As a result, for any $\mu$, the welfare is decreasing in $p$, so the equilibrium welfare decreases in $p$.

Now consider the case with selfish customers. Fixing $\mu$, the selfish threshold decreases in $p$ and there exists a price interval $(\underline{p}(\mu), \bar{p}(\mu))$ such that the selfish threshold is $n_0(\mu)$ for

https://doi.org/10.1017/apr.2019.64 Published online by Cambridge University Press

$p \in (\underline{p}(\mu), \bar{p}(\mu))$, resulting in the first-best welfare for given $\mu$. As in the collective customers' case, the gap between the selfish threshold and $n_0(\mu)$ increases as the price increases from $\bar{p}(\mu)$. Proposition 1 tells us that this gap is not greater than the gap between the collective threshold and $n_0(\mu)$, which implies that the resultant welfare at $\mu$ with selfish customers is not smaller than that with collective customers. Furthermore, as the price increases from 0 to $R$, the optimal service rate for collective customers changes, and we denote the corresponding optima by $\mu_0, \mu_1, \ldots, \mu_n, \ldots, \mu_N$. Then, for a sufficiently high $p \geq \max\{\underline{p}(\mu_n), \ldots, \underline{p}(\mu_N)\}$ with the corresponding optima $\mu_n, \ldots,$ and $\mu_N$ for collective customers, we know that the welfare of selfish customers is not smaller than that of collective customers so the equilibrium welfare with selfish customers is not smaller than that with collective customers. □

*Proof of Lemma 5 (welfare-maximizing service rate).* With selfish customers, we observe three cases of the Nash equilibrium:

(a) If $\mu \leq \mu_S^0$, then $q_S(\mu) = 0$. The best value of $\mu$ in this range is $\mu = 0$.

(b) If $\mu \geq \mu_S^1$, then

$$q_S(\mu) = 1 \quad \text{and} \quad \mathcal{W}_S(\mu) = \lambda\left(R - \frac{h}{\mu - \lambda}\right) - \bar{c}\mu.$$

The first and second derivatives of $\mathcal{W}_S(\mu)$ with respect to $\mu$ within this range are

$$\frac{h\lambda}{(\mu - \lambda)^2} - \bar{c} \quad \text{and} \quad -\frac{2h\lambda}{(\mu - \lambda)^3}.$$

Thus the welfare function is concave for $\mu > \lambda$ with a stationary point of $\bar{\mu} := \lambda + \sqrt{h\lambda/\bar{c}}$. The best value of $\mu$ in this range is $\bar{\mu}$ with $\mathcal{W}_S(\bar{\mu}) = \lambda(R - \bar{c}) - 2\sqrt{\bar{c}h\lambda}$ if $\bar{\mu} > \mu_S^1$ (i.e. $\bar{c} < (\lambda(R-p)^2)/h$), otherwise $\mu_S^1$.

(c) If $\mu \in (\mu_S^0, \mu_S^1)$, then

$$q_S(\mu) = \frac{\mu - 1/b}{\lambda} \quad \text{and} \quad \mathcal{W}_S(\mu) = (\mu - 1/b)p - \bar{c}\mu.$$

The welfare is linear and increasing in $\mu$. Thus the best value of $\mu$ in this range is $\mu_S^1$ with $\mathcal{W}_S(\mu_S^1) = \lambda p - \bar{c}(1/b + \lambda)$.

Combining these results, the social firm will choose $\mu_S^W = \bar{\mu} = \lambda + \sqrt{h\lambda/\bar{c}}$ if $\bar{c} < (\lambda(R-p)^2)/h$ and $\lambda(R - \bar{c}) - 2\sqrt{\bar{c}h\lambda} > 0$, $\mu_S^W = \mu_S^1$ if $c > (\lambda(R-p)^2)/h$ and $\lambda p - \bar{c}(1/b + \lambda) > 0$, and $\mu_S^W = 0$ otherwise.

Now consider the collective customers' case. We observe three cases:

(a) If $\mu \leq \mu_C^0$, then $q_C(\mu) = 0$. The best value of $\mu$ in this range is $\mu = 0$.

(b) If $\mu \geq \mu_C^1$, then $q_C(\mu) = 1$. As in the selfish customers' case, the best value of $\mu$ in this range is $\bar{\mu}$ with $\mathcal{W}_C(\bar{\mu}) = \lambda(R - \bar{c}) - 2\sqrt{\bar{c}h\lambda}$ if $\bar{\mu} > \mu_C^1$ (i.e. $\bar{c} < h\lambda/((\mu_C^1 - \lambda)^2)$), otherwise $\mu_C^1$.

(c) If $\mu_C^0 < \mu < \mu_C^1$, then

$$q_C(\mu) = \frac{\mu - \sqrt{\mu/b}}{\lambda} \quad \text{and} \quad \mathcal{W}_C(\mu) = (\mu - \sqrt{\mu/b})\left(R - \frac{h}{\sqrt{\mu/b}}\right) - \bar{c}\mu.$$

https://doi.org/10.1017/apr.2019.64 Published online by Cambridge University Press

It can be shown that the welfare is convex, that is,

$$\mathcal{W}_C''(\mu) = \frac{1/b^2 \cdot (R/b + h)\mu}{4(\mu/b)^{5/2}} > 0.$$

The best value of $\mu$ is either $\mu_C^1$ or $\mu_C^0$.

Therefore, for collective customers, the social firm will choose $\mu_C^W = \bar{\mu}$ if $\bar{c} < h\lambda/((\mu_C^1 - \lambda)^2)$ and $\lambda(R - \bar{c}) - 2\sqrt{\bar{c}h\lambda} > 0$, $\mu_C^W = \mu_C^1$ if $\bar{c} > h\lambda/(\mu_C^1 - \lambda)^2$ and $\lambda(R - h/(\mu_C^1 - \lambda)) - \bar{c}\mu_C^1 > 0$, and $\mu_C^W = 0$ otherwise. □

*Proof of Proposition 6(ii) (SBoA with the social firm).* We have three regions of $\bar{c}$:

(a) If $\bar{c} < 4b^2h\lambda/(1 + \sqrt{1 + 4b\lambda})^2$, then $\max\{\mu_S^1, \bar{\mu}\} = \max\{\mu_C^1, \bar{\mu}\} = \bar{\mu}$, so the equilibrium welfare is the same.

(b) If $4b^2h\lambda/(1 + \sqrt{1 + 4b\lambda})^2 < \bar{c} < b^2h\lambda$, then $\mu_C^1 < \bar{\mu} < \mu_S^1$, so the first-best welfare cannot be achieved with collective customers. If $\mu_S^W = \bar{\mu}$ and $\mu_C^W = \mu_C^1$ or 0, the welfare with selfish customers is higher. If $\mu_S^W = 0$ (because the first-best welfare at $\bar{\mu}$ is negative), $\mu_C^W = 0$.

(c) If $\bar{c} > b^2h\lambda$, the first-best welfare at $\bar{\mu}$ and $q = 1$ is not achievable with both selfish and collective customers. We first compare $\mu_S^W = \mu_S^1$ and $\mu_C^W = \mu_C^1$. We can show that

$$\lambda\left(R - \frac{h}{\mu_S^1 - \lambda}\right) - \bar{c}(\mu_S^1) > \lambda\left(R - \frac{h}{\mu_C^1 - \lambda}\right) - \bar{c}(\mu_C^1)$$

if

$$\bar{c} > b^2h\lambda \Big/ \left(\frac{1 + \sqrt{1 + 4b\lambda}}{2}\right),$$

which is always true. For the same reason, $\mu_S^W = 0$ and $\mu_C^W = \mu_C^1$ is not possible.

Therefore, for $\bar{c} > 4b^2h\lambda/(1 + \sqrt{1 + 4b\lambda})^2$, the welfare with selfish customers is higher than that with collective customers as long as the firm operates. □

*Proof of Lemma 6 (profit-maximizing price).* For a fixed $p$, we have three cases for selfish customers:

(a) If $p \geq R - h/\mu =: p_S^0$, then $q_S(p) = 0$, so the firm will not serve any customers, and $U_S(p) = 0$.

(b) If $p \leq R - h/(\mu - \lambda) =: p_S^1$, $R > h/(\mu - \lambda)$, and $\mu > \lambda$, then $q_S(p) = 1$. If $p_S^1 > 0$, because $\Pi_S(p) = \lambda p - \bar{c}\mu$ is increasing in $p \in [0, p_S^1]$, the firm chooses $p = p_S^1$, with profit $\Pi_S(p_S^1) = \lambda p_S^1 - \bar{c}\mu$, and $U_S(p_S^1) = 0$.

(c) If $p \in (p_S^1, p_S^0)$, the firm's profit function is

$$\Pi_S(\mu) = p\left(\mu - \frac{h}{R - p}\right) - \bar{c}\mu,$$

which, in contrast to the capacity-control case, is *concave*. The first-order condition gives the local maximum $\bar{p}_S := R - \sqrt{hR/\mu}$. By assumption, $R > h/\mu$, which means

https://doi.org/10.1017/apr.2019.64 Published online by Cambridge University Press

that $\bar{p}_S < p_S^0$. If $\bar{p}_S < p_S^1$ then $p_S^P = p_S^1$ and $q_S(p_S^P) = 1$. If $p_S^1 \le \bar{p}_S < p_S^0$, then

$$p_S^P = \bar{p}_S \quad \text{and} \quad q_S(p_S^P) = \frac{1}{\lambda}\left(\mu - \sqrt{\frac{h\mu}{R}}\right).$$

Therefore $p_S^P = \max\{\bar{p}_S, p_S^1\}$, and the utility of selfish customers is zero, i.e. $U_S(p_S^P) = 0$.

For a fixed $p$, we have three cases for collective customers:

(a) If $p \ge R - h/\mu =: p_C^0$, then $q_C(p) = 0$, so the firm will not serve any customers, and $U_C(p) = 0$.

(b) If

$$p \le R - \frac{h\mu}{(\mu-\lambda)^2} =: p_C^1, \quad R - \frac{h\mu}{(\mu-\lambda)^2} > 0 \quad \text{and} \quad \mu > \lambda,$$

then $q_C(p) = 1$ and $\Pi(p) = \lambda p - \bar{c}\mu$, so the firm chooses the largest $p$ in this range,

$$p_C^1 = \left(R - \frac{h\mu}{(\mu-\lambda)^2}\right) I\{\mu > \lambda\},$$

and the utility will be

$$U_C(p_C^1) = \frac{\lambda^2 h}{(\mu-\lambda)^2} > 0.$$

(c) If $p \in (p_C^1, p_C^0)$, then $q_C(p) \le 1$ and

$$\Pi_C(p) = \lambda p q_C(p) - \bar{c}\mu = p\left(\mu - \sqrt{\frac{h\mu}{R-p}}\right) - \bar{c}\mu.$$

This can be shown to be concave in $p$, as was the case for selfish customers. Let $\bar{p}_C$ be the solution to the first-order condition: $h(2R - \bar{p}_C)^2 = 4\mu(R - \bar{p}_C)^3$. Note that $0 \le \bar{p}_C \le p_C^0$, and for $\bar{p}_C \le p_C^1$, $p_C^P = p_C^1$ (because we are in case (b)), while for $\bar{p}_C > p_C^1$, $p_C^P = \bar{p}_C$. Therefore $p_C^P = \max\{\bar{p}_C, p_C^1\}$, as long as $\lambda p_C^P\, q(p_C^P) \ge \bar{c}\mu$. We also have that $p_C^P < p_S^P$ because $\bar{p}_C < \bar{p}_S$ and $p_C^1 < p_S^1$. □

*Proof of Lemma 7 (welfare-maximizing price).* For customer type $\alpha \in \{S, C\}$, the welfare function of $p$ is given by

$$\mathcal{W}_\alpha(p) = \lambda q_\alpha(p)\left(R - \frac{h}{\mu - \lambda q_\alpha(p)}\right) - \bar{c}\mu.$$

Note that the price only affects the welfare through $q_\alpha(p)$. We first solve for the optimal $q$ that maximizes the welfare. The first and second derivatives of $\mathcal{W}$ with respect to $q$ are

$$\mathcal{W}_q = \lambda\left(R - \frac{h}{\mu - \lambda q} - \frac{\lambda h q}{(\mu - \lambda q)^2}\right) \quad \text{and} \quad \mathcal{W}_{qq} = -\frac{2\lambda^2 h\mu}{(\mu - \lambda q)^3} < 0.$$

Note that the welfare function is concave for any $\mu > \lambda q$ with a stationary point of

$$q^W = \frac{1}{\lambda}\left(\mu - \sqrt{\frac{h\mu}{R}}\right).$$

https://doi.org/10.1017/apr.2019.64 Published online by Cambridge University Press

Given that $q_\alpha(p)$ is monotonically decreasing in $p$, the social firm will choose $p_\alpha^W$ to make the resultant $q$ to be $q^W$.

With selfish customers and $p \in \{p_S^1, p_S^0\}$, the joining probability is

$$q_S(p) = \frac{1}{\lambda}\left(\mu - \frac{h}{R-p}\right).$$

The social firm wants $q_S(p) = q^W$, resulting in the solution $p_S^W = R - \sqrt{hR/\mu}$.

With collective customers, $p_C^W = 0$ always generates the socially optimal equilibrium. This is because $p = 0$ makes the maximization of collective utility equivalent to the maximization of social welfare. □

*Proof of Lemma 8 (upper-envelope function).* (i) Similar to the proof of Lemma 2, the relaxed profit function can be shown to be continuous using L'Hôpital's rule. To show that the relaxed profit function is an upper envelope, note that $\lambda_e(\mu, \tilde{n}(\mu)) \geq \lambda_e(\mu, n(\mu))$ holds because $\lambda_e(\mu, n)$ is increasing in $n$ and $\tilde{n}(\mu) \geq n(\mu)$ for any $\mu$, and the equality holds only for every $\underline{\mu}(n)$ because $\tilde{n}(\mu) = n(\mu)$ only for $\mu = \underline{\mu}(n)$.

(ii) Consider a profit maximizer $\mu^P$ with corresponding threshold $n^P$, i.e. $\mu^P \in [\underline{\mu}(n^*), \underline{\mu}(n^*+1))$. Suppose there exists no stationary point of the relaxed profit function within $(\underline{\mu}(n^*-1), \underline{\mu}(n^*+1))$. This implies that the relaxed profit function is monotone within $[\underline{\mu}(n^*-1), \underline{\mu}(n^*+1)]$. If the relaxed profit function is decreasing, $\Pi(\underline{\mu}(n^*-1), n^*-1)$ is greater than $\Pi(\mu^*, n^*)$, which is a contradiction. Likewise, if the relaxed profit function is increasing, $\Pi(\underline{\mu}(n^*+1), n^*+1)$ is greater than $\Pi(\mu^*, n^*)$, which is a contradiction. □

*Proof of Lemma 9 (upper-envelope function with selfish customers).* We first show that the relaxed profit function with selfish customers is concave for any $\mu$ if $\tilde{n} \geq 2$. Then we will show that (4) is a sufficient condition for the equilibrium threshold to be greater than or equal to 2.

The relaxed effective arrival rate with the selfish customer is

$$\lambda_e(\mu, \tilde{n}_S(\mu)) = \lambda \cdot \frac{1-\rho^{b\mu}}{1-\rho^{b\mu+1}}.$$

The first and second derivatives of $\lambda_e(\mu, \tilde{n}_S(\mu))$ with respect to $\mu$ are given by

$$\frac{\mathrm{d}\lambda_e(\mu, \tilde{n}_S(\mu))}{\mathrm{d}\mu} = \lambda \cdot \frac{\rho^{b\mu}}{(1-\rho^{b\mu+1})^2} \cdot (b\mu(1-\rho)(1-\log\rho) - \rho(1-\rho^{b\mu})) \geq 0,$$

and

$$\begin{aligned}
&\frac{\mathrm{d}^2\lambda_e(\mu, \tilde{n}_S(\mu))}{\mathrm{d}\mu^2} \\
&= -\frac{\rho^{b\mu}}{(1-\rho^{b\mu+1})^3}\Bigg[-2\frac{\rho}{\mu}(1-\rho^{b\mu}) - b(1+\rho-5\rho^{b\mu+1}+3\rho^{b\mu+2}) \\
&\quad + b^2\mu(1-\rho)(1+\rho^{b\mu+1}) + 2b(\rho-2\rho^{b\mu+1}+\rho^{b\mu+2} - b\mu(1-\rho)(1+\rho^{b\mu+1}))\log\rho \\
&\quad + b^2\mu(1-\rho)(1+\rho^{b\mu+1})(\log\rho)^2\Bigg].
\end{aligned}$$

https://doi.org/10.1017/apr.2019.64 Published online by Cambridge University Press

We want to show that the second derivative is non-positive for any $b$ and $\mu > 1/b$. After replacing $b$ with $\tilde{n}/\mu$ and letting $x = \rho^{-1}$, it is equivalent to showing the following:

$$F(x,\tilde{n}) = -\tilde{n}(1+x^{\tilde{n}+1})(1+\log x)^2 + 2\left(\frac{x^{\tilde{n}+1}-1}{x-1} - 2\right)(1+\log x) + (x^{\tilde{n}+1}-1) + \frac{2}{\tilde{n}}\frac{x^{\tilde{n}+1}-x}{x-1} \le 0$$

for any $x > 0$ and $\tilde{n} > 1$.

$F(x,\tilde{n})$ can be shown to be decreasing in $\tilde{n}$ as follows:

$$\begin{aligned}
\frac{\partial F(x,\tilde{n})}{\partial \tilde{n}} &= \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[(-(\tilde{n}+\epsilon)x^{\tilde{n}+1+\epsilon} + \tilde{n}x^{\tilde{n}+1} - \epsilon)(1+\log x)^2 + 2x^{\tilde{n}+1}\frac{x^\epsilon - 1}{x-1}(1+\log x)\right.\\
&\quad \left. + x^{\tilde{n}+1}(x^\epsilon - 1) + \frac{2x}{(x-1)}\left(\frac{x^{\tilde{n}+\epsilon}-1}{\tilde{n}+\epsilon} - \frac{x^{\tilde{n}}-1}{\tilde{n}}\right)\right]\\
&\le \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[(-(\tilde{n}+\epsilon)x^{\tilde{n}+1+\epsilon} + \tilde{n}x^{\tilde{n}+1} - \epsilon)(1+\log x)^2 + 2x^{\tilde{n}+1}\frac{x^\epsilon - 1}{x-1}(1+\log x)\right.\\
&\quad \left. x^{\tilde{n}+1}(x^\epsilon - 1) + x^{\tilde{n}+1}(x^\epsilon - 1)\right] \qquad (11)\\
&= \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[(-(\tilde{n}+\epsilon)x^{\tilde{n}+1+\epsilon} + \tilde{n}x^{\tilde{n}+1} - \epsilon)(1+\log x)^2 + 2x^{\tilde{n}+1}\frac{x^\epsilon - 1}{x-1}(x+\log x)\right]\\
&= \lim_{\epsilon\to 0}\frac{x^{\tilde{n}+1}}{\epsilon}\left[(-(\tilde{n}+\epsilon)x^\epsilon + \tilde{n} - \epsilon x^{-\tilde{n}-1})(1+\log x)^2 + 2\frac{x+\log x}{x-1}(x^\epsilon - 1)\right]\\
&\le \lim_{\epsilon\to 0}\frac{x^{\tilde{n}+1}}{\epsilon}\left[(-(\tilde{n}+\epsilon)x^\epsilon + \tilde{n} - \epsilon x^{-\tilde{n}-1})(1+\log x)^2 + \frac{4x-2}{x-1}(x^\epsilon - 1)\right] \qquad (12)\\
&\le \lim_{\epsilon\to 0}\frac{x^{\tilde{n}+1}}{\epsilon}\left[(-(1+\epsilon)x^\epsilon + 1 - \epsilon x^{-2})(1+\log x)^2 + \frac{4x-2}{x-1}(x^\epsilon - 1)\right] \qquad (13)\\
&= \lim_{\epsilon\to 0} x^{\tilde{n}+1}\left[-(x^\epsilon + (1+\epsilon)x^\epsilon\log x + x^{-2})(1+\log x)^2 + \frac{4x-2}{x-1}x^\epsilon\log x\right]\\
&= x^{\tilde{n}+1}\left\{-(1+\log x + x^{-2})(1+\log x)^2 + \frac{4x-2}{x-1}\log x\right\}\\
&\le 0,
\end{aligned}$$

where the upper bounds in (11), (12), and (13) are derived from the following.

(i) The upper bound (11) is derived from the fact that $\partial x^{\tilde{n}}/\partial\tilde{n}$ is greater than

$$\left(\partial\frac{2}{\tilde{n}}\frac{x^{\tilde{n}}-1}{x-1}\right)\Big/\partial\tilde{n},$$

that is,

$$x^{\tilde{n}}\log x \ge -\frac{2(x^{\tilde{n}}-1)}{\tilde{n}^2(x-1)} + \frac{2\tilde{n}x^{\tilde{n}}\log x}{\tilde{n}^2(x-1)},$$

https://doi.org/10.1017/apr.2019.64 Published online by Cambridge University Press

which is equivalent to

$$\log x \begin{cases} > \dfrac{2(x^{-\tilde{n}}-1)}{\tilde{n}^2(x-1)-2\tilde{n}} & \text{if } x<1 \text{ or } x>1+\dfrac{2}{\tilde{n}}, \\ \leq \dfrac{2(x^{-\tilde{n}}-1)}{\tilde{n}^2(x-1)-2\tilde{n}} & \text{if } 1<x<1+\dfrac{2}{\tilde{n}}. \end{cases} \tag{14}$$

If $x > 1 + 2/\tilde{n}$, (14) is trivially true since the left-hand side is positive while the right-hand side is negative.

If $x < 1 + 2/\tilde{n}$, we first show that the right-hand side of (14) is decreasing in $\tilde{n}$ if $x < 1$ and increasing in $\tilde{n}$ if $x > 1$. The derivative of the right-hand side with respect to $\tilde{n}$ is

$$\frac{4(x^{\tilde{n}}-1)(\tilde{n}(x-1)-1)+2\tilde{n}(2-\tilde{n}(x-1))\log x}{\tilde{n}^2x^{\tilde{n}}(\tilde{n}(x+1)^2-2)^2}. \tag{15}$$

The derivative of the numerator in (15) with respect to $x$ can be shown to be non-negative:

$$\begin{aligned} \frac{2\tilde{n}}{x}\{(x-1)(2(\tilde{n}+1)x^{\tilde{n}}-\tilde{n}-2)-\tilde{n}x\log x\} &\geq \frac{2\tilde{n}}{x}\{(x-1)(2(\tilde{n}+1)x^{\tilde{n}}-\tilde{n}-2)-\tilde{n}x(x-1)\} \\ &= \frac{2\tilde{n}(x-1)}{x}\{2(\tilde{n}+1)(x^{\tilde{n}}-1)-\tilde{n}(x-1)\} \\ &\geq 0. \end{aligned}$$

Given that the numerator of (15) is increasing in $x$ and equal to 0 when $x = 1$, it follows that the right-hand side of (14) is decreasing in $\tilde{n}$ if $x < 1$ and increasing in $\tilde{n}$ if $x > 1$. Thus it suffices to show that (14) is true when $\tilde{n} = 1$, and this can be shown as follows:

$$\begin{aligned} \log x > \frac{-2(x-1)}{x(x-3)} \geq \frac{2(x^{-\tilde{n}}-1)}{\tilde{n}^2(x-1)-2\tilde{n}} \quad &\text{if } x<1, \\ \log x \leq \frac{-2(x-1)}{x(x-3)} \leq \frac{2(x^{-\tilde{n}}-1)}{\tilde{n}^2(x-1)-2\tilde{n}} \quad &\text{if } 1<x<1+\frac{2}{\tilde{n}}. \end{aligned}$$

(ii) The upper bound (12) is derived from $\log x \leq x - 1$.

(iii) The upper bound (13) holds because the value in the square bracket of (12) can be shown to be decreasing in $n$. Let $l(x, \tilde{n}) := -(\tilde{n}+\epsilon)x^{\epsilon} + \tilde{n} - \epsilon x^{-\tilde{n}-1}$ and $l_{\tilde{n}}(x, \tilde{n}) = 1 - x^{\epsilon} + \epsilon x^{-\tilde{n}-1}\log x$, the derivative of $l(x, \tilde{n})$ with respect to $\tilde{n}$. For small enough $\epsilon > 0$, both $l(x, \tilde{n})$ and $l_{\tilde{n}}(x, \tilde{n})$ are negative. Thus $\tilde{n} = 1$ gives the upper bound.

Since $F(x, \tilde{n})$ is decreasing in $\tilde{n}$, the upper bound of $F(x, \tilde{n})$ can be derived by plugging in the smallest value of $\tilde{n}$. If $x > 1$ (i.e. $\rho < 1$), we can show that $F(x, \tilde{n})$ is negative for any $\tilde{n}$:

$$\begin{aligned} F(x,\tilde{n}) \leq F(x,1) &= (2+\log x)[2(x-1)-(x^2+1)\log x] \\ &\leq (2+\log x)[2(x-1)-(x^2+1)(x-1)/x] \\ &= (2+\log x)(-(x-1)^3/x) \\ &\leq 0. \end{aligned}$$

If $x < 1$ (i.e. $\rho > 1$), we need $\tilde{n} \geq 2$ for the concavity of the relaxed profit function:

$$\begin{aligned} F(x,\tilde{n}) &\leq F(x,2) \\ &= -2(x^3+1)(1+\log x)^2 + 2(x^2+x-1)(1+\log x) + (x^3+x^2+x-1) \\ &\leq 0. \end{aligned}$$

https://doi.org/10.1017/apr.2019.64 Published online by Cambridge University Press

Thus, for any $\rho > 0$, the sufficient condition for the concave relaxed profit function is that the equilibrium threshold, $n_{SP}$, is greater than or equal to 2. A sufficient condition for $n_{SP} \geq 2$ is that any profit in the threshold interval for $n = 1$ is less than the starting point of the threshold interval for $n = 2$, given by

$$\max_{\mu \in [1/b, 2/b)} \Pi(\mu, 1) < p\left(1 - \frac{2/b - 1}{(2/b)^2 - 1}\right) - c\left(\frac{1}{b}\right) \leq p\left(1 - \frac{2/b - 1}{(2/b)^3 - 1}\right) - c\left(\frac{2}{b}\right).$$

With some algebra, this is equivalent to (4). Then $n_{SP} \geq 2$ and the envelope method specifies the unique interval that contains the profit-maximizer with selfish customers. □

## Acknowledgements

We appreciate the contributions of Tyler Maxey of UC Berkeley who did numerical work and helped with the analysis in Section 7. We thank Refael Hassin for his help in pointing us towards related literature.

## References

[1] Borgs, C., Chayes, J. T., Doroudi, S., Harchol-Balter, M. and Xu, K. (2014). The optimal admission threshold in observable queues with state dependent pricing. *Prob. Eng. Inf. Sci.* **28**, 101–119.
[2] Chen, H. and Frank, M. (2004). Monopoly pricing when customers queue. *IIE Trans.* **36**, 569–581.
[3] De Vany, A. (1976). Uncertainty, waiting time, and capacity utilization: a stochastic theory of product quality. *J. Political Econom.* **84** (3), 523–541.
[4] Dewan, S. and Mendelson, H. (1990). User delay costs and internal pricing for a service facility. *Manag. Sci.* **36**, 1502–1517.
[5] Edelson, N. M. and Hilderbrand, D. K. (1975). Congestion tolls for Poisson queuing processes. *Econometrica* **43**, 81–92.
[6] Gilboa-Freedman, G., Hassin, R. and Kerner, Y. (2014). The Price of Anarchy in the Markovian single server queue. *IEEE Trans. Automatic Control* **59**, 455–459.
[7] Grassmann, W. K. (1979). The economic service rate. *J. Operat. Res. Soc.* **30**, 149–155.
[8] Hassin, R. (1986). Consumer information in markets with random product quality: the case of queues and balking. *Econometrica* **54**, 1185–1195.
[9] Hassin, R. (2016). *Rational Queueing*. CRC press.
[10] Hassin, R. and Haviv, M. (2003). *To Queue or Not to Queue: Equilibrium Behavior in Queueing Systems*. Springer Science and Business Media.
[11] Haviv, M. and Roughgarden, T. (2007). The Price of Anarchy in an exponential multi-server. *Operat. Res. Lett.* **35**, 421–426.
[12] Knudsen, N. C. (1972). Individual and social optimization in a multiserver queue with a general cost-benefit structure. *Econometrica* **40**, 515–528.
[13] Lippman, S. A. and Stidham, S. Jr (1977). Individual versus social optimization in exponential congestion systems. *Operat. Res.* **25**, 233–247.
[14] Littlechild, S. (1974). Optimal arrival rate in a simple queueing system. *Internat. J. Production Research* **12**, 391–397.
[15] Maxey, T., Chung, H., Ahn, H.-S. and Righter, R. (2017). When is anarchy beneficial? *ACM SIGMETRICS Performance Evaluation Review* **45**, 18–20.
[16] Mendelson, H. (1985). Pricing computer services: queueing effects. *Commun. Assoc. Comput. Mach.* **28**, 312–321.
[17] Naor, P. (1969). The regulation of queue size by levying tolls. *Econometrica* **37**, 15–24.
[18] Roughgarden, T. (2005). *Selfish Routing and the Price of Anarchy*. MIT Press, Cambridge, MA.
[19] Stidham, S. Jr (1992). Pricing and capacity decisions for a service facility: stability and multiple local optima. *Manag. Sci.* **38**, 1121–1139.
[20] Yechiali, U. (1971). On optimal balking rules and toll charges in the GI/M/1 queuing process. *Operat. Res.* **19**, 349–370.
[21] Yechiali, U. (1972). Customers' optimal joining rules for the GI/M/s queue. *Manag. Sci.* **18**, 434–443.
[22] Zhang, J., Pourazarm, S., Cassandras, C. G. and Paschalidis, I. C. (2018). The Price of Anarchy in transportation networks: data-driven evaluation and reduction strategies. *Proc. IEEE* 106, 538–553.

https://doi.org/10.1017/apr.2019.64 Published online by Cambridge University Press