Pricing and waiting time decisions in a health care market with private and public provision

Alves, Vasco

University of Birmingham

8 April 2019

Pricing and Waiting Time Decisions in a Health Care Market with Private and Public Provision

Vasco F. Alves[a*]

[a]Department of Economics, University of Birmingham, Edgbaston, Birmingham, Warwickshire, B15 2TT, UK

ARTICLE HISTORY
Compiled April 25, 2019

ABSTRACT
This paper describes a duopoly market for healthcare where one of the two providers is publicly owned and charges a price of zero, while the other sets a price so as to maximize its profit. Both providers are subject to congestion in the form of an *M/M/1* queue, and they serve patient-consumers who have randomly distributed unit costs of time. Consumer demand (as market share) for both providers is obtained and described. The private provider's pricing decision is explored, equilibrium existence is proven, and conditions for uniqueness presented. Comparative statics for demand are presented. Social welfare functions are described and the welfare maximizing condition obtained. More detailed results are then obtained for cases when costs follow uniform and Kumaraswamy distributions. Numerical simulations are then performed for these distributions, employing several parameter values, demonstrating the private provider's pricing decision and its relationship with social welfare.

KEYWORDS
waiting times; queueing; private health care; competition

JEL CLASSIFICATION
I11, L13, D43

1. Introduction

Waiting times are a problem affecting most health care systems. This is especially true where rationing is required, although its root cause is simply that capacity is limited while treatments take time to perform. Queueing imposes a further cost on the patient-consumer, and in this way, health care fits into a wide body of literature which deals with the economic causes, behaviour and consequences of queues.

CONTACT V. F. Alves. Email: v.alves@bham.ac.uk

*Grateful thanks are extended to Tim Worrall for the many helpful discussions, to the University of Edinburgh School of Economics, for funding my PhD, during the course of which this paper was written, and to anonymous referees for the helpful suggestions for improvement. Any remaining errors are, of course, my own.

Rationing is common in public health care systems, such as the British NHS. Where systems of this type are present, a private sector often remains in operation, being able to offer shorter waiting times to those willing to pay. The present paper deals with the pricing decisions of a single profit maximizing private provider of healthcare, which operates in a market also served by a public provider offering free treatment. The approach to modelling the health care market follows a simplified version of that outlined in (Goddard, Malek, & Tavakoli, 1995). However, unlike that model, where the focus was on the NHS sector and its outcomes, the present paper focuses on the pricing decisions of the private sector: in particular, in (Goddard et al., 1995) the private sector set a market clearing price by design, and was not subject to congestion, whereas here the sole private provider exercises market power and is subject to congestion in a manner similar to the public sector—although in equilibrium, congestion will be lower in the former. The present model yields similar results to (Goddard et al., 1995) regarding the responses of demand for public and private sector treatment when capacity increases, namely, that an increase in capacity reduces expected waiting time in the public sector, and therefore increases demand for that sector while reducing it for the private. However, in the present model price is a decision variable, so that its response to an increase in capacity cannot be compared with the market-clearing response in that model.

The private provider's decision problem is solved with recourse to the basic model of competition in waiting times presented in (Levhari & Luski, 1978; Luski, 1976), taking a special case where one provider charges price zero (i.e., the NHS). This work's chief contribution is the analysis of private sector decisions when it is subject to congestion and possesses market power. It is appropriate to use queueing theory to model the waiting time features of the Health Care market, as argued persuasively by (Goddard et al., 1995). Queueing has been a subject of economic analysis since the seminal work of (Naor, 1969), which dealt with optimal queue sizes in $M/M/1$ FCFS queues, and sketched a framework for individual decisions taken by impatient consumers which has been almost universally followed. These results were extended to an arbitrary number of queues by (Knudsen, 1972).

1.1. Related Literature

Both (Levhari & Luski, 1978; Luski, 1976) employed the queueing framework to model duopolistic competition between two identical providers selling a good whose provision requires queueing. The two providers simultaneously choose price, and this decision affects demand both directly and indirectly through waiting time. Consumers are differentiated through a randomly distributed unit cost of waiting time. The key result of (Levhari & Luski, 1978) is to show there is a separating equilibrium where one provider specializes in serving consumers with high waiting costs at a high price, and the other provider serves consumers with low waiting costs at low prices. This framework is adapted for the present paper, with the public sector charging a price of zero, and the outside option being constructed such that no consumer will choose it in equilibrium.

A model similar to that in (Levhari & Luski, 1978) was devised by (Chen & Frank, 2004) for a monopolist provider, and then extended to a duopoly in (Chen & Wan, 2003). Their model assumes the presence of an outside option, and allows for heterogeneous firms. (Lederer & Li, 1997) has some relevance for the present problem, as it deals with consumer heterogeneity in delay cost, which is important in the health care context. (Li & Lee, 1994), describe a model of competition in three product characteristics: price, quality and service speed, where the good's value declines with time. This value-decay assumption is often used in the health economics literature.

More recently, (Melo, 2014) established the existence and uniqueness of a pure strategy price equilibrium in a congested market with free entry and price competition. His results predict congested markets will have a greater number of firms than is socially optimal. (Dube & Jain, 2014) revisited the Bertrand equilibrium between providers competing on cost and waiting time in a congestible system, found the conditions for the existence of a Nash equilibrium, then considered a setting where differentiated services with quality of service guarantees were introduced. However, this was found not to result in any global welfare improvement. (Deck, Kimbrough, & Mongrain, 2014) set up a model of duopoly competition on time and price, where one provider sets up an express and a regular checkout, charging for the former, competing with another with only one checkout. This was found to be harmful to sellers and beneficial to consumers, with the single queue seller servicing patient shoppers, driving down prices and profits, but increasing consumer surplus.

Models in this family have been applied to markets other than health care as well. (Zhou, Albuquerque, & Grewal, 2016) presents a model of competition in the airline industry where delay is one of several factors firms compete on, and analyse how strategic interactions affect the optimal delay and reliability measures. (Mazalov & Melnik, 2016) also consider a market for transport, where providers compete in price and service time, and solve the pricing problem for providers, defining the equilibrium intensity flows.

Moving away from industrial organization to the fields of health economics and the economics of publicly provided private goods, (Barzel, 1974) presents a theory of rationing "free" goods through waiting which is relevant to public health care provision, but does not have queueing as an important feature. (Iversen, 1993) has a model of resource allocation to meet waiting times within a national health service setting. (Hoel & Sæther, 2003) and (Marchand & Schroyen, 2005) also consider competition between private and public health providers—making important contributions to the welfare economics of the issue. This draws on previous work by (Lindsay & Feigenbaum, 1984) (and see (Cullis & Jones, 1986)), where rationing by waiting lists is used, with value decay rather than cost of time used to represent consumer impatience. Private providers have no capacity constraints, and therefore no queueing in equilibrium.

This value decay approach is followed by (Iversen, 1997), who however is chiefly concerned with whether rationing occurs in the public sector, and does not use queuing models.

(Propper, 2000) continued in the same vein as (Lindsay & Feigenbaum, 1984) by considering no capacity constraints in the private sector, and using this to analyse empirically the demand for private care in the UK. (Dimakou, Dimakou, & Basso, 2015) also develop an empirical model to find how hospitals characteristics determine waiting, finding the waiting time distribution for UK hospitals in the 1997-2005 period, and how changes to that distribution are affected by resource allocation, and priorities given to different kinds of patients.

More closely related to the present work, (Gravelle & Schroyen, 2016) model a setting of a public health system with different possible rules for hospital payment, where providers can choose quality. They develop a stochastic model of rationing by waiting and use it to derive welfare maximising payment to hospitals, which are linked to output, expected waiting times, quality, and a few other features. Crucially, they show that quality and waiting times must not be the only factors affecting price for optimal outcomes to occur. (Andritsos & Aflaki, 2015) use a game-theoretic queueing model to examine the effect of a hospital's objective (i.e., non-profit vs. for-profit) in hospital markets for elective care. They find the presence of competition can preclude a hospital from achieving economies of scale, and therefore increase waiting times. Moreover, they find that whether hospitals will engage in homogeneous and heterogeneous competition depends on the patients' willingness to wait before receiving care and the reimbursement level of the non-profit sector.

Finally, (Farnworth, 2003) is quite similar to the present paper, including capacity constraints on the fee-charging provider. The difference between it and the present paper is that while there is a fee-charging provider "competing" with a "free" one, the former has the same 'social' objectives as the latter and is therefore not profit maximizing, presenting a completely different problem. Indeed, the author specifically describes the model as presenting "two private hospitals that are publicly funded," and pricing decisions are made by "policy makers," for "equity reasons,"—it is the service rate that is the hospitals' decision variable. Conversely, the sole objective of the private sector provider in the present paper is profit maximization, and it operates without reference to externally set social policy. Nevertheless, he reaches a similar result in regard to waiting times' response to a price increase: a decrease (increase) in the fee-charging (free) sector, providing corroboration of one of the present model's results.

2. The Model

The model combines (Goddard et al., 1995)'s approach to queueing in health care markets with the framework for duopoly competition in price and waiting time in

(Levhari & Luski, 1978). Adaptations to the latter include: i) restricting the public provider to charge a price of zero; and ii) instead of there being a disutility of waiting in the queue, and positive utility coming from a good bought after being served, disutility accrues regardless of whether one chooses to queue or not (reflecting the suffering caused by disease) and the 'good' to be purchased is a treatment which eliminates the suffering. This is similar to what has been developed in (Farnworth, 2003; Iversen, 1997; Lindsay & Feigenbaum, 1984), and others. However, the framework represents an ongoing disutility rather than an exponential decay of treatment value, as generally assumed. While it can be shown that the two approaches are formally equivalent (for which, see Appendix A), the one adopted here is more in keeping with the approach adopted in the queueing literature.

2.1. Consumers

Consumers are risk neutral expected utility maximizers, whose utility function is linear in waiting time, and whose expected lifespan from the time of disease occurrence is τ.[1] The intuition behind the specification adopted here is that in healthcare, the value of a treatment is not so much a pure benefit, but the removal of a pre-existing condition causing disutility. The assumption that the disease does not cause a shortening of expected lifespan, but only disutility during that lifespan, obviously exchanges some generality for simplicity of treatment, but that accurately reflects many conditions, including some that cause the most problems with waiting times (hip fractures requiring a hip replacement are a notorious example).

Consumers suffer a disutility of c per unit of time, which can be interpreted as the severity of the consumer's illness. As illness severity varies across consumers, c is a random variable following a continuous and bounded probability distribution function:

$$f(c),\ x \in [0, \bar{c}], \tag{1}$$

where $\bar{c}$ is finite and $F(c)$ is the corresponding cumulative distribution function.

Disease cost c is suffered until removed by treatment at T_i, the waiting time for provider $i = \{n, p\}$, where n denotes the public and p the private providers. If the consumer does not seek treatment, they will suffer c until death, i.e., for τ units of time. The expected utility of seeking treatment from provider i is then:

$$U_i = -cT_i - P_i, \tag{2}$$

where $P_n = 0$, and $P_p = P$ is the price charged by the private provider. The expected utility of foregoing treatment, U_o, is:

$$U_o = -c\tau.$$

[1] As consumers are risk neutral, the expectations term is omitted throughout.

Total potential demand for health care is given by the exogenous parameter λ. This can be thought of as representing a patient population which is fixed in the short-run. The two providers will each service a share of this total, which will be the arrival rate for that provider's queue.[2] Following the literature (e.g., (Levhari & Luski, 1978)), this parameter is normalized such that $\lambda = 1$, without any loss of generality, because the unit of time t is arbitrary, so that it is always possible to find one measure of time for which the arrival rate is 1.

Unlike (Goddard et al., 1995), there is no consideration of consumer income—all consumers can afford the private sector treatment. While this is a restrictive assumption, made in the interest of simplicity, the problem is treated in this way to allow for a clear focus on the private provider's pricing decisions in response to the public sector provision. Interaction with consumer income constraints and heterogeneity, while doubtlessly important, is left for further research.

2.2. Providers

It is assumed that neither the private nor the public sectors have costs, or in what is perhaps a preferable interpretation, those costs are sunk in the short-run. It is further assumed that for both the private and public sectors, service times are identically and independently distributed along an exponential distribution with rate μ. The assumption is made that $\mu > 1$, meaning that either server is capable of, by itself, serving the entire demand stream $\lambda = 1$. The assumption is required in order to avoid explosive growth of waiting times. This rate is exogenous and common to both servers. This parameter can be taken to reflect the state of technology. For example, improvements in surgical techniques allowing for hospital stays to be reduced would lead to an increase in the service rate. In what is a simplifying departure from (Goddard et al., 1995), it is assumed that only one consumer can be treated at a time. This turns the queueing process into the well known $M/M/1$ system.

Expected waiting time for each provider, as a function of arrival and service rates, can be obtained from a well known result in queueing theory (see, *inter alia*, (Gross, Shortle, Thompson, & Harris, 2008)):

$$T_i = \frac{1}{\mu - \lambda_i}, \; i \in \{n, p\}. \tag{3}$$

As the present paper is concerned with the analysis of competition between public and private providers in health care, the assumption will be made that both the public and private providers exist and are used by at least one consumer. For this to happen, it must be the case that $U_n = -cT_n > U_o = -c\tau$, which implies that $\tau > T_n$. Since, by assumption, the public sector does not charge for treatment, regardless of demand,

[2] It is not required that the two shares add up to 1, as patients can opt for not seeking treatment. Nevertheless, given the way the model is constructed, the shares of the two providers will always add up to 1 in equilibrium. An explanation follows below.

and $\tau > T_n$, no consumers will forego treatment, and the public sector will serve its entire demand stream at zero monetary cost.

Meanwhile, private sector demand, given the exogenous parameter μ, is a function of P, which is set at the value that maximizes the provider's instantaneous profit function:

$$\max_P \pi = \max_P \lambda_p P. \tag{4}$$

3. Demand

3.1. Individual Choice

An individual consumer chooses to seek care from a particular provider if two conditions are satisfied: his expected utility is larger than that of joining the waiting list for the other provider, and it is larger than that of foregoing treatment.

For a consumer with a given value of c, expected utility when seeking care from the private or public provider can be obtained from eq. (2):

$$U_p = -cT_p - P, \tag{5}$$

and

$$U_n = -cT_n. \tag{6}$$

A consumer will then seek care from the private sector when:

$$c > \frac{P}{T_n - T_p}. \tag{7}$$

On the other hand, the consumer will choose the public provider when (assuming indifferent consumers will opt for the public sector):[3]

$$c \leq \frac{P}{T_n - T_p}. \tag{8}$$

Let c^* be the critical value of c for which a consumer is indifferent between the two providers:

$$c^*(P, T_n, T_p) = \frac{P}{T_n - T_p}. \tag{9}$$

[3]Returning to the assumption that both a public and a private sector exist, and are used by at least one consumer, it must be the case that $T_p < T_n$, and therefore $T_p < \tau$. Proof by contradiction: for $U_p > U_n$, it is necessary that $c > \frac{P}{T_n - T_p}$. But if $T_p > T_n$, this requires that $c < 0$, which is not a possible value of c. Cf. a similar discussion in (Luski, 1976).

3.2. Market Demand

Market demand for the private and public providers is obtained from the individual optimization process described in subsection 3.1. Potential demand λ divides itself between the two providers, with the share of consumers with $c \leq c^*$ choosing the public provider, and the share with $c > c^*$ the private provider, so that as c^* decreases, the private provider's market share increases. The two demand functions are then:

$$\lambda_p = \int_{c^*}^{\bar{c}} f(c)\,\mathrm{d}c \Rightarrow \lambda_p(c^*) = 1 - F(c^*), \text{ and} \tag{10}$$

$$\lambda_n = \int_{0}^{c^*} f(c)\,\mathrm{d}c \Rightarrow \lambda_n(c^*) = F(c^*). \tag{11}$$

It can be seen from eq. (9) that c^* is itself a function of T_n, T_p, and P. The former two values (as set out in eq. (3)) are themselves a function of demand and the parameter μ. Ultimately, demand is then determined by the shape of $f(c)$, and the values of μ and P, the latter being the private provider's decision variable, as set out in subsection 4.3 below.

4. Supply

4.1. Providers

Waiting times T_n and T_p depend on P, but only indirectly through λ_n and λ_p, both of which are a function of c^*. So it is possible, through implicit differentiation, to show that c^* is increasing in P even when taking into account the indirect effects, where c^*, λ_p and λ_n are as presented in (9)-(11). The direct effect corresponds to the direct increase of c^* in response to P, while the indirect effects are those mediated by the increase in T_n and decrease in T_p caused by the shift in the respective demands in response to an increase in price. Let ${c^*}'_P \equiv \frac{\partial c^*(P, T_n(\lambda_n(c^*),\mu), T_p(\lambda_p(c^*),\mu))}{\partial P}$, the total derivative of c^* with regard to P. This is given by the following expression:[4]

$${c^*}'_P = \left[(T_n - T_p) + c^* f(c^*)(T_n^2 + T_p^2)\right]^{-1} > 0. \tag{12}$$

This result leads to the following two lemmas, which establish that private sector demand is bounded between 0 and half of the market:[5]

Lemma 1. *For any continuous distribution function $f(c)$, $x \in [0, \bar{c}]$ meeting the conditions in* (1)*, there is a price $\bar{P}$ which at the limit makes demand for the private provider equal to zero.*

---

[4]See Appendix B for the more detailed derivation of this and following expressions.

[5]See Appendix B for the proofs of these, and all following Lemmas and Theorems.

Lemma 2. *For any distribution function $f(c)$, $x \in [0, \bar{c}]$ meeting the conditions in (1), when $P \to 0$, demand for the private sector equals that for the public sector: $\lambda_n(c^*) = \lambda_p(c^*) = \frac{1}{2}$.*

As $\partial\lambda_p/\partial P < 0$, this further implies that demand for private sector treatment will never exceed that for the public sector.

4.2. Comparative Statics

The same process can be used to ascertain the effects a shock to μ will have on c^*, if P does not change. Let ${c^*}'_\mu \equiv \frac{\partial c^*(P, T_n(\lambda_n(c^*),\mu), T_p(\lambda_p(c^*),\mu))}{\partial \mu}$, the total derivative of c^* with regard to μ. Then:

$$ {c^*}'_\mu = \frac{P\left(T_n^2 - T_p^2\right)}{(T_n - T_p)^2 + Pf(c^*)\left(T_n^2 + T_p^2\right)} > 0, \tag{13} $$

As demand for either provider, as set out in (10)-(11), is a direct function of c^* alone, it is easy to sign their comparative statics in respect of both P and μ, considering P as a parameter, which yield fairly intuitive results (see Appendix B for the detailed results): as expected, a price increase will cause demand for the private sector to fall and that for the public sector to increase. This effect reduces the expected waiting time for the private sector, for which the consumers with higher time costs are willing to pay the increased price. Section 4.3 describes the private sector's pricing decision.

On the other hand, when the service rate increases, expected waiting time falls for both sectors. This reduces the private sector's comparative advantage over the public sector, so that for a constant value of P, its demand would fall—similarly, to retain the same market share, it would have to reduce P.

4.3. Private Provider Optimization

The profit maximization process follows the model presented in (Levhari & Luski, 1978; Luski, 1976), but with the public provider constrained to charge a price of zero, so that the private provider will not interact strategically with the public sector. Rather, it will set P at the level P^* which maximizes its instantaneous profit function:

$$ P^* = \arg\max_P \pi = \arg\max_P \lambda_p P. \tag{14} $$

Note that waiting time in both sectors is then determined by the private provider, as it is a function of P and parameter μ (via c^*). Let $\pi'_P \equiv \frac{\partial \pi(\lambda_p(c^*(P, T_n(\lambda_n,\mu), T_p(\lambda_p,\mu))), P)}{\partial P}$, the total derivative of π with regard to P. Then the maximization problem will have the following first order condition:

$$ \pi'_P = \lambda_p(c^*(P, T_n(\lambda_n, \mu), T_p(\lambda_p, \mu)) + P{\lambda_p}'_P = 0, \tag{15} $$

where ${\lambda_p}'_P$ is known to be negative. P^* then follows from solving the first order condition for P:

$$P^* = \frac{1 - F(c^*)}{f(c^*)\, {c^*}'_P} = \frac{\lambda_p}{f(c^*)\, {c^*}'_P}. \tag{16}$$

The existence of P^* can be easily derived from the intermediate value theorem:

Theorem 1. *There exists at least one value of P which maximizes the private provider's profits.*

The firm's profit in equilibrium, π^* is then given by:

$$\pi^* = P^* \lambda_p(c^*(P^*, T_n(\lambda_n, \mu), T_p(\lambda_p, \mu)) = \frac{{\lambda_p}^2}{f(c^*)\, {c^*}'_P} = \frac{(1 - F(c^*))^2}{f(c^*)\, {c^*}'_P} \tag{17}$$

While existence could be shown from the first order condition alone, consideration of the uniqueness of P^* requires engaging with the second order condition as well:

Lemma 3. *It is a sufficient condition for P^* to be a unique local maximum in the range $(0, \bar{P})$ for:*

$$\frac{\partial f(c^*)}{\partial c^*} \left({c^*}'_P\right)^2 + f(c^*) \frac{\partial^2 c^*}{\partial P^2} > 0.$$

Lemma 3 presents only a sufficient condition, so one cannot guarantee uniqueness for distributions which do not meet the conditions presented. However, numerical simulations have failed to produce a counter-example where P^* is not an unique maximum on the relevant range, even when the sufficient conditions set out in the lemma were not met.

5. Welfare

In the present context, welfare for the consumers seeking treatment from provider i can be defined as the gain in disease-free time from seeking treatment $(\tau - T_i)$, across all consumers seeking treatment from that provider, i.e., multiplied by the expected value of c for each demand stream. To this is subtracted the price of seeking treatment in the case of those consumers choosing the private sector. Therefore, welfare accruing to consumers in unit time, W_i, $i \in \{n, p\}$, is given by the following expressions:

$$W_n(c^*, T_n) = \left(\int_0^{c^*} c f(c)\, \mathrm{d}c \right) (\tau - T_n) \Rightarrow \tag{18}$$

$$W_p(c^*, T_p, P) = \left(\int_{c^*}^{\bar{c}} c f(c)\, \mathrm{d}c \right) (\tau - T_p) - \lambda_p P. \tag{19}$$

Let W be aggregate social welfare, obtained from the sum of the firm's profit and the welfare of the two demand streams. The second term in W_p cancels out the firm's profit, so that W is given by:

$$
\begin{aligned}
W = W_n + W_p + \pi &= \left(\int_0^{c^*} cf(c)\,\mathrm{d}c\right)(\tau - T_n) + \left(\int_{c^*}^{\bar{c}} cf(c)\,\mathrm{d}c\right)(\tau - T_p) \Rightarrow \\
W(c^*, T_n, T_p) &= \left(\int_0^{\bar{c}} cf(c)\,\mathrm{d}c\right)\tau - \left(\int_0^{c^*} cf(c)\,\mathrm{d}c\right)T_n - \left(\int_{c^*}^{\bar{c}} cf(c)\,\mathrm{d}c\right)T_p. \qquad (20)
\end{aligned}
$$

The first term $\left(\int_0^{\bar{c}} cf(c)\,\mathrm{d}c\right)\tau = E[c]\tau$ is the expected disutility of seeking no treatment, for all consumers. It is a constant term determined exogenously, the product of the expected value of c across the consumer population and expected lifespan τ. The two subsequent terms are the product of the expected value of the share of consumers seeking demand from each provider, and the expected waiting time for that provider, so that as waiting time falls, disease-free time increases, as does welfare. Equations (18)-(20) can be re-written using integration by parts, for which, see Appendix B. These alternate expressions are more mathematically tractable, while those presented in the main body are more intuitively understandable.

Let $W'_P \equiv \frac{\partial W(c^*(P, T_n(\lambda_n,\mu), T_p(\lambda_p,\mu)), T_n(\lambda_p,\mu), T_n(\lambda_p,\mu))}{\partial P}$, the total derivative of W with regard to P. Further, let P^W be the social welfare maximizing price, such that $W'_P|_{P=P^W} = 0$; substituting this value back into W yields the largest possible social welfare W^M. Then W'_P is given by:

$$
W'_P = c^{*\prime}_P f(c^*)\left[c^*(T_p - T_n) + T_p^2 E(c \geq c^*)[1 - F(c^*)] - T_n^2 E(c \leq c^*)F(c^*)\right]. \qquad (21)
$$

The last line of equation (21) expresses the tradeoffs at play particularly clearly. By increasing its price, the private provider increases the welfare accruing to those consumers who still choose to buy care from the private sector after the price increase, as the fall in demand causes a fall in expected waiting time (these consumers are represented by the positive term $T_p^2\left(E(c \geq c^*)[1 - F(c^*)]\right)$). On the other hand, welfare is lost from the increase in expected waiting time in the share of consumers seeking care from the public provider (represented by the negative term $-T_n^2\left(E(c \leq c^*)F(c^*)\right)$). (Recall from eq. (9) that the term $c^*(T_p - T_n)$ is the price.) These changes are mediated via the derivative of the critical value, $c^{*\prime}_P$, and the distribution function $f(c^*)$.

Equation (21) can be used to show it is welfare increasing for the private provider to charge a strictly positive price:

Theorem 2. *The welfare maximizing private sector price is strictly positive (i.e. $P^W > 0$) for all well behaved distributions where $f(c^*) > 0\,\forall\, c^* \in (0, \bar{c})$.*

Note, however, that there are positive prices which yield worse outcomes, and this result does not guarantee that the profit maximizing price will improve welfare when compared to a price of 0 on both sectors.

Like several other expressions presented in the foregoing, (21) above is not easily tractable through analytical means as long as $f(c)$ is not specified. Therefore, section 6 below presents more detailed results for a selection of tractable distributions.

6. Numerical Simulations

This section will present numerical results for a set of probability distributions for the disutility of waiting, $f(c)$. The Kumaraswamy distribution will be used for this purpose. This is a bounded continuous distribution, where virtually every continuous shape can be obtained in the interval $[0,1]$ [6] by varying its parameters $a,\, b > 0$ (for instance, the Uniform distribution is a special case where $a = b = 1$). It is much more tractable than the more widely known Beta distribution. The general probability density function, and cumulative distribution function, are defined as follows:

$$f(c; a, b) = a\, b\, c^{a-1}(1 - c^a)^{b-1}, \tag{22}$$
$$F(c; a, b) = 1 - (1 - c^a)^b. \tag{23}$$

Numerical results will be presented for three different specifications: the uniform distribution, a left-skewed distribution, and a right-skewed distribution.

Specifically, define the following three probability distribution functions:

$$\begin{aligned} f_1(c) =& 1 \\ f_2(c) =& 2c \\ f_3(c) =& 2(1 - c), \end{aligned}$$

where $f_1(c)$ is the uniform distribution, $f_2(c)$ is a Kumaraswamy distribution with parameters $a = 2$ and $b = 1$ (left-skewed), and $f_3(c)$ is a Kumaraswamy distribution with parameters $a = 1$ and $b = 2$ (right-skewed).

To these p.d.f.s correspond the following cumulative distribution functions:

$$\begin{aligned} F_1(c) &= c \\ F_2(c) &= c^2 \\ F_3(c) &= 2c - c^2. \end{aligned}$$

See Appendix C for the development of the model for the distributions used.

The tables below show, for each of the three foregoing distributions, the results obtained for waiting times, market share, price, profit and social welfare. For the latter results, let W_n, W_p, and W denote social welfare when the private provider is charging the profit maximizing price P^*, for the share of consumers choosing the

[6] Without loss of generality, as the measure of consumer disutility is arbitrary.

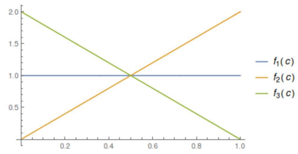

Figure 1.: Probability Distribution Functions $f_1(c)$-$f_3(c)$

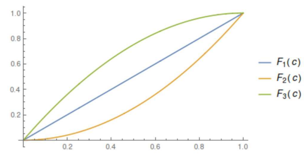

Figure 2.: Cumulative Distribution Functions $F_1(c)$-$F_3(c)$

Table 1.: Numerical Simulations for Distribution Function $f_1(c)$, $\tau = 8$.

μ	T_n	T_p	λ_n	λ_p	P^*	π	W_n	W_p	W	P^W	W^M	W_{2n}
1.2	2.719	0.969	0.832	0.168	1.457	0.244	1.829	0.836	2.909	0.163	3.304	3.286
1.5	1.443	0.765	0.807	0.193	0.547	0.106	2.134	1.157	3.397	0.090	3.510	3.500
2	0.831	0.557	0.796	0.204	0.218	0.045	2.272	1.318	3.635	0.045	3.672	3.667
4	0.312	0.264	0.790	0.210	0.038	0.008	2.399	1.446	3.853	0.010	3.858	3.857

public provider, the share choosing the private provider, and the total, respectively; W^M denotes social welfare when the private provider is charging the social welfare maximizing price P^W, and W_{2n} denotes the social welfare when there are two public providers charging a price of 0 and sharing demand equally. Each table corresponds to one of the distributions. There follows a section discussing the results presented here.

6.1. Discussion

Analysis of the simulation results yields some notable findings. It is noteworthy that, for all distributions, as μ increases, W approaches W^M even as P and π fall, though the latter two values remain strictly positive. This can be intuitively explained as the increase in μ improving the service of the public sector *vis-a-vis* the private sector, therefore forcing the latter to reduce its price. This is in line with the result in eq. (B6) showing λ_p is decreasing in μ for a constant price—clearly, even when the private sector changes its price to respond to the fall in demand, the downward pressure on demand still dominates.

There is a unique value of P^*—in fact, numerical simulations beyond those displayed here were not able to produce any counter-example where multiple equilibria emerged.

Comparing the results for the three distributions, one observes similar movements in welfare, prices and market shares when μ increases. It is interesting to note that

Table 2.: Numerical Simulations for Distribution Function $f_2(c)$, $\tau = 8$.

μ	T_n	T_p	λ_n	λ_p	P^*	π	W_n	W_p	W	P^W	W^M	W_{2n}
1.2	2.596	0.985	0.815	0.185	1.454	0.269	2.650	0.967	3.886	0.122	4.389	4.381
1.5	1.406	0.776	0.789	0.211	0.559	0.118	3.078	1.325	4.522	0.068	4.671	4.667
2	0.818	0.563	0.778	0.222	0.225	0.050	3.284	1.507	4.841	0.034	4.891	4.889
4	0.310	0.265	0.771	0.229	0.039	0.009	3.474	1.653	5.136	0.007	5.143	5.143

Table 3.: Numerical Simulations for Distribution Function $f_3(c)$, $\tau = 8$.

μ	T_n	T_p	λ_n	λ_p	P^*	π	W_n	W_p	W	P^W	W^M	W_{2n}
1.2	2.897	0.948	0.855	0.145	1.206	0.175	1.148	0.589	1.912	0.127	2.208	2.190
1.5	1.490	0.753	0.829	0.171	0.432	0.074	1.363	0.825	2.262	0.071	2.343	2.333
2	0.846	0.550	0.818	0.182	0.170	0.031	1.452	0.941	2.423	0.035	2.449	2.444
4	0.314	0.262	0.811	0.189	0.029	0.006	1.532	1.032	2.569	0.008	2.572	2.571

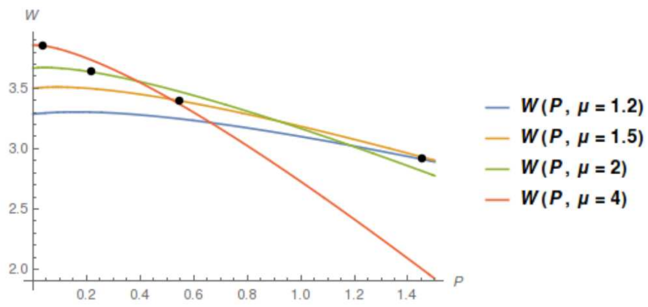

Figure 3.: W as a function of P

the private sector's profit maximizing price is higher for the uniform distribution than for either of the skewed distributions. This indicates the private sector's market power is higher under these circumstances. This is interesting and unexpected. While a full explanation of this phenomenon if left for further research, one intuitively plausible explanation is that there are countervailing forces weighing on the private sector's market power. On the one hand, if there are relatively few consumers with high values of c, then the firm must lower the price to increase revenue directly. On the other hand, there are relatively more consumers with high values of c, the direct effect on revenue of lowering the price by a small amount will be larger. On the other hand, this will have a countervailing effect, as more consumers seek to use the private provider, increasing waiting times and reducing demand. The uniform distribution would seem to be closer to the best distribution for higher profit than either of the skewed distributions.

Turning to social welfare, this is highest on the left-skewed distribution $f_2(c)$, and lowest on the right-skewed distribution $f_3(c)$, presumably because a greater mass of consumers with a lower time cost has less to gain from the private sector's presence. This intuition is confirmed when it is noted $f_3(c)$ presents the lowest market share and profits for the private sector.

When the social welfare is compared with the situation where there are two public providers charging a price of 0, regardless of the distribution, it is the case that the private sector equilibrium price produces an inferior outcome to the case where there are two public providers instead. However, this is always inferior to the socially optimal outcome where the private provider charges price P^W. Figure 3 illustrate this by showing how, for $f_1(c)$, social welfare varies with price for different values of μ, where the dark dot on each line represents welfare at the profit maximizing price P^*.

This suggests an argument for price regulation of the private sector. Presumably this is due to the private sector functioning as an escape valve for consumers with very high disease costs who are willing to pay a lot to be served more quickly.

7. Conclusion

Waiting times affect competition in healthcare in a way that is not adequately captured by traditional models. The present paper, using methods derived from the applications of queueing theory to Industrial Organization and Health Economics, contributes to the understanding of this phenomenon, following the lead of (Farnworth, 2003; Lindsay & Feigenbaum, 1984), among others.

When a profit maximizing private provider acts as a competitor to a public provider of health care, and that private provider faces capacity constraints resulting in queues, in a similar manner to the public sector, but can vary its demand by varying the price, there will always exist at least one price which maximizes the private sector's profit. This equilibrium's uniqueness will depend on the distribution of consumers' cost of waiting time/illness severity.

Comparative statics can be obtained for demand functions for any given distribution of c. Some of the most noteworthy results include that an increase (decrease) in the price charged by the private sector causes an increase (decrease) in demand for the public sector, as well as their waiting time; that a positive price for the private sector is welfare enhancing when compared with free treatment in two providers, and that an increase in the service rate μ decreases waiting times for both sectors, reducing the attractiveness of the private sector and lowering its demand. This forces the private sector to reduce its price, approximating the welfare maximizing value with higher values of μ. The most promising lead for further research is relaxing the assumption that both providers have the same service rates, and allowing at least the private sector to choose its own rate.

Further, the numerical results strongly suggest that the private sector equilibrium price is too high from a social welfare point of view. Further research should explore the effects of the public sector charging a regulated small price compared to a private sector charging a profit maximizing price.

Other warranted topics for further research include, in the first instance, to incorporate consumer income constraints and distribution into the decision process. Other interesting extensions are to incorporate long run capacity decisions by providers, and allowing for a plurality of private (and possibly public) providers. Numerical simulations of model outcomes for different distributions are also warranted, as is a further examination of how intertemporal considerations and discounting might affect consumer decisions.

Disclosure Statement

The author states that there are no conflicts of interest.

References

Andritsos, D. A., & Aflaki, S. (2015). Competition and the operational performance of hospitals: The role of hospital objectives. *Production and Operations Management*, *24:11*, 1812-1832.

Barzel, Y. (1974). A theory of rationing by waiting. *Journal of Law and Economics*, *17:1*, 73-95.

Chen, H., & Frank, M. (2004). Monopoly pricing when customers queue. *IIE Transactions*, *36:6*, 569-581.

Chen, H., & Wan, Y.-W. (2003). Price competition of make-to-order firms. *IIE Transactions*, *35:9*, 817-832.

Cullis, J. G., & Jones, P. R. (1986). Rationing by waiting lists: An implication. *The American Economic Review*, *76:1*, 250-256.

Deck, C., Kimbrough, E. O., & Mongrain, S. (2014). Paying for express checkout: Competition and price discrimination in multi-server queuing systems. *PloS one*, *9:3*, e92070.

Dimakou, S., Dimakou, O., & Basso, H. S. (2015). Waiting time distribution in public health care: empirics and theory. *Health economics review*, *5:1*, 25.

Dube, P., & Jain, R. (2014). Bertrand equilibria and efficiency in markets for congestible network services. *Automatica*, *50:3*, 756-767.

Farnworth, M. G. (2003). A game theoretic model of the relationship between prices and waiting times. *Journal of Health Economics*, *22:1*, 47–60.

Goddard, J. A., Malek, M., & Tavakoli, M. (1995). An economic model of the market for hospital treatment for non-urgent conditions. *Health Economics*, *4:1*, 41-55.

Gravelle, H., & Schroyen, F. (2016). *Optimal hospital payment rules under rationing by random waiting.* (Working Paper)

Gross, D., Shortle, J. F., Thompson, J. M., & Harris, C. M. (2008). *Fundamentals of queueing theory.* Hoboken, New Jersey: John Wiley and Sons, Inc.

Hoel, M., & Sæther, E. M. (2003). Public health care with waiting time: the role of supplementary private health care. *Journal of Health Economics*, *22:4*, 599–616.

Iversen, T. (1993). A theory of hospital waiting lists. *Journal of Health Economics*, *12:1*, 55–71.

Iversen, T. (1997). The effect of a private sector on the waiting time in a national health service. *Journal of Health Economics*, *16:4*, 381–396.

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

Lederer, P. J., & Li, L. (1997). Pricing, production, scheduling, and delivery-time competition. *Operations Research*, *45:3*, 407-420.

Levhari, D., & Luski, I. (1978). Duopoly pricing and waiting lines. *European Economic Review*, *11:1*, 17–35.

Li, L., & Lee, Y. S. (1994). Pricing and delivery-time performance in a competitive environ-

ment. *Management Science*, *40:5*, 633-646.
Lindsay, C. M., & Feigenbaum, B. (1984). Rationing by waiting lists. *The American Economic Review*, *74:3*, 404-417.
Luski, I. (1976). On partial equilibrium in a queuing system with two servers. *The Review of Economic Studies*, *43:3*, 519-525.
Marchand, M., & Schroyen, F. (2005). Can a mixed health care system be desirable on equity grounds? *The Scandinavian Journal of Economics*, *107:1*, 1–23.
Mazalov, V. V., & Melnik, A. V. (2016). Equilibrium prices and flows in the passenger traffic problem. *International Game Theory Review*, *18:01*(01), 1650001.
Melo, E. (2014). Price competition, free entry, and welfare in congested markets. *Games and Economic Behavior*, *83*, 53-72.
Naor, P. (1969). The regulation of queue size by levying tolls. *Econometrica*, *37:1*, 15-24.
Propper, C. (2000). The demand for private health care in the uk. *Journal of Health Economics*, *19:6*, 855–876.
Zhou, C., Albuquerque, P., & Grewal, R. (2016). Competition and firm service reliability decisions: A study of the airline industry. *INSEAD Working Paper*, *2016/18/MKT*.

Appendix A. Equivalence of Exponential and Linear Formulation

Many works addressing waiting times in health care have formulated consumer utility using exponential decay. The following is an example, taken from (Lindsay & Feigenbaum, 1984):

$$U_i = V(\bar{u}, p)e^{-gt}. \tag{A1}$$

In (A1), $V(\bar{u}, p)$ is the value of the good, in this case the treatment, which is a function of a vector of parameters $\bar{u}$, and of price p. This value decays at rate g per unit of time t. Applying this function to the present problem, if $\bar{u}$ is held constant and V is a linear function of p such that $V = v - p$, where v is the value of the good before its price is deducted, (A1) becomes:

$$U_i = (v - p)e^{-gt}. \tag{A2}$$

This can then be transformed by taking logs:

$$\tilde{U_i} = \ln U_i = \ln(v - p) - gt. \tag{A3}$$

Both v and p are arbitrary, so they can be redefined to new values such that $\Lambda = \ln(v - p)$:

$$\tilde{U_i} = \Lambda - gt, \tag{A4}$$

where g is equivalent to parameter c in the utility function at equation (2).

There is still one outstanding issue. The presence of value parameter Λ introduces some mathematical complications. Moreover, as discussed above, medical treatments, especially of chronic diseases, can hardly be said to possess intrinsic value for consumers, unless perhaps they suffer from Münchausen syndrome. Rather, they are valuable in so far as they removes an illness. Therefore, take price p as being paid to remove disutility g. It's perfectly possible to postulate a function V of this kind, say $V = -p$, i.e., there is no benefit from the treatment itself other than it causing g to stop. This yields the following utility function:

$$U_i = -pe^{-gt}. \tag{A5}$$

Taking logs of (A5) yields

$$\tilde{U_i} = \ln U_i = \ln(-p) - gt. \tag{A6}$$

Then once P is defined such that $P = \ln(-p)$, the utility function from (2) emerges.

Appendix B. Proofs and Derivations

Proof of Lemma 1. As c^* is a continuous and strictly increasing function of P (${c^*}'_P > 0$), and $\lambda_p(c^*)$ is a continuous and decreasing function of c^*, as defined at (9), it follows that:

$$\lim_{c^* \to \bar{c}} \lambda_p = 0, \tag{B1}$$

where $\bar{P}$ is such that $c^*(\bar{P}, T_n, T_p) = \bar{c} = F^{-1}(1)$. $\square$

Proof of Lemma 2. This is the reverse of Lemma 1. As c^* is a continuous and strictly increasing function of P (${c^*}'_P > 0$), and $\lambda_p(c^*)$ is a continuous and decreasing function of c^*, as defined at (9), it follows that when $P \to 0$, the two providers are indistinguishable, as none of them charges for treatment and they offer the same service rate μ. In this case the market outcome is a Bertrand equilibrium with the two providers splitting the market equally: $\lim_{P\to 0} c^*(0, T_n, T_p) = F^{-1}\left(\frac{1}{2}\right)$. $\square$

Derivation of Equation (12): Recall ${c^*}'_P$ is the total derivative of c^* with regard to P:

$${c^*}'_P = \frac{1}{T_n - T_p} + \frac{\partial c^*}{\partial T_n}\frac{\partial T_n}{\partial \lambda_n}\frac{\partial \lambda_n}{\partial c^*}\frac{\partial c^*}{\partial P} + \frac{\partial c^*}{\partial T_p}\frac{\partial T_p}{\partial \lambda_p}\frac{\partial \lambda_p}{\partial c^*}\frac{\partial c^*}{\partial P} > 0.$$

Solving, it yields equation (12):

$${c^*}'_P = \left[(T_n - T_p) + c^* f(c^*)(T_n^2 + T_p^2)\right]^{-1} > 0, \tag{B2}$$

as

$$\frac{\partial c^*}{\partial T_p} = -\frac{\partial c^*}{\partial T_n} = \frac{P}{(T_n - T_p)^2}; \; \frac{\partial T_i}{\partial \lambda_i} = T_i^2, i = \{n, p\};$$
$$\frac{\partial \lambda_n}{\partial c^*} = -\frac{\partial \lambda_p}{\partial c^*} = f(c^*); \text{ and } (T_n - T_p) > 0.$$

Derivation of Equation (13): Recall ${c^*}'_\mu$ is the total derivative of c^* with regard to μ. Then:

$${c^*}'_\mu = \frac{\partial c^*}{\partial T_n}\left(\frac{\partial T_n}{\partial \mu} + \frac{\partial T_n}{\partial \lambda_n}\frac{\partial \lambda_n}{\partial c^*}\frac{\partial c^*}{\partial \mu}\right) + \frac{\partial c^*}{\partial T_p}\left(\frac{\partial T_p}{\partial \mu} + \frac{\partial T_p}{\partial \lambda_p}\frac{\partial \lambda_p}{\partial c^*}\frac{\partial c^*}{\partial \mu}\right) > 0, \forall\ P > 0.$$

Solving, it yields equation (13):

$${c^*}'_\mu = \frac{P\left(T_n^2 - T_p^2\right)}{(T_n - T_p)^2 + Pf(c^*)\left(T_n^2 + T_p^2\right)} > 0, \tag{B3}$$

where it follows from the waiting time expression at (3) that:

$$\frac{\partial T_i}{\partial \mu} = -\frac{1}{(\mu - \lambda_i)^2} = -T_i^2 < 0.$$

Comparative Statics: Let ${\lambda_p}'_P \equiv \frac{\partial \lambda_p(c^*(P, T_n(\lambda_n, \mu), T_p(\lambda_p, \mu)))}{\partial P}$, ${\lambda_n}'_P \equiv \frac{\partial \lambda_n(c^*(P, T_n(\lambda_n, \mu), T_p(\lambda_p, \mu)))}{\partial P}$, ${\lambda_p}'_\mu \equiv \frac{\partial \lambda_p(c^*(P, T_n(\lambda_n, \mu), T_p(\lambda_p, \mu)))}{\partial \mu}$, ${\lambda_n}'_\mu \equiv \frac{\partial \lambda_n(c^*(P, T_n(\lambda_n, \mu), T_p(\lambda_p, \mu)))}{\partial \mu}$, the derivatives of λ_p and λ_n in regard to P and μ, respectively.

$${\lambda_p}'_P = \frac{\partial \lambda_p}{\partial c^*}\frac{\partial c^*}{\partial P} = -f(c^*){c^*}'_P < 0, \tag{B4}$$
$${\lambda_n}'_P = \frac{\partial \lambda_n}{\partial c^*}\frac{\partial c^*}{\partial P} = f(c^*){c^*}'_P > 0, \tag{B5}$$
$${\lambda_p}'_\mu = \frac{\partial \lambda_p}{\partial c^*}\frac{\partial c^*}{\partial \mu} = -f(c^*){c^*}'_\mu < 0, \tag{B6}$$
$${\lambda_n}'_\mu = \frac{\partial \lambda_n}{\partial c^*}\frac{\partial c^*}{\partial \mu} = f(c^*){c^*}'_\mu > 0. \tag{B7}$$

Proof of Theorem 1. At least one value P^* exists if the first order condition (15) has at least one zero in the domain $P \in (0, \bar{P})$, where $\bar{P}$ is the value for which $\lambda_p = 0$, or $c^* = \bar{c}$. When $P \to 0$, (15) takes the form:

$$\lim_{P \to 0}\left[(1 - F(c^*)) + P{\lambda_p}'_P\right] > 0,$$

whereas when $P \to \bar{P}$, $c^* \to \bar{c}$, and (15) takes the form:

$$\lim_{P \to \bar{P}}\left[(1 - F(\bar{c})) + \bar{P}{\lambda_p}'_P\right] = 0 - \bar{P}f(\bar{c}){c^*}'_P < 0.$$

Since the FOC takes a positive value at one end of the domain, a negative value at the other, and is continuous across the domain, the intermediate value theorem implies there is at least one zero. □

Proof of Lemma 3. If the derivative of the first order condition is negative across the specified range, P^* will be unique in that range. This follows from the proof of theorem 1: as one end of the domain is negative and the other positive, and the FOC is continuous, if its derivative is negative it will only have one zero along that domain.

This condition is given by:

$$-2f(c^*){c^*}'_P - P\left[\frac{\partial f(c^*)}{\partial c^*}\left({c^*}'_P\right)^2 + f(c^*)\frac{\partial^2 c^*}{\partial P^2}\right] < 0, \tag{B8}$$

evaluated at P^*. As ${c^*}'_P > 0$, (B8) holds under the condition presented in lemma 3. □

Alternate Versions of Welfare Functions: Both $\int_0^{c^*} cf(c)\,\mathrm{d}c$ and $\int_{c^*}^{\bar{c}} cf(c)\,\mathrm{d}c$ can be rewritten using integration by parts:

$$\int_0^{c^*} cf(c)\,\mathrm{d}c = c^*F(c^*) - \int_0^{c^*} F(c)\,\mathrm{d}c$$
$$\int_{c^*}^{\bar{c}} cf(c)\,\mathrm{d}c = \bar{c} - c^*F(c^*) - \int_{c^*}^{\bar{c}} F(c)\,\mathrm{d}c.$$

Then W_i and W can be written as follows:

$$W_n = \left(c^*F(c^*) - \int_0^{c^*} F(c)\,\mathrm{d}c\right)(\tau - T_n) \tag{B9}$$

$$W_p = \left(\bar{c} - c^*F(c^*) - \int_{c^*}^{\bar{c}} F(c)\,\mathrm{d}c\right)(\tau - T_p) - \lambda_p P \tag{B10}$$

$$\begin{aligned} W = &\left(\bar{c} - \int_0^{\bar{c}} F(c)\,\mathrm{d}c\right)\tau + c^*F(c^*)(T_p - T_n) - \bar{c}T_p + \\ &\left(\int_0^{c^*} F(c)\,\mathrm{d}c\right)T_n + \left(\int_{c^*}^{\bar{c}} F(c)\,\mathrm{d}c\right)T_p. \end{aligned} \tag{B11}$$

The final expression is especially useful to determine how welfare responds to changes in price and service rate, as $\bar{c}$ and τ are constants, while the integrand is the cumulative distribution function. Nevertheless, while the foregoing set is more tractable, it is preferable to use eqs. (18)-(20) for intuitive reasoning about welfare, as the interpretation of eqs. (B9)-(B11) is not straightforward.

Derivation of Equation (21)

Recall ${W'}_P$ is the total derivative of W with regard to P, and that P^W is the social welfare maximizing price. Then ${W'}_P$ is given by:

$$
\begin{aligned}
&W'_P = \left(c^{*\prime}{}_P F(c^*) + c^* f(c^*) c^{*\prime}{}_P\right)(T_p - T_n) + c^* F(c^*)\left(\frac{\partial T_p}{\partial \lambda_p}\frac{\partial \lambda_p}{\partial c^*} c^{*\prime}{}_P - \frac{\partial T_n}{\partial \lambda_n}\frac{\partial \lambda_n}{\partial c^*} c^{*\prime}{}_P\right)\\
&- \bar{c}\frac{\partial T_p}{\partial \lambda_p}\frac{\partial \lambda_p}{\partial c^*} c^{*\prime}{}_P + \frac{\partial}{\partial P}\left(\int_0^{c^*} F(c)\,\mathrm{d}c\right) T_n + \left(\int_0^{c^*} F(c)\,\mathrm{d}c\right)\frac{\partial T_n}{\partial \lambda_n}\frac{\partial \lambda_n}{\partial c^*} c^{*\prime}{}_P\\
&+ \frac{\partial}{\partial P}\left(\int_{c^*}^{\bar{c}} F(c)\,\mathrm{d}c\right) T_p + \left(\int_{c^*}^{\bar{c}} F(c)\,\mathrm{d}c\right)\frac{\partial T_p}{\partial \lambda_p}\frac{\partial \lambda_p}{\partial c^*} c^{*\prime}{}_P \Leftrightarrow\\
&W'_P = c^{*\prime}{}_P f(c^*)\Bigg[c^*(T_p - T_n) - T_p^2\left(c^* F(c^*) - \bar{c} + \left(\int_{c^*}^{\bar{c}} F(c)\,\mathrm{d}c\right)\right) +\\
&T_n^2\left(\left(\int_0^{c^*} F(c)\,\mathrm{d}c\right) - c^* F(c^*)\right)\Bigg].\\
&W'_P = c^{*\prime}{}_P f(c^*)\left[c^*(T_p - T_n) + T_p^2\left(E(c \geq c^*)[1 - F(c^*)]\right) - T_n^2\left(E(c \leq c^*) F(c^*)\right)\right].
\end{aligned}
\tag{B12}
$$

Proof of Theorem 2. Recall the derivative of social welfare w.r.t. price given at eq. (21):

$$
W'_P = c^{*\prime}{}_P f(c^*)\left[c^*(T_p - T_n) + T_p^2 E(c \geq c^*)[1 - F(c^*)] - T_n^2 E(c \leq c^*) F(c^*)\right].
$$

If the private sector provider were to charge a price of 0, both providers would split the market equally, and their expected waiting times would be identical ($T_n = T_p = T$), as per Lemma 2. This implies that $W'_P|_{P=0}$ is given by:

$$
W'_P\Big|_{P=0} = c^{*\prime}{}_P f(c^*) T \frac{1}{2}\left[E(c \geq c^*) - E(c \leq c^*)\right]. \tag{B13}
$$

As the term in brackets is obviously positive, and as per eq. (13), so is $c^{*\prime}{}_P$, $W'_P|_{P=0} > 0$ for all well behaved distributions where $f(c^*) > 0\,\forall\, c^* \in (0, \bar{c})$.

This implies that it is always welfare increasing to raise P above 0, thus completing the proof. □

Appendix C. Results for Selected Distributions

This section presents the development of the model described in the main body for the distributions used in the numerical simulations.

C.1. Uniform Distribution

Consider first a uniform distribution for $c \in [0, 1]$, such that:

$$
f(c) = 1 \tag{C1}
$$

$$
F(c) = c. \tag{C2}
$$

Under this distribution, demand and waiting time functions take the following forms:

$$\lambda_n = c^* \ , \qquad\qquad \lambda_p = 1 - c^*,$$
$$T_n = \frac{1}{\mu - c^*} \ , \qquad\qquad T_p = \frac{1}{\mu - (1 - c^*)},$$

which taken together form a system of four equations in four unknowns, easily solvable analytically.

The derivative of c^* with regard to price follows easily from c^*:

$$c^{*\prime}{}_P = \left[T_n - T_p + c^*(T_n^2 + T_p^2)\right]^{-1},$$

which allows analytical derivation of the comparative statics outlined in eqs. (B4)-(B5):

$$\begin{aligned}\lambda_p{}'{}_P &= -c^{*\prime}{}_P = -\left[T_n - T_p + c^*(T_n^2 + T_p^2)\right]^{-1}\\ \lambda_n{}'{}_P &= c^{*\prime}{}_P = \left[T_n - T_p + c^*(T_n^2 + T_p^2)\right]^{-1}.\end{aligned}$$

The private provider's optimization problem is presented and solved below:

$$\begin{aligned}&\max_P \pi = \max_P \lambda_p P = \max_P \left(1 - c^*\right) P\\ &\pi'{}_P = 0 \Leftrightarrow \lambda_p + P\lambda_p{}'{}_P = 0\\ &P^* = (1 - c^*)\left(T_n - T_p + c^*(T_n^2 + T_p^2)\right).\end{aligned}$$

Once P^* is known, equilibrium profit follows easily:

$$\pi = P^*\lambda_p(P^*) = (1 - c^*)^2\left(T_n - T_p + c^*(T_n^2 + T_p^2)\right).$$

For price P^*, welfare levels W_n, W_p, and W are:

$$\begin{aligned}W_n &= \frac{1}{2}(c^*)^2(\tau - T_n)\\ W_p &= \frac{1}{2}\left(1 - (c^*)^2\right)(\tau - T_p) - \pi\\ W &= \frac{1}{2}\left[\tau - T_n(c^*)^2 + T_p\left((c^*)^2 - 1\right)\right].\end{aligned}$$

The derivative of W with regard to P is then as follows:

$$W'{}_P = c^{*\prime}{}_P\left[T_p\left(c^* + \frac{T_p}{2}\left(1 - (c^*)^2\right)\right) - T_n\left(c^* + \frac{T_n}{2}\left(c^*\right)^2\right)\right].$$

This can be set equal to 0 and solved for P, yielding P^W, the social welfare maximizing price.

When c is uniformly distributed, the uniqueness of P^* follows straightforwardly:

Lemma 4. *When c is uniformly distributed, P^* is unique.*

Proof. Recall the sufficient condition for uniqueness given at Lemma 3. This is met when $f(c)$ is the uniform distribution, as the first term $\partial f(c^*)/\partial c^* \left(c^{*\prime}{}_P\right)^2 = 0$ and the second term is positive, as shown below:

$$f(c^*)\frac{\partial^2 c^*}{\partial P^2} = (c^{*\prime}{}_P)^{-1}\left[2(T_n^2+T_p^2)+2c^*(T_n^3-T_p^3)\right] > 0, \tag{C3}$$

where $(T_n^3 - T_p^3) > 0$ since $T_n > T_p$. □

C.2. Kumaraswamy distributions

This subsection performs the same exercise as above, but for a Kumaraswamy distribution, as defined in eqs. (22)-(23).

Under this distribution, demand and waiting time take the following forms:

$$\lambda_n = 1-(1-(c^*)^a)^b\ , \qquad \lambda_p = (1-(c^*)^a)^b,$$
$$T_n = \frac{1}{\mu - [1-(1-(c^*)^a)^b]}\ , \qquad T_p = \frac{1}{\mu-(1-(c^*)^a)^b},$$

which taken together form a system of four equations in four unknowns, easily solvable analytically.

The derivative of c^* with regard to price follows:

$$c^{*\prime}{}_P = \left[T_n - T_p + a\,b\,(c^*)^a(1-(c^*)^a)^{b-1}(T_n^2+T_p^2)\right]^{-1},$$

which allows analytical derivation of the comparative statics outlined in eqs. (B4)-(B5):

$$\lambda_p{}'_P = a\,b\,(c^*)^{a-1}(1-(c^*)^a)^{b-1}\left[T_n - T_p + a\,b\,(c^*)^a(1-(c^*)^a)^{b-1}(T_n^2+T_p^2)\right]^{-1}$$
$$\lambda_n{}'_P = a\,b\,(c^*)^{a-1}(1-(c^*)^a)^{b-1}\left[T_n - T_p + a\,b\,(c^*)^a(1-(c^*)^a)^{b-1}(T_n^2+T_p^2)\right]^{-1}.$$

The private provider's optimization problem is presented and solved below:

$$\max_P \pi = \max_P \lambda_p P = \max_P \left(1-(c^*)^a)^b\right) P$$
$$\pi'_P = 0 \Leftrightarrow \lambda_p + P\lambda_p{}'_P = 0$$
$$P^* = (1-(c^*)^a)^b\left(\frac{T_n - T_p}{a\,b\,(c^*)^{a-1}(1-(c^*)^a)^{b-1}} + c^*(T_n^2+T_p^2)\right).$$

Once P^* is known, equilibrium profit follows easily:

$$\pi = P^*\lambda_p(P^*) = (1-(c^*)^a)^{2b}\left(\frac{T_n - T_p}{a\,b\,(c^*)^{a-1}(1-(c^*)^a)^{b-1}} + c^*(T_n^2+T_p^2)\right).$$

When price is P^*, welfare levels W_n, W_p, and W are:

$$
\begin{aligned}
W_n =& \left(\frac{B\left((c^*)^a,\, a^{-1},\, b+1\right)}{a} - (1-(c^*)^a)^b c^* \right)(\tau - T_n),\\
W_p =& \left((1-(c^*)^a)^b c^* - \frac{B\left((c^*)^a,\, a^{-1},\, b+1\right)}{a} + \frac{\Gamma(a^{-1}+1)\Gamma(b+1)}{\Gamma(1+a^{-1}+b)} \right)(\tau - T_p) - \pi,\\
W =& \frac{\Gamma(a^{-1}+1)\Gamma(b+1)}{\Gamma(1+a^{-1}+b)}\tau - c^*(1-(c^*)^a)^b(T_p - T_n) - T_n \frac{B\left((c^*)^a,\, a^{-1},\, b+1\right)}{a} + \\
& T_p \left(\frac{B\left((c^*)^a,\, a^{-1},\, b+1\right)}{a} - \frac{\Gamma(a^{-1}+1)\Gamma(b+1)}{\Gamma(1+a^{-1}+b)} \right),
\end{aligned}
$$

where $B(x, y, z)$ is the incomplete Beta function, and $\Gamma(x)$ is the Gamma function.

The derivative of W with regard to P is then as follows:

$$
\begin{aligned}
W'_P =& c^{*\prime}{}_P f(c^*) \Bigg[c^*\left(T_p - T_n\right) + T_n^2 \left(\left(s - \frac{B\left((c^*)^a,\, a^{-1},\, b+1\right)}{a} \right) - c^* F(c^*) \right)\\
& - T_p^2 \left(c^* F(c^*) - 1 + \left(1 - s + \frac{B\left((c^*)^a,\, a^{-1},\, b+1\right)}{a} - \frac{\Gamma\left(1+a^{-1}\right)\Gamma(b+1)}{\Gamma\left(1+a^{-1}+b\right)} \right) \right) \Bigg],
\end{aligned}
$$

This can be set equal to 0 and solved for P, yielding P^W, the social welfare maximizing price.