Pricing, Quality and Competition at On-Demand Healthcare Service Platforms

Yixuan Liu  
McCombs School of Business, The University of Texas, Austin, yixuan.liu@utexas.edu

Xiaofang Wang  
School of Business, Renmin University, China, xiaofang.wang@ruc.edu.cn

Stephen Gilbert  
McCombs School of Business, The University of Texas, Austin, steve.gilbert@mccombs.utexas.edu

Guoming Lai  
McCombs School of Business, The University of Texas, Austin, guoming.lai@mccombs.utexas.edu

Abstract

Problem definition: We consider on-demand healthcare platforms that allow patients to seek care online from distributed doctors. Academic/Practical relevance: Healthcare costs have been steadily increasing, while patient experience continues to sour with costly (many times unnecessary) commute and waiting. To alleviate the costs, various on-demand healthcare platforms have emerged but have been little investigated in academic research. Methodology: We develop a strategic queueing model where the platform decides the commission rate upon which potential doctors make their participation, service quality and pricing decisions and potential patients make their service acquisition decisions independently. Results: We find that in equilibrium a higher commission rate always lowers doctor participation as well as service quality, but it may increase the service price if it significantly softens the competition. Moreover, as the quality sensitivity of the patients increases, the service quality improves, accompanied largely with a higher price. We further investigate the effect of platform price control. We find that allowing the platform to control the service price in addition to the commission rate may result in more doctor participation, higher service quality and price, higher platform profit, and surprisingly even higher profit for the doctors. This generally occurs when the quality sensitivity of the patients is either low or high, the waiting cost is low, or the doctor heterogeneity is low. Managerial implications: Our results are useful to understand the performance of on-demand healthcare platforms.
1 Introduction

While the healthcare costs have been steadily increasing worldwide, reaching $7.7 trillion in 2015 (Emergo, 2016), common problems of conventional healthcare systems that rely on face-to-face interaction between a patient and a doctor continue to affect the patient experience. Such problems include the access of doctors, the prolonged waiting time and the unpleasantness of clinic and hospital visits. In the developed countries such as Canada, the UK and the U.S., the waiting time for doctor appointments can range from days to months (or even longer), whereas, in the developing countries such as China and India, many major healthcare facilities are overcrowded with poor environment. To readily access healthcare services is often challenging through the conventional healthcare service systems, especially for the rural areas. In addition, visiting healthcare facilities itself can be costly. As estimated by a healthcare research (Ray et al., 2015), during the period of 2003-2010, it took on average 121 minutes each time a patient seeking medical care in the U.S., including 37 minutes of travel time and 84 minutes of clinic time; the opportunity cost per medical visit was estimated at $43 in contrast to an average $32 out-of-pocket co-pay.

One contributor to such unsatisfactory patient experience is the misuse of the healthcare resources. For instance, in 2015, about 83.2% of the adults in the U.S. had contact with a health professional, with 125.7 million outpatient hospital visits and 928.6 million physician office visits, about 3 visits per person; however, studies show that over 70% of such visits were unnecessary and could be mitigated by remote consultations (Cowell, 2016). Another contributor is the mismatch between the healthcare resources and the demand. The best healthcare resources are often concentrated in the large cities, especially in the developing countries. For instance, in China, most of the renowned doctors are affiliated with the large public hospitals in the big cities. The eagerness for high-quality services and the lack of confidence in the small healthcare providers drive most of the patients to the large hospitals. It results in an imbalance between the healthcare resources and the patient demand across the regions and hospitals.

In the face of the above problems, on-demand service platforms have emerged in the healthcare industry, just like in the transportation and hospitality sectors (e.g., Uber, Airbnb). Examples include the ChunYu-Doctor mobile platform in China, and the Amwell and Doctor-on-Demand mobile platforms in the U.S. These platforms mainly focus on primary care and consultation services. The doctors registered on these platforms can either be affiliated with local clinics and hospitals, providing such online services during their spare time, or conduct services primarily on the platforms. Most of the doctors are board certified, some being renowned specialists (many
doctors that provide services on the ChunYu-Doctor platform in China are renowned specialists, affiliated with the large public hospitals). The doctors can conduct medical diagnoses, prescribe medicines, and provide advice and referrals to the local healthcare facilities. On some platforms such as ChunYu-Doctor, the patients can choose one-time services with text or image messages or video calls, or choose subscription service plans, at different prices set by individual doctors. Whereas, some platforms such as Amwell provide only the video call option at a uniform price. Most platforms provide detailed introductions of the doctors as well as patient reviews. Although the platforms may differ in the way they operate, they share a common goal to connect the distributed healthcare resources and patient demand with convenient and timely services.

On these healthcare platforms, the doctors are independent individual contractors. They provide services in exchange for fees, and they compete against each other in the same category of care. Different from the services provided on the ride-sharing and hospitality platforms, for the healthcare services, the service time influences not only the waiting time but also the service quality. Although a longer service time increases the average waiting time, it may entail more detailed inquiries and more in-depth diagnosis, which can result in a higher service quality. Therefore, on the healthcare platforms that allow the patients to choose their doctors, they will trade off the service quality against the service fee and the waiting time. On the other hand, the platform owners usually profit from the commissions they charge for the services delivered on their platforms. As a result, the competition between the doctors is affected not only by the patients’ sensitivities to the service quality and waiting time but also by the platform commission rate. These factors together further determine the number of participating doctors as well as the number of patients requesting services. While on-demand service platforms have been rapidly growing, related research has just emerged in the operations field. Most of the existing studies focus on the commute and hospitality sectors. These studies typically take the platform’s perspective and investigate either dynamic matching between supply and demand or the optimal service prices and wages to balance supply and demand. However, the decentralized pricing and service quality competition at on-demand healthcare platforms has not yet been investigated.

To fill this gap, we consider one representative service listed on an on-demand healthcare platform. Although the quality of healthcare service may depend upon both a doctor’s skill level and the amount of time spend with a patient, to focus on the trade-off between congestion and quality, we assume that all doctors have the same level of skill, but may have different reservation profits. Given a commission rate charged by the platform, the doctors make their participation decisions
simultaneously, and those who join the platform compete on both price and service quality. The potential patients, who are sensitive to the service quality as well as the waiting time, arrive continually. We consider homogeneous patients who have the same reservation utility in our model (the extended analysis with heterogeneous patients is available upon request). They will purchase the service from a doctor if the perceived utility from the service, after factoring in the service quality, service fee and waiting cost, is the highest among all doctors and also greater than their reservation utility. While the price can be observed directly, the quality can often be inferred from the doctor’s reputation as well as the patient reviews. We consider unobservable queues where the patients use the long-run average waiting time to derive the waiting cost. This is the case at the platforms like Chunyu-Doctor where the number of patients in the queue is not displayed. Although some platforms may display the number of waiting patients, observable queueing games are in general very challenging to solve. It is well documented in prior literature that there is no equilibrium even for the simplest pricing competition game with observable queues (Hassin and Haviv, 2003). Nevertheless, the unobservable queueing model we study can provide a reasonable proxy to deliver managerial insights.

We first analyze the case where the commission rate is fixed (such a setting might be reasonable for platforms with a great variety of services so that to have complete knowledge of the characteristics of all the services and to customize the commission rate targeting each service are practically or legally challenging). We derive the distributive equilibrium where the doctors make their participation, pricing and service time decisions independently, and then the patients make their service acquisition decisions. The equilibrium can be divided into two scenarios with either partial or full market coverage. When the market is partially covered, each participating doctor is in essence a local monopolist; whereas, when the market is fully covered, the doctors compete for the common patient pool. Our analysis reveals that with a higher commission rate, fewer doctors will participate on the platform and their service quality will be lower. However, the service price may either increase or decrease. In particular, when the commission rate is small, an increase of the commission softens the competition, which allows the remaining doctors to increase their price. When the commission rate is intermediate, a further increase of the commission does not soften much the competition. On the other hand, the service quality becomes lower, and thus the remaining doctors need to lower their price to attract the patients. When the commission rate is large, the market is partially covered and the remaining doctors become local monopolists whose price no longer depends on the commission rate. We also find that a larger variability of the doctors’
reservation profits may either increase or decrease the service quality and the service price. When the doctor participation ratio is low, an increase of the variability will improve participation as well as the service quality. When the participation ratio is high, an increase of the variability will lead to an opposite outcome. The quality sensitivity of the patients affects the equilibrium outcome too. Interestingly, we find that the market is more likely to be fully covered when the perceived quality of care is either very sensitive or very insensitive to the time spent with each patient. When the sensitivity is low, there is a small number of participating doctors but they implement very fast service; when the sensitivity is high, the doctors implement a low service rate to improve quality but the number of participating doctors increase significantly in equilibrium.

To deepen the understanding, we further extend our model to allow the platform to optimize its commission rate and compare it with the setting where the platform controls both the commission rate and the service price, as commonly debated in the articles (e.g., Chen et al., 2018). (In practice, the service prices on the ride-sharing platforms such as Uber, Lyft and Didi are often set by the platforms, while the prices on the hospitality platforms such as Airbnb and HomeAway are often decentralized. For the healthcare platforms, we observe both types. For instance, at the Chunyu-Doctor platform, the prices are decentralized. In contrast, at the platforms such as Amwell, a uniform pricing scheme is implemented, most likely by the platforms.) We find that in both settings, the platform will set the commission rate above some thresholds to constrain the doctor participation ratio, to avoid excessive competition, and that the same equilibrium arises in the two settings when the market is partially covered. However, when the market is fully covered, the platform, if it controls the service price, can use the service price to extract surplus, and thus it can set a lower commission rate to induce more doctor participation. In equilibrium, the service quality and the service price are both higher in the setting with platform price control. More interestingly, not only the platform’s profit, but the doctors’ profits can also be higher when the platform controls the price. Our analysis reveals that these two settings are more likely to differ when the quality sensitivity of the patients is either low or high, the unit waiting cost is low, or the variability of the doctors’ reservation profits is low. When the system parameters vary, the service patterns in the two settings can sometimes be quite different.

The remainder of the paper is organized as follows. Section 2 reviews the related literature. Section 3 presents the model. We analyze the setting with fixed commission rate in section 4, and compare the two settings with and without platform price control in section 5. We conclude in section 6.
2 Related Literature

Our paper is related to the two streams of literature on sharing economy and on healthcare/discretionary services. First, we discuss the papers on sharing economy. Cachon et al. (2017) study the pricing schemes for on-demand service platforms such as Uber and Lyft where the service providers self-schedule their time to work based on their opportunity costs. They show that surge pricing which prices the service dynamically based on the demand conditions and pays the service providers with a fixed commission of the dynamic price can significantly improve the profit of the platform relative to the pricing schemes with a fixed service price and/or fixed wage. In certain scenarios, such dynamic pricing schemes can also benefit the service providers and the consumers. Gurvich et al. (2016) study a similar problem but focus on capacity management in a multi-period setting. The firm decides the total number of agents to attract, and sets the compensation and the cap on the number of agents allowed to work in each period. They find that the self-scheduling flexibility of the agents will reduce the firm’s profit and the service level; moreover, a minimum wage requirement may reduce the number of agents that are allowed to work in certain periods. Hu and Zhou (2017) investigate a setting where the platform faces a set of possible market scenarios and in each scenario the supply and demand can be characterized as functions of the platform’s price and wage decisions. They show that for each given scenario the joint price and wage optimization can be reduced to a one-dimensional problem of solving for the optimal matching quantity, and that the optimal price has a U-shaped relationship with the wage. Moreover, with uncertainty of the market scenarios, the commission contract can be (near) optimal for the platform. Hu and Zhou (2016) also study a two-sided marketplace where the supply and demand arrive in random quantities. They focus on the platform’s dynamic matching decisions in a multiple period setting. Chen and Hu (2017) incorporate strategic behavior in the context where the agents and the customers arriving at a random time may strategically wait to monitor the platform’s price and wage levels before they accept a matching or leave the market.

While most of the above studies focus on quantity matching (analogous to the newsvendor setting), Taylor (2017) uses the M/M/k queueing framework to investigate the effects of consumer delay sensitivity and agent independence on the service platform’s pricing and wage decisions. He finds that delay sensitivity induces the platform to reduce price and increase wage when consumer valuation and agent opportunity cost are certain; however, when they are uncertain, delay sensitivity as well as agent independence can either increase or decrease the optimal price. Similarly, Bai et al. (2017) apply the M/M/k queueing framework to investigate the platform’s pricing schemes.
with delay sensitivity. The platform sets the service price as well as the payout ratio. They find that both the optimal price and the payout ratio increase in demand. Moreover, the payout ratio increases as the capacity decreases or the customers become more sensitive to the waiting time, but the optimal price is not necessarily monotone in either capacity or waiting time. Riquelme et al. (2016) examine the pricing schemes with a two-sided queueing framework, where both the agents and the consumers arrive dynamically, and the platform pays a fixed proportion of the service price to the agents. They compare dynamic pricing, which depends on the number of available agents, with static pricing. They find that static pricing can outperform dynamic pricing, but the latter is more robust to fluctuations in system parameters. Differently, Feng et al. (2018) compare the efficiencies between the on-demand ride-hailing platform and the traditional street-hailing system based on the M/M/k queueing framework. They find that the on-demand platform may sometimes lead to higher waiting time than the traditional street-hailing system.

In the sharing economy literature, most studies assume that the platform makes the pricing and wage decisions, upon which the agents and the customers decide whether to transact. One exception is Allon et al. (2012). In their study, a given group of agents on the platform make their own price decisions, while the consumers decide the agent from whom to seek service. They focus on the role of the platform in facilitating information gathering, operational efficiency, and communication among agents. There also exist several studies (Benjaafar et al., 2015; Jiang and Tian, 2016) that investigate consumers’ purchasing and rental decisions in the presence of product sharing platforms. The rental price is determined either by the platform or by a decentralized market-clearance mechanism. Differently, our study investigates healthcare platforms, where the platform firm decides the commission rate upon which potential doctors make their participation, price and service quality decision and potential patients make their service acquisition decisions. A critical difference from prior literature is that the doctors in our study compete on both price and service rate, while the patients are sensitive not only to the price and waiting time but also to the service quality. We also compare the platforms with and without price control.

Second, our study is related to the literature on healthcare/discretionary services. Wang et al. (2010) study a phone-based nurse line with discretionary services to examine the diagnostic accuracy and congestion tradeoff, while Anand et al. (2011) model the dependence of waiting time and service quality on service duration. Xu et al. (2015) examine the benefit for the service provider to randomly assign service rates, faced with customers who value longer service time but dislike waiting. Tong and Rajagopalan (2014) investigate the optimal price and service time for discretionary services,
where both the service rate and the demand are endogenous functions of the pricing schemes. Cui and Veeraraghavan (2016) study a queueing system where, instead of deciding the service rate, the service provider can determine whether to reveal the true service rate information to the uninformed customers, in order to influence their joining decisions. Their analysis shows that at times it is more beneficial to keep the customers blind. Guo et al. (2016) study a healthcare system in which the service time determines the patients’ readmission rate. They investigate the effects of healthcare reimbursement schemes (fee-for-service or bundle payment) on the healthcare provider’s service rate decision and the patients’ admission decisions. However, most of these studies focus on a centralized firm, instead of a platform with distributed heterogeneous doctors, and they don’t investigate simultaneous price and service competition, facing patients who care about not only price and waiting time but also service quality.

3 The Model

We consider a representative healthcare service listed on an on-demand platform, in which the platform collects a commission equal to fraction $r$ of the service fee that each patient pays. The number of potential doctors who can conduct this service is $N$. Let $i$ index the doctors. We assume that the doctors have the same skill level, and differ only in their reservation profits $\delta_i$. As such, the doctors will join the platform if and only if their expected profits to do so are not less than their reservation profits. Let $I_i \in \{0, 1\}$ indicate doctor $i$’s participation decision.

We assume that the service delivery process follows a standard M/M/1 queueing system at each participating doctor. The potential patients in the market experience sickness randomly and arrive according to a Poisson process with rate $\Lambda$. The patients are homogeneous, whose unit waiting cost is $c$ and reservation utility is zero (the results with patient heterogeneity are available upon request). Suppose doctor $i$ joins the platform ($I_i = 1$). He sets the service rate $\mu_i$ and the service price $p_i$. Healthcare services are typically “patient intensive,” in which a higher service rate implies a lower perceived quality. Therefore, similar as in Anand et al. (2011), we model the value of doctor $i$’s service to a patient by:

$$V(\mu_i) = V_0 + \alpha(\mu_0 - \mu_i),$$

where $V_0$ is the base value of the service, $\mu_0$ is a threshold service rate, and $\alpha$ reflects the patient’s sensitivity to the service quality. All the patients are rational. From the platform, they observe the service prices listed by the doctors and can also learn their service rates. While we assume that
the patients do not observe the queue length, they anticipate the other patients’ decisions to infer the expected waiting time.

We study distributive equilibria in which the doctors make their participation, price and service rate decisions simultaneously and independently and then the patients independently choose whether and where to acquire the service. As in Hassin and Haviv (2003) and Anand et al. (2010), we focus on symmetric queue joining strategies of the patients. Let $\gamma_i(p, \mu, I, \Lambda)$ be the probability of joining doctor $i$’s queue and $\gamma_0(p, \mu, I, \Lambda)$ be the probability of balking. Then, the demand rate for doctor $i$ is $\lambda_i(p, \mu, I, \Lambda) = \gamma_i(p, \mu, I, \Lambda) \Lambda$, under which the expected waiting time follows:

$$W(\mu_i, \lambda_i) = \begin{cases} 
\frac{1}{\mu_i - \lambda_i} & \text{if } \mu_i > \lambda_i, \\
\infty & \text{if } \mu_i \leq \lambda_i.
\end{cases}$$

Factoring in the service price and waiting cost, we can write a patient’s utility to acquire the service from doctor $i$ as $V(\mu_i) - p_i - cW(\mu_i, \lambda_i)$. Apparently, a patient will acquire the service from a doctor where his utility is the highest and no less than zero. Therefore, there are three possible cases:

1. $V(\mu_i) - p_i - cW(\mu_i, 0) < 0, \forall i : I_i = 1$. In this case, no patient will join any queue. Thus, the pure strategy of balking is optimal. That is, $\gamma_0(p, \mu, I, \Lambda) = 1$.

2. $\exists i : I_i = 1, V(\mu_i) - p_i - cW(\mu_i, \Lambda) \geq 0$ and $V(\mu_i) - p_i - cW(\mu_i, \Lambda) \geq V(\mu_j) - p_j - cW(\mu_j, 0), \forall j \neq i : I_j = 1$. In this case, even if all patients join doctor $i$’s queue, they all obtain a non-negative utility which is higher than from joining any other doctor’s queue. Therefore, the pure strategy of joining doctor $i$’s queue is optimal. Specifically, $\gamma_i(p, \mu, I, \Lambda) = 1$.

3. $\exists i : I_i = 1, V(\mu_i) - p_i - cW(\mu_i, 0) \geq 0$, and $\forall j : I_j = 1, \exists k \neq j : I_k = 1, V(\mu_j) - p_j - cW(\mu_j, \Lambda) < V(\mu_k) - p_k - cW(\mu_k, 0)$. In this case, no pure symmetric strategy exists so that the patients randomize their joining probabilities. An equilibrium will be reached when the expected utility from joining any queue is the same, i.e., $V(\mu_i) - p_i - cW(\mu_i, \lambda_i) = V(\mu_j) - p_j - cW(\mu_j, \lambda_j) \geq 0, \forall i, j : I_i = I_j = 1$.

Since the utility of a patient decreases in the demand rate at the doctor where he acquires the service, the queue-joining subgame equilibrium is unique in any of the above cases. Note that the doctors aim to maximize their profits given all the other doctors’ decisions and the corresponding patient choices. Therefore, if, in equilibrium, multiple doctors join the platform, only case 3 is possible under their optimal decisions, and each participating doctor $i$ will set his price and service rate to ensure:
(i) \( V_0 + \alpha(\mu_0 - \mu_i) - p_i - \frac{c}{\mu_i - \lambda_i} \geq 0; \)

(ii) \( V_0 + \alpha(\mu_0 - \mu_i) - p_i - \frac{c}{\mu_i - \lambda_i} = V_0 + \alpha(\mu_0 - \mu_j) - p_j - \frac{c}{\mu_j - \lambda_j}, \forall j \neq i : I_j = 1. \)

Hence, given any commission rate \( r \), a distributive equilibrium \((I^*, p^*, \mu^*)\) can be derived from the following set of simultaneous programs:

**P1:** \( \forall i \in \{1, ..., N\} \)

\[
\max_{I_i, p_i, \mu_i} I_i(1 - r)p_i\lambda_i + (1 - I_i)\delta_i \\
\text{s.t. } V_0 + \alpha(\mu_0 - \mu_i) - p_i - \frac{c}{\mu_i - \lambda_i} \geq 0, \\
V_0 + \alpha(\mu_0 - \mu_i) - p_i - \frac{c}{\mu_i - \lambda_i} = V_0 + \alpha(\mu_0 - \mu_j) - p_j - \frac{c}{\mu_j - \lambda_j}, \forall j \neq i : I_j = 1, \\
\lambda_i + \sum_{j \neq i; I_j = 1} \lambda_j \leq \Lambda, \\
0 \leq \lambda_i < \mu_i, \ \ i = 1, \cdots, N.
\]

Clearly, from the above programs, given \( I^*, p^* \) and \( \mu^* \), the demand rates \( \lambda^* \) in the subgame constitute a Nash equilibrium, satisfying the patients’ participation constraint, incentive compatibility constraint as well as the total demand rate constraint. Given the decisions, \( I_{-i}^*, p_{-i}^* \) and \( \mu_{-i}^* \), of the other doctors and their demand rates \( \lambda_{-i}^* \), \( I_i^* \), \( p_i^* \) and \( \mu_i^* \) constitute doctor \( i \)'s optimal decisions with the corresponding best demand rate \( \lambda_i^* \). Recall that in our model, the doctors are identical except for their reservation profits. Therefore, if there are multiple equilibria, we will always choose the symmetric equilibrium that offers the highest social welfare. For simplicity, the subscript \( i \) is dropped in the remainder of the paper.

### 4 Analysis with Fixed Commission Rate

In this section, we analyze the model with an exogenous commission rate. This may be suitable for the scenarios in practice where the commission rate set by the platform does not change frequently, or where there are many different categories of healthcare services and there are either legal or technical obstacles that prevent customizing the commission rate for each service. In addition, characterizing the equilibrium that is conditional upon the commission rate is a building block for choosing an optimal commission rate.
4.1 Equilibrium Derivation

We first focus on the price and service rate decisions by assuming a fixed number \( n \) of participating doctors. From the queueing model, we know that as the demand rate increases, the waiting cost will increase, which will lower the service value. As a result, the incremental number of patients that can be attracted by reducing the service price will decrease as the price continues decreasing. On the other hand, although an increase in the service rate can reduce the waiting cost, it also reduces the service value, and thus the number of patients that can be attracted will also decrease as the service rate continues increasing. This observation implies that it is not always beneficial for a doctor to reduce the price or increase the service rate to attract more patients. In particular, when \( n \) is sufficiently small relative to the aggregate demand rate, \( \Lambda \), rather than competing for more patients, the doctors may have no incentive to set prices low enough to attract all of the patients to purchase the service. When this occurs, the doctors are local monopolists in the sense that an incremental change in the price or service rate at one of the doctors would have no impact upon the other doctors’ demand rates.

Lemma 1. The optimal monopoly price and service rate are

\[
p_{M} = V_0 + \alpha \mu_0 - 2\sqrt{\alpha c} \quad \text{and} \quad \mu_{M} = \frac{V_0 + \alpha \mu_0}{2 \alpha},
\]

under which the demand rate is

\[
\lambda_{M} = \frac{V_0 + \alpha \mu_0 - 2\sqrt{\alpha c}}{2 \alpha} \quad \text{and} \quad \text{the participating doctor's profit rate is}
\]

\[
\pi_{M} = (1 - \gamma) \alpha \lambda_{M}^2.
\]

In the above lemma, the service rate balances the waiting cost and the service quality, achieving the highest service surplus. Then, the doctors, as local monopolists, use the service price to extract all the surplus. When \( n \) becomes sufficiently large, the market turns to be fully covered and competition arises. In the following, we present a necessary condition for the distributive equilibrium for a given \( n \).

Proposition 1. There exist two thresholds \( n_0 = \frac{\Lambda}{\lambda_{M}} \) and \( n_1 = \frac{1 + \frac{\Lambda}{\lambda_{M}}}{2} + \sqrt{\frac{1 + \frac{\lambda_{M}^2}{\Lambda}}{2}} \) such that the distributive equilibrium \( (p^*, \mu^*, \lambda^*) \) satisfies: i) If \( n < n_0 \), \( p^*(n) = p_{M} \), \( \mu^*(n) = \mu_{M} \), \( \lambda^*(n) = \lambda_{M} \); ii) If \( n_0 \leq n < n_1 \), \( p^*(n) = V_0 + \alpha \mu_0 - \alpha \Lambda/n - 2\sqrt{\alpha c}, \mu^*(n) = \Lambda/n + \sqrt{c/\alpha}, \lambda^*(n) = \Lambda/n \); and iii) If \( n \geq n_1 \), \( p^*(n) = \frac{\alpha \Lambda}{n-1}, \mu^*(n) = \Lambda/n + \sqrt{c/\alpha}, \lambda^*(n) = \Lambda/n \).

Proposition 1 shows two thresholds. When the number of participating doctors, \( n \), is less than the lower threshold, the equilibrium follows the monopoly solution. In this region, neither the service rate nor the service price depends on \( n \). Differently, when \( n \) exceeds the lower threshold, the market becomes fully covered, and each doctor’s market share decreases in \( n \). In equilibrium,
the service rate balances the waiting cost and service quality to achieve the highest service surplus. We can observe that the service rate now decreases in $n$. In other words, as the competition intensifies, the doctors reduce their service rate to improve the service quality. On the other hand, when $n$ is in the middle region, the service price still extracts full surplus, which in fact increases in $n$ as the service value increases. However, when $n$ exceeds the upper threshold, the competition becomes intense so that the doctors start to compete on price. Now the service quality improves, while the service price decreases, as $n$ increases. In this region, the patients are able to retain part of the service surplus. We illustrate this pattern in Figure 1.

![Figure 1: Illustration of equilibrium service price, quality, patient utility and doctor profit. The parameters: $V_0 = 20, \mu_0 = 5, c = 20, \alpha = 10, \Lambda = 10, N = 10, r = 0$.](image)

Next, we characterize the number of participating doctors in equilibrium. To that end, we sort the doctors' reservation profits. Without loss of generality, let $\delta_1 < \cdots < \delta_N$. Since the doctors are otherwise identical, those with large reservation profits will have a weaker incentive to join the platform than the ones with small reservation profits. Therefore, when the expected profit per doctor from the platform decreases as competition intensifies, those with large reservation
profits will first drop out the competition. Note that the optimal service rate is always at the level that balances the waiting cost and the service quality, maximizing the service surplus. Hence, the doctors with small reservation profits will use the price weapon to squeeze the market share left for the ones with large reservation profits.

**Lemma 2.** The doctors whose reservation profits $\delta_i \geq \delta_n$ will not join the platform if those with reservation profits $\delta_i < \delta_n$ set the service price lower than

$$
\hat{p}(n, r) = \begin{cases} 
2\sqrt{\frac{\alpha n \delta_n}{(1-r)(n-1)}} - \frac{\alpha \Lambda}{n-1}, & \text{if } \delta_n \leq \frac{(1-r)n(n-1)(V_0+\alpha \mu_0-2\sqrt{\alpha c})^2}{\alpha(2n-1)^2}, \\
\frac{(2n-1)(V_0+\alpha \mu_0-2\sqrt{\alpha c})}{2(n-1)} - \frac{\sqrt{(V_0+\alpha \mu_0-2\sqrt{\alpha c})^2 - \alpha \delta_n}}{n-1}, & \text{if } \frac{(1-r)n(n-1)(V_0+\alpha \mu_0-2\sqrt{\alpha c})^2}{\alpha(2n-1)^2} < \delta_n \leq \frac{(1-r)(V_0+\alpha \mu_0-2\sqrt{\alpha c})^2}{4\alpha}, \\
\infty, & \text{if } \delta_n > \frac{(1-r)(V_0+\alpha \mu_0-2\sqrt{\alpha c})^2}{4\alpha},
\end{cases}
$$

where $\hat{p}(n, r)$ increases in both $n$ and $r$.

Lemma 2 shows the threshold service price to deter the $n_{th}$ and the subsequent doctors. Clearly, it depends not only on the $n_{th}$ doctor’s reservation profit but also on $n$ and the commission rate $r$.

When $n$ is large, the market share among the doctors is small and thus it would not require a very low price to deter additional doctors from participating. Similarly, when $r$ is large, the revenues that the doctors obtain from the platform for each service are low and thus the incentive for the doctors to join the platform is weak. Given this threshold service price, we can further obtain the demand rate for the doctors who join the platform.

**Lemma 3.** When the doctors whose reservation profits $\delta_i < \delta_n$ set the service price at $\hat{p}(n, r)$, the demand rate at each of these doctors follows

$$
\hat{\lambda}(n, r) = \begin{cases} 
\frac{\Lambda}{n-1}, & \text{if } \delta_n < \frac{\alpha(n-1)(1-r)\lambda^2_{M}}{n}, \\
V_0+\alpha \mu_0-2\sqrt{\alpha c} - \hat{p}(n, r), & \text{if } \delta_n \geq \frac{\alpha(n-1)(1-r)\lambda^2_{M}}{n},
\end{cases}
$$

From the above lemma, we can observe that to deter the $n_{th}$ and the subsequent doctors, the market may not need to be fully covered by the participating doctors. This can arise, for instance, if the $n_{th}$ doctor’s reservation profit $\delta_n$ is large, the commission rate $r$ is high, or the monopoly demand rate is small. Combining the above two lemmas, we can evaluate the revenue for a participating doctor at the threshold service price.

**Lemma 4.** For any $n$ such that $\hat{p}(n, r)\hat{\lambda}(n, r) \leq p^*(n)\lambda^*(n)$, $\hat{p}(n, r)\hat{\lambda}(n, r)$ increases in $n$. 

13
Lemma 4 shows that the revenue \( \hat{p}(n,r) \hat{\lambda}(n,r) \) at the threshold service price is bigger for a bigger \( n \). This result implies that for the doctors who remain to participate in the platform, it is actually more beneficial to deter fewer opponents. This is because to intentionally deter more opponents requires more aggressive pricing, which hurts the revenue. On the other hand, from the results of Proposition 1, we know that as the number of participating doctors increases, the profitability of each doctor decreases. Hence, there is an equilibrium participation level. Let

\[
\hat{n}(r) = \min\{n| p^*(n) \lambda^*(n) = \hat{p}(n,r) \hat{\lambda}(n,r) \}
\]

which is the first intersection of \( \hat{p}(n,r) \hat{\lambda}(n,r) \) and \( p^*(n) \lambda^*(n) \).

**Proposition 2.** For any given commission rate \( r \in [0,1] \), if \( N > n_0 \), then there exists a threshold \( \hat{r} < 1 - \frac{\delta n_0}{\alpha \lambda M} \) such that the equilibrium number of participating doctors \( n^* \) follows:

\[
n^* = \begin{cases} 
\max\{n|(1-r)\alpha \lambda^2_M \geq \delta_n\}, & \text{if } r \geq 1 - \frac{\delta n_0}{\alpha \lambda M}, \\
\hat{n}(r), & \text{if } \hat{r} \leq r < 1 - \frac{\delta n_0}{\alpha \lambda M}, \\
N, & \text{if } r < \hat{r};
\end{cases}
\]

if \( N \leq n_0 \), then \( n^* = \max\{n|(1-r)\alpha \lambda^2_M \geq \delta_n\} \).

Proposition 2 shows the equilibrium number of participating doctors. In particular, when either the commission rate is high or the total number of doctors is small, the equilibrium is held at the monopoly solution. The doctors whose reservation profits are less than the monopoly profit join the platform. When the commission rate is intermediate, the equilibrium number of participating doctors is held at \( \hat{n}(r) \) as formed in the above. Lastly, when the commission rate is low, all the doctors may join the platform.

### 4.2 Sensitivity Analysis

In this subsection, we discuss several sensitivity results of the equilibrium.

**Proposition 3.** As the commission rate \( r \) increases, the number of participating doctors \( n^* \) decreases while the service rate \( \mu^* \) increases. For the service price \( p^* \), there exists a threshold \( \hat{r}_1 \) such that it increases in \( r \) for \( r \in [0,\hat{r}_1] \), decreases in \( r \) for \( r \in (\hat{r}_1,1 - \frac{\delta n_0}{\alpha \lambda M}) \), and equals \( p_M \) for \( r \in \left[1 - \frac{\delta n_0}{\alpha \lambda M}, 1\right] \).

If the commission rate increases, the profit to provide the service on the platform will decrease and thus fewer doctors will participate. With less competition, the demand rate for each participating doctor will (weakly) increase, which leads the doctors to increase their service rate, even though
this adversely affects the quality. Therefore, Proposition 3 implies that all else equal, for those platforms that charge higher commissions, there will be fewer doctors serving on the platforms with lower quality. Moreover, from Proposition 1, we can observe that the total served demand (weakly) increases in the number of participating doctors and thus the market coverage will be lower too. For the service price, however, the effect can vary. When the commission rate is in a low region, the number of participating doctors is large. In such a scenario, facing intensive competition, the doctors compete not only on service but also on price. Since an increase of the commission rate can mitigate the competition, the service price will increase. In contrast, when the commission rate is in a middle region, the doctors compete only on service. Given that a higher commission rate will result in a higher service rate, the service value will decrease and so will the service price that the doctors can charge. When the commission rate exceeds a threshold, the participating doctors become local monopolists and their service price is no longer affected by the commission rate. We illustrate these effects in Figure 2.

It is also of interest to explore how the equilibrium is affected by the quality sensitivity of the patients as well as their unit waiting cost. From our numerical analysis, we find that under an exogenous commission rate, an increase of $\alpha$ (i.e., the service value becomes more sensitive to the service quality) will generally lead to more participating doctors, except in some middle region where the market coverage changes from full coverage to partial coverage (see the left subplot of Figure 3). On the other hand, as shown in the middle and right subplots of Figure 3, when $\alpha$ increases, the service rate decreases while the service price increases; that is, the service will have a higher quality but become more expensive to acquire. For the patients’ waiting cost, we find that when the unit cost $c$ is in a low region, the market is fully covered and an increase of $c$ does not

![Figure 2: Illustration of the effect of $r$. The parameters: $V_0 = 20, \mu_0 = 5, c = 20, \alpha = 10, \Lambda = 25, \delta_i \sim U[10, 30], N = 20.$](image-url)


Figure 3: Illustration of the effect of $\alpha$ when $r = 0.1$ (dotted curves) and $r = 0.3$ (solid curves). The parameters: $V_0 = 20, \mu_0 = 5, c = 20, \alpha = 10, \Lambda = 25, \delta_i \sim U[15, 45], N = 30$.

Figure 4: Illustration of the effect of $c$. The parameters: $r = 0.2, V_0 = 20, \mu_0 = 5, c = 20, \alpha = 10, \Lambda = 25, \delta_i \sim U[15, 45], N = 30$.

change the number of participating doctors (see Figure 4). Moreover, in this region, as waiting becomes more costly for the patients, the doctors will speed up their service, which lowers the waiting time as well as the service quality, but they will not alter their service price. In contrast, when $c$ exceeds a threshold, the doctors now will lower the service price but keep the service rate flat. Given that the revenue decreases, the number of participating doctors starts to decrease in $c$ and so does the market coverage.

For the doctors’ characteristics, when their reservation profits are uniformly distributed, we find that an increase of the distribution dispersion may have opposing effects depending on the participation ratio. In particular, if the given participation ratio is below a threshold, an increase of the distribution dispersion can increase the participation and further reduce the service rate as well as the service price. That is, the service will become cheaper while having a higher quality. In
Figure 5: Illustration of the effect of $\Delta \delta$ when $r = 0.1$ (dotted curves) and $r = 0.3$ (solid curves). The parameters: $V_0 = 20, \mu_0 = 5, c = 20, \alpha = 10, \Lambda = 25, \delta_i \sim U[15, 45], N = 30$.

contrast, if the given participation ratio is already higher than the threshold, now an increase of the dispersion may lead to completely different outcomes. We illustrate these effects in Figure 5.

5 Platform Commission and Price Control

Thus far, we have assumed that the commission rate is exogenous, which may fit some decentralized scenarios where the platform does not have tight control over the services. However, there are other healthcare platforms, such as Amwell, where the services are more standard and the platform may have more control over the services. Therefore, in this section, we investigate such scenarios and compare two settings. The first setting extends the one in the above section, allowing the platform to optimize the commission rate, while doctors continue to set prices independently. In the second setting, we allow the platform to control both the commission rate and the service price, while the participating doctors only make their service rate decisions. This comparison can shed light on the commonly raised question about the degree of platform control (Chen et al., 2018).

5.1 Only Commission Control

To optimize the commission rate, the platform solves: $\max_r r p^*(r) n^*(r) \lambda^*(r)$, based on the equilibrium derived in the previous section. Clearly, the main tradeoff is between the per unit service revenue $r p^*(r)$ and the market coverage $n^*(r) \lambda^*(r)$. From Proposition 1, we can observe that when $n < n_0$, the market coverage increases in $n$, while the service price is independent of $n$. Once $n$ reaches $n_0$, the market becomes always fully covered, while the service price increases in $n$ when $n_0 < n < n_1$ and decreases in $n$ when $n > n_1$. On the other hand, although a higher commission...
rate provides a larger share of the service revenue to the platform, it may decrease the number of participating doctors as well as the service price. Therefore, the platform needs to find a commission rate that leads to an ideal number of participating doctors. The following proposition derives such a solution for the case where the doctors’ reservation profits are evenly distributed.

**Proposition 4.** Suppose $\delta_1, \ldots, \delta_N$ are uniformly distributed in the interval $[\delta_0 - \Delta \delta, \delta_0 + \Delta \delta]$. Then, there exists a threshold $\bar{N}$ such that the optimal commission rate and the subsequent distributive equilibrium satisfy: (i) If $N < \bar{N}$, then

$$
\begin{align*}
    r^* &= \frac{1}{2} - \frac{\delta_0}{2\alpha \lambda_M^2} + \frac{N + 1}{N - 1} \frac{\Delta \delta}{2\alpha \lambda_M^2}, \\
    n^* &= \frac{1}{4} \left[ N + 1 + \frac{1}{\Delta \delta} (N - 1)(\alpha \lambda_M^2 - \delta_0) \right], \\
    p^* &= p_M, \\
    \mu^* &= \mu_M;
\end{align*}
$$

(ii) If $N \geq \bar{N}$, then

$$
\begin{align*}
    r^* &= \{ r | p^*(n^*(r))\lambda^*(n^*(r)) = \hat{\lambda}(n^*(r), r) \hat{\lambda}(n^*(r), r) \}, \\
    n^* \in [n_0, n_1], \\
    p^* &= V_0 + \alpha \mu_0 - \alpha \Lambda / n^* - 2\sqrt{\alpha c}, \\
    \mu^* &= \Lambda / n^* + \sqrt{c/\alpha},
\end{align*}
$$

where $n^*(r)$ follows Proposition 2.

To better explain the above proposition, we plot the profit and utility functions of the parties. From the left subplot of Figure 6, we can observe that the profit of the platform first increases and then decreases in the commission rate, while the total profit of the doctors and the total utility of the patients always decrease in the commission rate. In particular, the patients gain positive utility when the number of participating doctors is greater than the threshold $n_1$ (the region on the left side of $n_1$ in this figure). The above proposition shows that if the platform can customize the commission rate targeting the given service, it will either set a high commission rate such that the number of participating doctors is less than the threshold $n_0$, which leads to the local monopoly outcome, or set an intermediate commission rate such that the number of participating doctors is between the two thresholds $n_0$ and $n_1$. The former arises if the pool of potential doctors is small, and the latter arises otherwise. In either case, all the service surplus is extracted away from the
Figure 6: Illustration of the parties’ profit and utility functions under platform commission and price control. The parameters are: $N = 100, V_0 = 20, \mu_0 = 5, c = 15, \alpha = 10, \Lambda = 25, \delta_0 = 40, \Delta \delta = 10, n_0 = 11.0, n_1 = 11.5$.

patients in equilibrium. That is, when the platform has the control of the commission rate, it will never set a low commission rate that would lead to a large number of participating doctors with intensive competition and leave service surplus to the patients. Clearly, from the figure, we can see that such control benefits the platform but hurts the doctors and patients as a whole.

5.2 Both Commission and Price Control

When the platform controls both the commission rate and the service price, the doctors’ problems can be formulated as follows:

\[ \textbf{P2: } \forall i \in \{1, \ldots, N\} \]

\[ \max_{I_i, \mu_i, \lambda_i} I_i (1 - r) p \lambda_i + (1 - I_i) \delta_i \]

\[ \text{ s.t. } V_0 + \alpha(\mu_0 - \mu_i) - p - \frac{c}{\mu_i - \lambda_i} \geq 0, \]

\[ V_0 + \alpha(\mu_0 - \mu_i) - \frac{c}{\mu_i - \lambda_i} \geq V_0 + \alpha(\mu_0 - \mu_j) - \frac{c}{\mu_j - \lambda_j}, \forall j \neq i : I_j = 1, \]

\[ \lambda_i + \sum_{j \neq i: I_j = 1} \lambda_j \leq \Lambda, \]

\[ 0 \leq \lambda_i < \mu_i, \ i = 1, \ldots, N. \]

We again focus on the symmetric equilibrium. From the above programs, we can determine the
number of participating doctors $n^*(r, p)$, their service rate $\mu^*(r, p)$ and the demand rate $\lambda^*(r, p)$. Then, the platform solves: $\max_{r, p} \ r p n^*(r, p) \lambda^*(r, p)$). Below we obtain the solution for the case where the doctors’ reservation profits are evenly distributed (we use the subscript “$b$” to denote the association with this setting, and $\bar{N}$ below is the same threshold as in Proposition 4).

Proposition 5. Suppose $\delta_1, \cdots, \delta_N$ are evenly distributed in the interval $[\delta_0 - \Delta \delta, \delta_0 + \Delta \delta]$. Then, the optimal commission rate, the optimal service price and the subsequent distributive equilibrium satisfy: (i) If $N < \bar{N}$, then

$$r^*_b = \frac{1}{2} - \frac{\delta_0}{2 \alpha \lambda^2_M} + \frac{N + 1}{N - 1} \frac{\Delta \delta}{2 \alpha \lambda^2_M},$$

$$p^*_b = p_M,$$

$$n^*_b = \frac{1}{4} \left[ N + 1 + \frac{1}{N} (N - 1) (\alpha \lambda^2_M - \delta_0) \right],$$

$$\mu^*_b = \mu_M;$$

(ii) If $N \geq \bar{N}$, then

$$n^*_b : \frac{\alpha \Lambda^2}{n^*_b} - \frac{4 \delta n^*_b \Delta \delta}{N - 1} + \frac{\Delta \delta (N + 1)}{N - 1} - \delta_0 = 0,$$

$$r^*_b = 1 - \frac{\delta n^*_b}{n^*_b (\beta - \alpha \Lambda n^*_b - 2 \sqrt{\alpha c})},$$

$$p^*_b = V_0 + \alpha \mu_0 - \alpha \Lambda / n^*_b - 2 \sqrt{\alpha c},$$

$$\mu^*_b = \frac{\Lambda / n^*_b}{n^*_b} + \sqrt{\frac{c}{\alpha}}.$$

The right subplot of Figure 6 shows the profit and utility functions of the parties when the platform controls both the commission rate and the service price. From the figure, we can observe that similar to the case with only commission control, the platform’s profit first increases and then decreases in the commission rate (the service price is optimized accordingly). The total profit of the participating doctors always decreases in the commission rate, while the utility of the patients remains at zero in the entire region. Therefore, a unique optimal solution can be found from the platform’s perspective, which is provided in Proposition 5.

5.3 Performance Comparison

In this subsection, we compare the performances of the above two settings.
Proposition 6. i) When the pool of potential doctors is small \((N < \bar{N})\), the equilibrium outcomes of the two settings coincide. In particular, the monopoly solution arises in both settings and the market is partially covered; that is, \(\mu_b^* = \mu^* = \mu_M, p_b^* = p^* = p_M, n_b^* = n^* < n_0, n_b^* \lambda_b = n^* \lambda < \Lambda\). The profits and the utilities of all the parties are the same in the two settings.

ii) When the pool of potential doctors is large \((N \geq \bar{N})\), the market is fully covered in both settings. In the setting with both platform commission and price control, the number of participating doctors is larger \((n_b^* > n^*)\), the service rate is lower \((\mu_b^* < \mu^*)\) which implies that quality is higher, and the service price is higher \((p_b^* > p^*)\).

Clearly, if the market is partially covered, the monopoly solution is optimal from the perspective of the doctors as well as the platform. As such, the service price remains the same in these two settings, which implies that whether or not the platform has the price control would not affect the properties of the service. The differences occur when the market is fully covered. As we discussed in the above, if the doctors are allowed to set their own service prices, excessive price competition could arise when the number of participating doctors is large. In that scenario, the patients would obtain a part of the service surplus. As a result, to avoid such an outcome, the platform sets the commission rate high to limit the number of participating doctors. Differently, when the platform controls the service price, it can always use the price to extract the service surplus from the patients. Therefore, the commission rate does not need to be as high as in the setting with only commission control, which can allow a larger number of doctors to participate and to provide higher quality service. The platform is always better off by having more control; interestingly, the participating doctors may also be better off, since the commission rate and the intensity of the service competition decrease. This can be readily observed from Figure 6. Although the patients might be in a more disadvantageous situation if the platform controls the service price, in practice, service subsidies are often offered by many platforms, which can redistribute some of the increased surplus. Hence, the above proposition indicates that when there is abundant doctor supply, letting the platform control the service price can lead to a system-wide improvement in efficiency.

While the above comparison is from the perspective of the pool of doctors, the equilibrium outcomes also depend on the other system parameters. We start with the patient characteristics that measure their sensitivity to service quality and waiting delays.

Proposition 7. There exist two thresholds \(\alpha\) and \(\bar{\alpha}\) such that the market is fully covered if \(\alpha \in [0, \alpha] \cup [\bar{\alpha}, \infty)\), and partially covered otherwise.
The above proposition reveals that market full coverage and thus the differences of the two settings generally arise when the quality sensitivity of the patients (α) is either low or high. In our context, the service value is influenced by the service quality (reflected by \( \mu_0 - \mu \)) while α captures the sensitivity. When α is small, the value that can be added by improved service quality is small, which limits the service price as well as doctor participation. As such, it is optimal to implement fast service to lower the waiting cost and induce large demand at each participating doctor. In contrast, when α is large, the contribution of the service quality to the service value becomes large, which allows for a high service price. Moreover, a large α also means a high marginal effect of the service quality. As a result, it becomes optimal to implement a slow service but induce a large number of participating doctors. In either case, it is more beneficial to cover the market fully than partially, from the platform’s perspective, while it is the reverse when α is intermediate. We illustrate the effects of α in Figure 7. Similar as in the setting with exogenous commission rate that is shown in Figure 3, we also observe here that as α increases, the number of participating doctors first increases and then may decrease, before increasing again. The drop occurs at the transition point where the market coverage changes from full coverage to partial coverage. The participation ratio is higher in the setting where the platform controls the service price when the market is fully covered. The service rate always decreases and thus the service quality improves, as α increases. For the service price, although it generally increases in α, it may also decrease when α is small in the setting where the platform controls the price. From the figure, we can see that, if the platform controls the price, it sets a higher price than the doctors would, and the service rate is lower, which implies a higher quality. Figure 7 also shows that in both settings, the platform sets the commission rate at high levels when α is either small or large but at low levels when α is intermediate, and accordingly the platform’s profit function also exhibits a convex shape in α. The total profit of the doctors depends directly on the number of participating doctors and thus it may drop in some middle region of α. In other words, an increase of the patient sensitivity to time spent with the doctor does not always translate into an increase of the profit for the platform or the profit for the doctors. Comparing these two settings, we can readily observe that the commission rate is lower, while the profits of the platform and the doctors are both higher, if the platform controls the service price.

**Proposition 8.** There exists a threshold \( \bar{c} \) such that when \( c < \bar{c} \), the market is fully covered, and when \( c \geq \bar{c} \), the market is partially covered with the coverage decreasing in \( c \).

For the patients’ unit waiting cost, apparently, the service surplus will decrease as \( c \) increases.
Figure 7: Illustration of the effect of $\alpha$ with dotted curves for only commission control and solid curves for both commission and price control. The parameters are: $V_0 = 20, \mu_0 = 5, c = 20, \Lambda = 25, \delta_0 = 28, \Delta \delta = 10, N = 20$.

Therefore, it is more profitable to serve the whole market when $c$ is small than when it is large, as indicated by the above proposition. More interestingly, we find that the two types of platforms with and without platform price control may have different service patterns, as $c$ varies. From Figure 8, we can observe that when the market is fully covered, as $c$ increases, in the setting with only commission control, the number of participating doctors increases, while the service rate remains constant; in contrast, in the setting with both commission and price control, the number of participating doctors remains constant, while the service rate increases. In other words, the platforms in these two settings combat the increase of the waiting cost in different means. With control of only the commission, the platform enlists more doctors while maintaining the same service rate, which lowers the demand rate and thus the average waiting time at each doctor. With control over both the commission and the service price, the platform maintains the same number of participating doctors but induces them to increase their service rate, which lowers the average waiting time. Clearly, in the second setting, the service quality is sacrificed as the unit waiting cost increases. Figure 8 also shows that the service price, the commission rate and the profit of the platform always decrease in the patient waiting cost. However, it is interesting that an increase of
the patient waiting cost is not always detrimental for the doctors. In particular, in the setting with only commission control, the total profit of the doctors increases in the unit waiting cost when the cost is small, as the platform attracts more doctors to participate. Of course, this benefit varnishes once the platform gains the control of the service price.

Figure 8: Illustration of the effect of $c$ with dotted curves for only commission control and solid curves for both commission and price control. The parameters: $V_0 = 20, \mu_0 = 5, \alpha = 4, \Lambda = 25, \delta_0 = 28, \Delta \delta = 10, N = 20$.

In the following, we examine the effect of doctor characteristics from the perspective of the dispersion of their reservation profits.

**Proposition 9.** There exists a threshold $\Delta \delta$ such that when $\Delta \delta < \Delta \delta$, the market is fully covered, and when $\Delta \delta \geq \Delta \delta$, the market is partially covered.

The above proposition shows that full market coverage generally occurs when the dispersion of the doctors’ reservation profits is small. From Propositions 4-5, we can observe that the number of participating doctors in equilibrium decreases in $\Delta \delta$ when the market is partially covered. In other words, it is more beneficial for the platform to attract more doctors to participate as $\Delta \delta$ decreases, which implies the above result. Therefore, only when the heterogeneity of the doctors (in terms of their reservation profits) is relatively small would the platform price control affect the system.

24
performance. We further numerically examine the sensitivities with respect to the dispersion (see Figure 9). We find that in both settings with and without platform price control, when the quality sensitivity of the patients $\alpha$ is relatively low (the upper six subplots), an increase of the dispersion will increase participation, decrease the service rate, and increase the service price. That is, in this scenario, as the heterogeneity of the potential doctors increases, we can find more doctors providing service on the platform, and their service quality tends to be higher as well as their service price; in contrast, when $\alpha$ is high (the lower six subplots), the patterns are reversed and they appear to be more sensitive to the change of the doctor heterogeneity in the case with platform price control than without. In both scenarios, we can observe that at optimum, the platform always charges a higher commission rate for a larger $\Delta \delta$, which benefits the platform but hurts the doctors. Specifically, as heterogeneity among the doctors’ reservation profits increases, the “elasticity” of doctor participation decreases, i.e., it takes a larger percentage reduction in commission to induce a given percentage increase in doctor participation. In general, this would reduce the platform’s incentive to offer a low commission rate.

6 Conclusion

In conventional health care systems that require face-to-face interaction between patients and doctors, local imbalances between doctors and patients can lead to inefficiencies. To address this issue, emerging on-demand healthcare platforms allow patients to more conveniently seek advice and remote treatment from distributed healthcare resources. This study aims to provide a better understanding of the equilibrium outcomes for such healthcare platforms. Our analysis reveals that a higher platform commission rate always lowers doctor participation so that the service rate decreases, allowing doctors to spend more time with each patient to provide a higher quality of care. However, the service price may increase if the competition is significantly softened. As the quality of care depends more heavily on the time spent with each patient, the service quality and the service price generally increase, while the market coverage may first decrease and then increase.

We further extend our study to allow the platform to endogenize the commission rate, and compare it with the setting where the platform can even set the service price. We find that in both settings, the platform will set the commission rate above some thresholds to limit competition. The two settings will have identical outcomes if the market is partially covered; otherwise, to allow the platform to set the price can lead to more doctor participation, higher service quality and price, higher
Figure 9: Illustration of the effect of $\Delta \delta$ when $\alpha = 1.2$ (upper six plots) and $\alpha = 6$ (lower six plots), with dotted curves for only commission control and solid curves for both commission and price control. The other parameters: $V_0 = 20, \mu_0 = 5, c = 20, \Lambda = 25, \delta_0 = 25, N = 30$.  

26
platform profit, and even higher profit for the doctors. The latter scenario generally arises when the quality of care is either very sensitive or very insensitive to the time spent with each patient, the unit waiting cost is low, or the variability of the doctors’ reservation profits is low. Beyond these factors, we also investigate patient heterogeneity (the results are available upon request). We find that its effect depends critically on the market coverage. A higher heterogeneity generally increases the doctor participation, the service rate, the service price, and the profits of the platform and the doctors when the market coverage is low, and it may lead to opposite outcomes otherwise.

References


Cowell, Kari. 2016. Decreasing the amount of unnecessary medical visits would benefit everyone in the healthcare system. *Medelita URL*


Guo, Pengfei, Christopher S. Tang, Yulan Wang, Ming Zhao. 2016. The impact of reimbursement policy on patient welfare, readmission rate and waiting time in a public healthcare system: Fee-for-service vs. bundled payment. UCLA Anderson School of Management Working Paper.


Online Appendix

Proof of Lemma 1. A monopoly doctor’s problem can be formulated as

\[
\max_{p, \mu, \lambda} p \lambda \\
\text{s.t. } V_0 + \alpha(\mu_0 - \mu) - p - \frac{c}{\mu - \lambda} \geq 0, \\
\mu > \lambda > 0.
\]

It is equivalent to

\[
\max_{\lambda, \mu} \left( V_0 + \alpha(\mu_0 - \mu) - \frac{c}{\mu - \lambda} \right) \lambda.
\]

Note that the revenue is maximized when the service rate \( \mu = \lambda + \sqrt{c/\alpha} \). Thus, the optimal monopoly price, service rate and demand rate satisfy

\[
p_M = \frac{V_0 + \alpha \mu_0 - 2\sqrt{\alpha c}}{2} = \alpha \lambda_M, \\
\mu_M = \frac{V_0 + \alpha \mu_0}{2\alpha}, \\
\lambda_M = \frac{V_0 + \alpha \mu_0 - 2\sqrt{\alpha c}}{2\alpha}.
\]

The profit rate is \( \pi_M = (1 - r)p_M \lambda_M = (1 - r)\alpha \lambda_M^2 \).

Proof of Proposition 1. Finding the best response of doctor \( i \) is equivalent to solving the following problem:

\[
\max_{p_i, \mu_i, \lambda_i} p_i \lambda_i \\
\text{s.t. } -\beta + \alpha \mu_i + p_i + \frac{c}{\mu_i - \lambda_i} \leq 0, \\
-\beta + \alpha \mu_i + p_i + \frac{c}{\mu_i - \lambda_i} = -\beta + \alpha \mu_j + p_j + \frac{c}{\mu_j - \lambda_j}, \forall j \neq i, \\
\sum_{i=1}^{n} \lambda_i \leq \Lambda, \\
0 \leq \lambda_i < \mu_i, i = 1, \ldots, n,
\]

where \( \beta = V_0 + \alpha \mu_0 \). With \( z_i^j, x^i \) being the Lagrange multipliers, the KKT conditions for doctor
i’s problem are given by

\[ \lambda_i = \sum_{j=1}^{n} z_{ij}^i, \quad (1) \]

\[ 0 = \left[ \alpha - \frac{c}{(\mu_i - \lambda_i)^2} \right] \sum_{j=1}^{n} z_{ij}^i, \quad (2) \]

\[ p_i = \frac{c}{(\mu_i - \lambda_i)^2} \sum_{j=1}^{n} z_{ij}^i + x^i, \quad (3) \]

\[ 0 = -\frac{cz_{ij}^i}{(\mu_j - \lambda_j)^2} + x^i, \quad j \neq i. \quad (4) \]

Note that, since \( \lambda_i = \sum_{j=1}^{n} z_{ij}^i > 0 \), \( \alpha - \frac{c}{(\mu_i - \lambda_i)^2} = 0 \). Thus,

\[ \mu_i^* = \lambda_i^* + \sqrt{\frac{c}{\alpha}}. \]

Also, at equilibrium, the patient surplus of joining any queue will be the same. Then, either

\[ z_{ij}^i > 0, \quad \forall i \text{ or } z_{ij}^i = 0, \quad \forall i. \]

Case 1: \( z_{ij}^i = 0 \) and \( z_{ij}^i > 0, \quad \forall j \neq i. \)

By (4), \( x_i = 0 \), which implies that \( \sum_{i=1}^{n} \lambda_i < \Lambda \) and \( p_i = \alpha \lambda_i \). In addition, constraints 

\[ -\beta + \alpha \mu_i + p_i + \frac{c}{\mu_i - \lambda_i^*} \leq 0, \quad \forall i \text{ are binding. Thus, } p_i^* = p_M, \mu_i^* = \mu_M, \lambda_i^* = \lambda_M, \text{ given that } n\lambda_i^* < \Lambda, \]

which is equivalent to \( n < n_0 \triangleq \frac{\Lambda}{\lambda_M} \).

Case 2: \( z_{ij}^i = z_{ij}^i > 0, \quad \forall i, j. \)

Constraints 

\[ -\beta + \alpha \mu_i + p_i + \frac{c}{\mu_i - \lambda_i^*} \leq 0, \quad \forall i \text{ are binding. Thus, } \]

\[ p_i^* = \beta - \alpha \lambda_i^* - 2\sqrt{\alpha c}. \]

By (3),

\[ x^i = \beta - 2\alpha \lambda_i^* - 2\sqrt{\alpha c}. \]

Thus,

\[ z_{ij}^i = x^i / \alpha \]

\[ = \beta / \alpha - 2\lambda_i^* - 2\sqrt{\alpha c}, \quad \forall j \neq i; \]

\[ z_{ij}^i = \lambda_i^* - \sum_{j \neq i} z_{ij}^i \]

\[ = (2n - 1)\lambda_i^* + 2(n - 1)\sqrt{c/\alpha} - (n - 1)\beta / \alpha. \]
Since $z_i^j > 0$, $\forall j$, we have
\[
\frac{2n-2}{2n-1} \lambda_M \leq \lambda_i^* \leq \lambda_M.
\]
In summary, the equilibrium satisfies
\[
\mu_i^* = \lambda_i^* + \sqrt{c/\alpha},
\]
\[
p_i^* = \beta - \alpha \lambda_i^* - 2\sqrt{\alpha c},
\]
\[
\sum_{i=1}^{N} \lambda_i^* = \Lambda,
\]
\[
\frac{2n-2}{2n-1} \lambda_M \leq \lambda_i^* \leq \lambda_M.
\]
When the market is equally divided among $n$ doctors, the equilibrium satisfies
\[
\mu_i^* = \Lambda/n + \sqrt{c/\alpha},
\]
\[
p_i^* = \beta - \alpha \Lambda/n - 2\sqrt{\alpha c}.
\]

Case 3: $z_i^i = 0$, $\forall i$.

By (1) and (4),
\[
\begin{align*}
z_i^j &= \frac{\lambda_i^*}{n-1}, \quad \forall j \neq i, \\
x_i^i &= \frac{\alpha \lambda_i^*}{n-1}.
\end{align*}
\]
Then, by (3),
\[
p_i^* = \alpha \lambda_i^* \frac{n}{n-1}.
\]
Since
\[
\beta - \alpha \mu_i - p_i - \frac{c}{\mu_i - \lambda_i} = \beta - \alpha \mu_j - p_j - \frac{c}{\mu_j - \lambda_j}, \quad \forall j \neq i,
\]
we have
\[
\beta - \alpha (\lambda_i^* + \sqrt{c/\alpha}) - \alpha \lambda_i^* \frac{n}{n-1} - \sqrt{\alpha c} = \beta - \alpha (\lambda_j^* + \sqrt{c/\alpha}) - \alpha \lambda_j^* \frac{n}{n-1} - \sqrt{\alpha c} \geq 0.
\]
Thus,
\[
\lambda_i^* = \lambda_j^*, \quad \forall j,
\]
\[
\lambda_i^* \leq \frac{2n-2}{2n-1} \lambda_M.
\]
In summary,

\[
\lambda_i^* = \frac{\Lambda}{n}, \\
\mu_i^* = \frac{\Lambda}{n} + \sqrt{c/\alpha}, \\
p_i^* = \frac{\alpha \Lambda}{n - 1}, \\
\lambda_i^* \leq \frac{2n - 2}{2n - 1} \lambda_M.
\]

The threshold \(n_1\) can be obtained by solving

\[
\frac{\Lambda}{n_1} = \frac{2n - 2}{2n - 1} \lambda_M \quad \Rightarrow \quad n_1 = \frac{1 + \frac{\Lambda}{\lambda_M} + \sqrt{1 + \frac{\Lambda^2}{\lambda_M^2}}}{2}.
\]

**Proof of Lemma 2.** For each participating doctor \(i\), the optimal service rate will be set to \(\lambda_i + \sqrt{c/\alpha}\). Thus, in the following analysis, we only focus on the pricing decisions of the doctors.

Given \(p_i, i \leq n - 1\), the demand function of the \(n\)th doctor is

\[
\lambda_n(p_n) = \left\{ \begin{array}{ll}
\frac{\Lambda}{n} + \frac{1}{\alpha} \sum p_{n-1} - \frac{n-1}{\alpha n} p_n, & \text{if } p_n + \sum p_{n-1} \leq n(\beta - 2\sqrt{\alpha c}) - \alpha \Lambda, \\
\beta - 2\sqrt{\alpha c} - p_n, & \text{if } p_n + \sum p_{n-1} > n(\beta - 2\sqrt{\alpha c}) - \alpha \Lambda,
\end{array} \right.
\]

where \(\sum p_{n-1} = \sum_{i \leq n-1} p_i\). Thus, the optimal price set by the \(n\)th doctor is

\[
p_n^0(\{p_i, i \leq n-1\}) = \left\{ \begin{array}{ll}
\frac{\alpha \Lambda + \sum p_{n-1}}{2(n-1)}, & \text{if } \sum p_{n-1} \leq \frac{2n-2}{2n-1} n(\beta - 2\sqrt{\alpha c}) - \alpha \Lambda, \\
n(\beta - 2\sqrt{\alpha c}) - \alpha \Lambda - \sum p_{n-1}, & \text{if } \frac{2n-2}{2n-1} n(\beta - 2\sqrt{\alpha c}) - \alpha \Lambda < \sum p_{n-1} \\
\beta - 2\sqrt{\alpha c} - \alpha \Lambda, & \text{if } \sum p_{n-1} > \frac{2n-1}{2} (\beta - 2\sqrt{\alpha c}) - \alpha \Lambda,
\end{array} \right.
\]

The optimal profit of the \(n\)th doctor is

\[
\pi_n^0(\{p_i, i \leq n-1\}) = \left\{ \begin{array}{ll}
\frac{(\alpha \Lambda + \sum p_{n-1})^2}{2n(n-1)}, & \text{if } \sum p_{n-1} \leq \frac{2n-2}{4n(n-1)} n(\beta - 2\sqrt{\alpha c}) - \alpha \Lambda, \\
\frac{1}{n} [n(\beta - 2\sqrt{\alpha c}) - \alpha \Lambda - \sum p_{n-1}] [\alpha \Lambda + \sum p_{n-1} - (n-1)(\beta - 2\sqrt{\alpha c})], & \text{if } \frac{2n-2}{2n-1} n(\beta - 2\sqrt{\alpha c}) - \alpha \Lambda < \sum p_{n-1} \leq \frac{2n-1}{2} (\beta - 2\sqrt{\alpha c}) - \alpha \Lambda, \\
\alpha \Lambda n M, & \text{if } \sum p_{n-1} > \frac{2n-1}{2} (\beta - 2\sqrt{\alpha c}) - \alpha \Lambda.
\end{array} \right.
\]

Thus, if \(\delta_n \leq (1 - r) \frac{\alpha \Lambda^2}{4n(n-1)}\), the \(n\)th doctor will always join the platform; if \(\delta_n > (1 - r) p_M \lambda_M\), the \(n\)th doctor will never join the platform. For \(\delta_n\) in between, by solving

\[
(1 - r)\pi_n^0(\{p_i, i \leq n-1\}) = \delta_n
\]

33
for \( p_i = p_j, \forall i, j \leq n - 1 \), we have the results.

**Proof of Lemma 3.** We first prove that \( \hat{p}(n, r) \) increases in \( r \) and \( n \). Obviously, \( \hat{p}(n, r) \) increases in \( r \). Note that, the function

\[
\tilde{\pi}(p, n) = \begin{cases} 
\frac{(\alpha \Lambda + (n-1)p)^2}{4\alpha \Lambda (n-1)}, & \text{if } p \leq \frac{2n - 2\alpha \Lambda}{2n-1}(eta - 2\sqrt{\alpha c}) - \frac{\alpha \Lambda}{n-1}, \\
\frac{1}{\alpha}n(\beta - 2\sqrt{\alpha c}) - \alpha \Lambda - (n - 1)p\left[\alpha \Lambda + (n-1)p - (n-1)(\beta - 2\sqrt{\alpha c})\right], & \text{if } \frac{2n - 2\alpha \Lambda}{2n-1}(eta - 2\sqrt{\alpha c}) - \frac{\alpha \Lambda}{n-1} < p \leq \frac{2n - 2\alpha \Lambda}{2n-2}(eta - 2\sqrt{\alpha c}) - \frac{\alpha \Lambda}{n-1}, \\
\frac{\beta n}{\alpha}, & \text{if } p > \frac{\beta n}{\alpha}.
\end{cases}
\]

decreases in \( n \) for any \( p \geq 0 \). Also, \( \delta_n \) increases in \( n \). Thus, we have \( \hat{p}(n, r) \) increases in \( n \).

Each doctor will set price to fully extract all patient surplus if

\[
\frac{\beta - 2\sqrt{\alpha c} - \hat{p}(n, r)}{\alpha} \leq \frac{\Lambda}{n-1} \iff \hat{p}(n, r) \geq \beta - 2\sqrt{\alpha c} - \frac{\alpha \Lambda}{n-1} \iff \delta_n \geq (1-r)\frac{\alpha \Lambda^2(n-1)}{n}.
\]

Thus, if \( \delta_n \geq (1-r)\frac{\alpha \Lambda^2(n-1)}{n} \), \( \hat{\lambda}(n, r) = \frac{\Lambda}{n-1} \); otherwise, \( n - 1 \) doctors evenly split the market and \( \hat{\lambda}(n, r) = \frac{\Lambda}{n-1} \).

**Proof of Lemma 4.** Let

\[
n_{21}(r) = \arg \max_n \{\hat{p}(n, r)\frac{\Lambda}{n-1}\} \quad \text{and} \quad n_{22}(r) = \arg \max_n \{\hat{p}(n, r)\frac{\beta - 2\sqrt{\alpha c} - \hat{p}(n, r)}{\alpha}\}.
\]

\[
\partial \left(\hat{p}(n, r)\frac{\Lambda}{n-1}\right)/\partial n = \Lambda \frac{\partial \hat{p}(n, r)}{\partial n} \frac{(n-1) - \hat{p}(n, r)}{(n-1)^2} = 0 \iff \frac{\partial \hat{p}(n, r)}{\partial n} = \hat{p}(n, r)\frac{1}{n-1}.
\]

Thus, for \( \delta_n < (1-r)\frac{\alpha \Lambda^2(n-1)}{n} \),

\[
\hat{p}(n_{21}(r), r)\frac{\Lambda}{n_{21}(r) - 1} = \Lambda \frac{\partial \hat{p}(n, r)}{\partial n} \bigg|_{n=n_{21}(r)}.
\]

Note that when \( \delta_n < (1-r)\frac{\alpha \Lambda^2(n-1)}{n(n-1)} \), \( \hat{p}(n, r) = 2\sqrt{\frac{\alpha \Lambda^2 n}{(1-r)(n-1)}} - \alpha \frac{\Lambda}{n-1} \). Then,

\[
\frac{\Lambda}{n(n-1)} \frac{\partial \hat{p}(n, r)}{\partial n} = \frac{\alpha \Lambda^2}{n(n-1)} \propto \sqrt{\frac{(1-r)\alpha \delta_n}{n(n-1)}} - \delta_n + (n-1)n\delta'_n,
\]

where we treat \( \delta_n \) as a differentiable function in \( n \) and \( \delta'_n \) is its derivative. Since if \( \delta_n > (1-r)\frac{\alpha \Lambda^2}{n(n-1)} \), the \( n_{th} \) doctor will not join the platform. Thus,

\[
\Lambda \sqrt{\frac{(1-r)\alpha \delta_n}{n(n-1)}} - \delta_n + (n-1)n\delta'_n \geq 0.
\]

Therefore,

\[
\hat{p}(n_{21}(r), r)\frac{\Lambda}{n_{21}(r) - 1} > p^*(n)\lambda^*(n).
\]

34
For \( \hat{p}(n, r) \frac{\beta - 2\sqrt{\alpha c}}{\alpha} \), it will be maximized at \( \hat{p}(n, r) = p_M \). Since if \( n < n_0 \),

\[
p_M > \frac{2n - 1}{2n - 2} (\beta - 2\sqrt{\alpha c}) - \frac{\alpha \Lambda}{n - 1},
\]

and thus, \( \hat{p}(n, r) < p_M \). Therefore, \( \hat{p}(n, r) = p_M \) will only be satisfied for some \( n \geq n_0 \). Also, for \( n \geq n_0 \), \( p^*(n)\lambda^*(n) \leq p_M \lambda_M \). Hence,

\[
\hat{p}(n, r) \frac{\beta - 2\sqrt{\alpha c} - \hat{p}(n, r)}{\alpha} \geq p^*(n)\lambda^*(n).
\]

In summary, both \( \hat{p}(n, r) \frac{\Lambda}{n - 1} \) and \( \hat{p}(n, r) \frac{\beta - 2\sqrt{\alpha c} - \hat{p}(n, r)}{\alpha} \) increase in \( n \) for any \( n \) such that \( \hat{p}(n, r)\lambda(n, r) \leq p^*_i(n)\lambda^*_i(n) \). Therefore, \( \hat{p}(n, r)\lambda(n, r) \) increases in \( n \) in such region. \( \square \)

**Proof of Proposition 2.** If \( \pi_M = (1 - r)\alpha \lambda^2_M \leq \delta_{n_0} \), only the first \( n_0 \) doctors can possibly join the platform. Whoever joins the platform charges the monopoly price and serves the monopoly demand rate. Thus, the exact number of participating doctors is the number of doctors who enjoy nonnegative profits. Thus, if

\[
(1 - r)\alpha \lambda^2_M \leq \delta_{n_0} \iff r \geq 1 - \frac{\delta_{n_0}}{\alpha \lambda^2_M},
\]

\( n^* = \max\{n | (1 - r)\alpha \lambda^2_M \geq \delta_n \} \).

Let

\[
\hat{r}(N) = \{r | p^*(N) \frac{\Lambda}{N} = \hat{p}(N, r)\hat{\lambda}(N, r) \}.
\]

Then, if \( r < \hat{r}(N) \), all doctors will join the platform.

If \( r \in [\hat{r}, 1 - \frac{\delta_{n_0}}{\alpha \lambda^2_M}] \), the equilibrium number of doctors is given by \( \hat{n}(r) \).

\( \square \)

**Proof of Proposition 3.** Note that as \( r \) increases, \( \hat{n}(r) \) and \( \max\{n | (1 - r)\alpha \lambda^2_M \geq \delta_n \} \) decrease. Thus, \( n^*(r) \) decreases in \( r \).

When \( r \leq 1 - \frac{\delta_{n_0}}{\alpha \lambda^2_M} \), the service rate is \( \frac{\Lambda}{n^*(r)} + \sqrt{c/\alpha} \); otherwise, the service rate is \( \mu_M \). Thus, the service rate increases in \( r \).

Let \( \tilde{r}_1 = \{r | n^*(r) = n_1 \} \). By the monotonicity of the equilibrium price, as \( n^*(r) \) decreases to \( n_1 \), the equilibrium price increases; when \( n^*(r) \) decreases between \( n_1 \) and \( n_0 \), the equilibrium price decreases; when \( n^*(r) < n_0 \), the equilibrium price is \( p_M \).

\( \square \)

**Proof of Proposition 4.** First, consider the case \( n < n_0 \). The platform’s problem is

\[
\max_{0 \leq n < n_0} n r \alpha \lambda^2_M
\]

s.t. \( (1 - r)\alpha \lambda^2_M = \frac{n - 1}{N - 1} \Delta \delta + \delta_0 - \Delta \delta. \)
Thus,
\[
r = 1 - \frac{n-1}{N-1}2\Delta\delta + \delta_0 - \Delta\delta\]
and the platform’s problem is equivalent to
\[
\max_{0 \leq n < n_0} n \left( 1 - \frac{n-1}{N-1}2\Delta\delta + \delta_0 - \Delta\delta \right) \alpha\lambda^2_M.
\]
Therefore, the optimal solution is
\[
n^* = \frac{1}{4} \left[ N + 1 + \frac{1}{\Delta\delta} (N - 1) (\alpha\lambda^2_M - \delta_0) \right],
\]
\[
r^* = \frac{1}{2} - \frac{\delta_0}{2\alpha\lambda^2_M} + \frac{N + 1}{N - 1} \frac{\Delta\delta}{2\alpha\lambda^2_M},
\]
if \(n^* < n_0\), which is equivalent to
\[
N < \bar{N} \triangleq \frac{\alpha\lambda^2_M - \delta_0 - \Delta\delta}{\alpha\lambda^2_M - \delta_0 - 3\Delta\delta}.
\]

Note that when the market is fully covered, \(n^*\) cannot exceed \(n_1\). The reason is that, as shown by Proposition 1, the equilibrium price \(p^*(n)\) decreases in \(n\) for \(n > n_1\). Thus, when the market is fully covered, \(n^* \in [n_0, n_1]\). By Proposition 1, the optimal price and service rate are given by \(p^*(n^*)\) and \(\mu^*(n^*)\), respectively.

**Proof of Proposition 5.** If \(N < \bar{N}\), the problem is exactly the same as in the case where doctors set prices. Thus, \(n_b^* = n^*, r_b^* = r^*, p_b^* = p^*\) and \(\mu_b^* = \mu^*\).

If \(N \geq \bar{N}\), the platform will set price to fully extract the patient surplus. Thus, \(p_b^* = \beta - \alpha\Lambda/n_b^* - 2\sqrt{\alpha c}\). Then, the platform’s problem is
\[
\max_r r\Lambda(\beta - \alpha\frac{\Lambda}{n} - 2\sqrt{\alpha c})
\]
\[
s.t. (1 - r)\frac{\Lambda}{n}(\beta - \alpha\frac{\Lambda}{n} - 2\sqrt{\alpha c}) = \frac{n - 1}{N - 1}2\Delta\delta + \delta_0 - \Delta\delta,
\]
\[
0 \leq r \leq 1.
\]
It is equivalent to solve
\[
\max_{n \geq n_0} (\beta - \alpha\frac{\Lambda}{n} - 2\sqrt{\alpha c})\Lambda - (\frac{n - 1}{N - 1}2\Delta\delta + \delta_0 - \Delta\delta)n,
\]
whose F.O.C. is given by
\[
\frac{\alpha\Lambda^2}{n^2} - \frac{4n\Delta\delta}{N - 1} + \frac{\Delta\delta(N + 1)}{N - 1} - \delta_0 = 0.
\]
Thus, \(n_b^*\) satisfies the above condition. The optimal commission rate \(r_b^*\) satisfies
\[
(1 - r_b^*)\frac{\Lambda}{n_b^*}(\beta - \alpha\frac{\Lambda}{n_b^*} - 2\sqrt{\alpha c}) = \delta_n^*.
\]

The optimal price and service rate are given by
\[ p_b^* = V_0 + \alpha \mu_0 - \alpha \Lambda / n_b^* - 2\sqrt{ac}, \]
\[ \mu_b^* = \Lambda / n_b^* + \sqrt{c/\alpha}. \]

**Proof of Proposition 6.** If \( N < \bar{N}, \) clearly \( \mu_b^* = \mu^* = \mu_M, \) \( p_b^* = p^* = p_M, \) \( n_b^* = n^* < n_0, \)
\( n_b^* \lambda_b^* = n^* \lambda^* < \Lambda. \)

In the following, we show that for \( n > n_0 \) such that \( p^*(n) \lambda^*(n) = \delta_n / (1 - r), \) \( p^*(n) \lambda^*(n) < \hat{p}(n, r) \hat{\lambda}(n, r). \)

If \( \frac{\delta_n}{1 - r} \geq \frac{2\alpha^2 \Lambda}{\Lambda + 2\alpha^2 \Lambda^2 + \Lambda M}, \) then \( \frac{\delta_n}{1 - r} \geq \frac{4n(n-1)}{(2n-1)} \alpha \Lambda^2 M \) for \( n \in [n_0, n_1]. \) Thus, \( \hat{p}(n, r) = \frac{2n-1}{2n-2}(\beta - 2\sqrt{ac}) - \alpha \frac{\Lambda}{n-1} - \frac{1}{n-1} \sqrt{\frac{(\beta - 2\sqrt{ac})^2}{4} - \frac{\alpha \Lambda}{n-1}} \) and \( \hat{\lambda}(n, r) = \frac{\beta - 2\sqrt{ac} - \hat{p}(n, r)}{\alpha}. \) In addition, \( p^*(n) = \beta - \alpha \frac{\Lambda}{n} - 2\sqrt{ac} \)
and \( \lambda^*(n) = \frac{\Lambda}{n}. \)

Note that when \( \frac{\delta_n}{1 - r} = (\beta - \alpha \frac{\Lambda}{n} - 2\sqrt{ac}) \frac{\Lambda}{n}, \)
\[ \hat{p}(n, r) - p_M = p_M - \left( \beta - \alpha \frac{\Lambda}{n} - 2\sqrt{ac} \right) \geq 0. \]
Thus, \( \hat{p}(n, r) \hat{\lambda}(n, r) = p^*(n) \lambda^*(n) \) at \( \frac{\delta_n}{1 - r} = (\beta - \alpha \frac{\Lambda}{n} - 2\sqrt{ac}) \frac{\Lambda}{n}. \) Denote this intersection of \( \hat{p}(n, r) \hat{\lambda}(n, r) \) and \( p^*(n) \lambda^*(n) \) by \( \bar{n}. \) Since \( \hat{p}(n, r) \hat{\lambda}(n, r) \) decreases in \( n \) when \( \hat{p}(n, r) \geq p_M, \) there is another intersection of \( \hat{p}(n, r) \hat{\lambda}(n, r) \) and \( p^*(n) \lambda^*(n) \) between \( n_0 \) and \( \bar{n}, \) which will be the equilibrium number of participating doctors. Thus, we have \( (1 - r^*) p^*(n^*) \lambda^*(n^*) > \delta_{n^*}. \)

If \( \frac{\delta_n}{1 - r} < \frac{2\alpha^2 \Lambda}{\Lambda + 2\alpha^2 \Lambda^2 + \Lambda M}, \) then \( \frac{\delta_n}{1 - r} = \frac{\alpha \Lambda^2}{n(n-1)} \) for some \( n > n_1. \) Then, \( \hat{p}(n, r) = 2 \sqrt{\frac{\alpha n \delta_n}{(1 - r)(n-1)}} - \alpha \frac{\Lambda}{n-1}, \)
\[ \hat{\lambda}(n, r) = \frac{\beta - 2\sqrt{ac} - \hat{p}(n, r)}{\alpha} \text{ or } \frac{\Lambda}{n-1}. \] Plugging \( \frac{\delta_n}{1 - r} = \frac{\alpha \Lambda^2}{n(n-1)} \) into \( \hat{p}(n, r), \) we have
\[ \hat{p}(n, r) \frac{\Lambda}{n-1} - \frac{\alpha \Lambda^2}{n(n-1)} = \frac{\alpha \Lambda^2}{(n-1)^2} > 0, \]
\[ \hat{p}(n, r) \frac{\beta - 2\sqrt{ac} - \hat{p}(n, r)}{\alpha} - \frac{\alpha \Lambda^2}{n(n-1)} = \frac{\alpha \Lambda(n - 2\Lambda n + 2\Lambda M(n-1)n)}{(n-1)^2 n} > 0 \text{ if } n > n_1. \]
Therefore, we have \( (1 - r^*) p^*(n^*) \lambda^*(n^*) > \delta_{n^*}. \) It implies that when the platform has full control over the commission and the price, the number of participating doctors is always higher. It implies a higher price and a lower service rate. \( \square \)

**Proof of Proposition 7.** When the market is partially covered, the market coverage is
\[ n^* \lambda_M = \frac{1}{4} \left[ N + 1 + \frac{1}{\Delta \delta} (N - 1)(\alpha \lambda_M^2 - \delta_0) \right] \lambda_M. \]
Take derivatives w.r.t. \( \alpha, \)
\[ \frac{\partial n^* \lambda_M}{\partial \alpha} \propto (\sqrt{ac} - V_0) \left( \alpha \lambda_M^2 - (\delta_0 - \frac{N + 1}{N - 1} \Delta \delta) \right) + (\alpha \mu_0 - V_0) \alpha \lambda_M^2. \]

37
The above equation (treated as a function of $\alpha$) has no more than one zero point when $\alpha > 0$. If there is one, then $n^*\lambda_M$ is unimodal, which decreases in $\alpha$ and then increases in $\alpha$. Therefore, $n^*\lambda_M = \Lambda$ has at most two solutions. Let $\underline{\alpha} \leq \bar{\alpha}$ be the solutions to $n^*\lambda_M = \Lambda$. Then, the market is fully covered if $\alpha \in [0, \underline{\alpha}] \cup [\bar{\alpha}, \infty)$, and partially covered otherwise.

Proof of Proposition 8. Note that

$$n^*\lambda_M = \frac{1}{4} \left[ N + 1 + \frac{1}{\Delta \delta} (N - 1)(\alpha \lambda_M^2 - \delta_0) \right] \lambda_M$$

decreases in $c$ since $\lambda_M$ decreases in $c$. Let $\bar{\alpha}$ be the solution to $n^*\lambda_M = \Lambda$. Then, when $c < \bar{\alpha}$, the market is fully covered, and when $c \geq \bar{\alpha}$, the market is partially covered.

Proof of Proposition 9. Note that

$$n^*\lambda_M = \frac{1}{4} \left[ N + 1 + \frac{1}{\Delta \delta} (N - 1)(\alpha \lambda_M^2 - \delta_0) \right] \lambda_M$$

decreases in $\Delta \delta$. Let $\overline{\Delta \delta}$ be the solution to $n^*\lambda_M = \Lambda$. Then, when $\Delta \delta < \overline{\Delta \delta}$, the market is fully covered, and when $\Delta \delta \geq \overline{\Delta \delta}$, the market is partially covered.

38