Inter-regional Competition and Quality in Public Hospital Care

Hiroshi Aiura
Faculty of Economics, Oita University
700, Dannoharu, Oita, 870-1192, Japan.

E-mail address: office@aiura.info

Abstract

This paper analyzes the effect of capitation payment and patient choice on waiting time and the comprehensive quality of hospital care. The study assumes that two local public hospitals are located in two cities with different population sizes and compete with each other. We find that the comprehensive quality of hospital care as well as waiting time of both hospitals improve with an increase in capitation payment. However, we also find that the extent of these improvements differs according to the population size of the cities where the hospitals are located. Under the realistic assumptions that hospitals involve significant labor-intensive work, we find the improvements in comprehensive quality and waiting time in the hospital located in the small city to be greater than those in the hospital located in the large city. The result implies that regional disparity in the quality of hospital care decreases with an increase in capitation payment.

Keywords: Patient choice; Waiting time; Hotelling-type spatial competition model; Multi-region model

JEL classification: I18, L32

1 Introduction

This paper analyzes the effect of inter-regional competition on waiting time and comprehensive quality in public hospital care. During the past few decades, several European countries—Norway, Switzerland, the United Kingdom, etc.—have reformed their health care systems; they have introduced free choice in hospitals (so-called patient choice) instead of limited or no choice, and a capitation payment in the reimbursement system. These changes aim to improve the quality of health care, especially waiting time, because they provide hospitals with an incentive to compete for acquiring patients. Since hospital care is horizontally differentiated according to geographical location, hospitals experiencing these changes compete geographically in terms of the quality of hospital care.\(^1\)

Several theoretical studies using spatial competition models have attempted to determine the manner in which an incentive to compete for acquiring patients influences the quality of hospital care.

\(^1\)Tay (2003) empirically shows that the demand responses to both distance and quality are substantial.
Gravelle and Masiero (2000), Karlsson (2007), and Brekke et al. (2010) focus on how this incentive influences quality within the health-care system. Although waiting time is implicitly modeled as a part of quality (as mentioned by Brekke et al., 2007), the abovementioned studies do not provide an answer regarding the effect on waiting time alone, that is, separate from quality. On the other hand, Xavier (2003), Siciliani (2005), and Brekke et al. (2008) focus on the impact of this incentive on waiting time; however, they do not consider comprehensive quality.

Moreover, these studies do not assume that patients undertake trips across provinces with different population sizes. If geographical constraints did not exist, patients would prefer hospitals located in large cities to those in small cities, because hospitals located in large cities enjoy economies of scale and/or scope and appear to provide higher quality health care than those in small cities.\(^2\) Therefore, when patient choice is introduced, a certain number of patients flow into hospitals located in large cities from small cities; the impact of this, directly on waiting time and secondarily on comprehensive quality, is different for the hospitals in both types of cities. Aiura and Sanjo (2010) consider this flow of patients and derive the competitive equilibrium qualities of local public hospitals located in two regions with different population sizes. However, economies of scale and scope do not work in their model, and a counterintuitive result—that a rural public hospital always offers higher quality of hospital care than an urban public hospital—is derived. Furthermore, they do not analyze the effect of the incentive to compete.

In this paper, we assume that the comprehensive quality of hospital care is reflected by several factors, such as waiting time, medical technology, and the skills and experience of medical staff. The model in this paper divides these factors into two: waiting time and factors that are irrelevant to waiting time. By this division, we can study the effect on not only the comprehensive quality of hospital care but also on waiting time alone—that is, waiting time separated from comprehensive quality; this has not been dealt with in previous studies.

When the demand on hospitals increases, hospitals get crowded, and additional resources are necessary to cut down congestion and maintain a certain amount of waiting time. Thus, we assume that the costs to maintain a certain amount of waiting time would increase with the demand on hospitals. On the other hand, we assume that the quality of hospital care brought in by factors irrelevant to waiting time does not get worse even if the demand on hospitals increases. For example, the medical technology and skills and experience of medical staff of a hospital remain the same regardless of their number of patients (although improvement of these factors may increase its waiting time by attracting more patients). Accordingly, we assume that there are economies of scale in the costs for improving the factors that are irrelevant to waiting time, but there are not economies of scale in the costs for reducing waiting time.

When these assumptions about the comprehensive quality of hospitals and costs as well as patient choice across cities are introduced, we derive the competitive equilibrium qualities of the local public hospitals located in the two cities with different population sizes. Under certain specified parameters, there exists an equilibrium in which a hospital in a large city is superior to a hospital in a small city in terms of comprehensive quality and the brevity of waiting time. This equilibrium is intuitive, and is not shown by Aiura and Sanjo (2010).

Further, we analyze the effect of capitation payment and find that both the local public hospitals

---


\(^3\) Aletras (1999) suggests that, apparently, economies of scale work effectively in acute care hospitals with 100-200 beds. Preyra and Pink (2006) show that economies of scale and scope through hospital consolidations are almost certainly possible.
in the two cities improve not only their comprehensive qualities but also their waiting times with an increase in capitation payment; however, the extent of these improvements is different for the hospitals. In an actual situation, in which hospitals involve significant labor-intensive work, these improvements are seen to be greater for the hospital in the small city than for the hospital in the large city. This result implies that regional disparity in the quality of hospital care decreases with an increase in capitation payment. Since the costs required by a hospital that accepts only a few patients for a certain decrease in waiting time is smaller than that required by a hospital that accepts many patients, the reasoning behind this result is that the hospital in the small city, which has a small demand, has a cost advantage in terms of improvement in waiting time. Further, on the basis of these assumptions, we can also infer that the reduction in the disparity in waiting time between the two hospitals in the large and small cities is greater than that in comprehensive quality. In other words, when patients are given a free choice of hospitals and capitation payment to hospitals is adequate, regional disparity in waiting time appears to be smaller than that in comprehensive quality.

OECD Regions at a Glance 2009 shows that Japan has a more balanced regional distribution of physicians than most European countries. This report also shows that the number of physicians in the urban regions of each European country is positively correlated with population share, whereas the number of physicians in the urban regions of Japan is negatively correlated with population share. Japan has been permitting patient choice in hospitals since the 1960s, and the Japanese are used to exercising this choice. Moreover, Japan is small in size. Therefore, Japanese hospitals seem to be highly competitive. Thus, the results from the model in this study would support the findings in OECD Regions at a Glance 2009.

The remainder of this paper is organized as follows: section 2 presents the model. Section 3 determines the first-best outcome when all hospitals are national hospitals. Section 4 determines the equilibrium at which local public hospitals compete on quality and investigates how this changes with an increase in capitation payment. Section 5 uses numerical analysis to support the implications of section 4. Section 6 presents several discussions resulting from section 4’s implications, and Section 7 presents the conclusion.

2 Model

In this study, we consider an economy extended over a linear segment with length 1. Two cities, city 1 and city 2, are located at the two endpoints of this segment. Geographically speaking, the measure of each city is 0; that is, each city is regarded as a point on the segment.4 The area between the two cites is assumed to be agricultural, and the population in this agricultural area is distributed uniformly. Hereafter, the agricultural area is referred to as the “village.” We indicate the populations in city 1, city 2, and the village as \( N_1 \), \( N_2 \), and 1, respectively. Further, we assume that \( N_1 > N_2 > 1 \), which implies that the populations of both the cities are larger than the population of the village, and the population of city 1 is larger than that of city 2. Each of the two cities and the village are governed by different local governments. Only the cities have local public hospitals; the village does not have one because the village government would not have sufficient funds to manage it. Therefore, the people in the village have to go to either of the two cities in order to receive hospital care.

Every resident in each of the cities and the village demands one unit of hospital care. When resident takes one unit of hospital care available in city \( i \), he/she gains benefits equal to \( v + q(w_i, \theta_i) \),

---

4Takahashi (2004) also considers a similar spatial economy. Even if the people in cities are spread over a segment with a certain length, the results do not change within the parameter domain of this paper.
where $v$ denotes the reservation benefits obtained by accessing hospital care, and $q(w_i, \theta_i)$ denotes the comprehensive quality of hospital care provided by the local public hospital in city $i (= 1, 2)$. In order to guarantee full market coverage, we assume that $v$ is sufficiently large. The comprehensive quality of hospital $i$ is a function of two factors: (1) waiting time, denoted by $w_i$, and (2) the amount of such resources that reflect the quality of hospital care but do not influence waiting time, denoted by $\theta_i$. $w_i$ and $\theta_i$ are substitutable, but not perfectly. For an extreme example, we would definitely not want a hospital in which waiting time is more than 10 years, even if it had the best medical technology in the world. Therefore, we assume that

$$q(w_i, \theta_i) = aw_i^{-\alpha}\theta^\beta,$$

where $a$, $\alpha$, and $\beta$ are constant and greater than 0. The elasticity of the comprehensive quality with regard to waiting time is equal to $\alpha$; that is, waiting time becomes more worth for patients as $\alpha$ increases. The residents of the cities and the village who receive hospital care from another city incur transportation costs for traveling from their homes to the city that provides hospital care. When a resident residing at $x \in [0, 1]$ receives hospital care in city 1, we assume that he/she incurs $tx$ as transportation costs, where $t$ is a constant and greater than 0. Similarly, when the resident receives hospital care in city 2, we assume that he/she incurs $t(1 - x)$ as transportation costs.$^5$

Therefore, a resident residing at $x$, who consumes one unit of hospital care that is available in city $i (= 1, 2)$ gains a surplus—denoted by $u_i(x)$—which is equal to

$$u_1(x) = v + q(w_1, \theta_1) - rp - tx,$$

$$u_2(x) = v + q(w_2, \theta_2) - rp - t(1 - x),$$

where $p$ and $r$ denote the capitation payment to the hospital and co-payment rate of the patient, respectively, in the reimbursement system that the central government constructs. The central government pays $(1 - r)p$ per capitation to hospitals. Since most patients consult their general practitioners (GPs) before accessing hospital care, they are well informed by their GPs about the hospitals they will access; thus, we assume that the residents have perfect information regarding the qualities of hospitals in both the cities before receiving hospital care.$^6$ Accordingly, the residents go to the city that offers a higher surplus and receive hospital care that is available in that city. Let $X$ denote the location of the resident who receives the same amount of surplus from both the cities. Thus, we obtain the following equation:

$$v + q(w_1, \theta_1) - rp - tX = v + q(w_2, \theta_2) - rp - t(1 - X),$$

which yields

$$X = \frac{1}{2} + \frac{q(w_1, \theta_1) - q(w_2, \theta_2)}{2t}.$$

The residents on the left side of $X$ consume one unit of hospital care available in city 1 and gain $u_1$, whereas those on the right side of $X$ consume one unit of hospital care available in city 2 and gain $u_2$.

---

$^5$Even if we assume quadratic transportation costs, the results remain unchanged.

$^6$If the residents do not have perfect information, but the errors that patients make with regard to the quality of hospitals are identically and independently distributed, then the results presented by this paper hold qualitatively; however, the effect of capitation payment weakens.
Accordingly, the demand for hospital care in city \(i\)\((= 1, 2)\), \(D_i\), is given by

\[
D_1(q(w_1, \theta_1), q(w_2, \theta_2)) = \begin{cases} 
0, & \text{if } X < 0 \\
N_1 + X, & \text{if } 0 \leq X \leq 1 \\
N_1 + N_2 + 1, & \text{if } X > 1
\end{cases}
\]

\[
D_2(q(w_1, \theta_1), q(w_2, \theta_2)) = (N_1 + N_2 + 1) - D_1(q(w_1, \theta_1), q(w_2, \theta_2)).
\]

The local public hospital in city \(i\) requires resources to provide \(D_i\) units of hospital care at its decided waiting time (which is \(w_i\)). If the hospital handled \(k\) times the demand without increasing resources, the waiting time would become \(k\) times longer. Conversely, if the hospital handled \(k\) times the demand without increasing waiting time, \(k\) times resources would be needed. Therefore, we assume the costs of the resources to provide \(D_i\) units of hospital care at its decided waiting time (which is \(w_i\)) as \(r_iD_i/w_i\), where \(r_i\) denotes the price of the resources in city \(i\) to provide one unit of hospital care per unit of time. In addition, the local public hospital in city \(i\) also requires \(\theta_i\) (which denotes the amount of hospital-care resources that do not influence waiting time); thus, the cost function is

\[
c(\theta_i, w_i, D_i) = r_i \frac{D_i}{w_i} + \hat{r}_i \theta_i,
\]

where \(\hat{r}_i\) denotes the unit price of \(\theta_i\). We assume that all resources to provide hospital care are mobile across regions and sectors, and their prices are determined in a competitive marketplace. Therefore, their prices are the same in cities 1 and 2, and we set \(\hat{r}_1 = \hat{r}_2 = 1\) as the numeraire; we also set \(r_1 = r_2 = r\).

Since the demand of both hospitals remains unchanged unless their comprehensive qualities change, the local public hospitals face the following cost minimization problem in providing hospital care with comprehensive quality level \(q_i\):

\[
\min_{w_i, \theta_i} \frac{r D_1(q_1, q_2)}{w_i} + \theta_i \quad \text{s.t. } q(w_i, \theta_i) = q_i \ (i = 1, 2).
\]

The first-order conditions for the minimization problem yield

\[
w_i(q_1, q_2) = \frac{1}{\alpha^\gamma} \left( \frac{\beta}{\alpha} r \right)^{\beta\gamma} D_i(q_1, q_2)^{\beta\gamma} q_i^{-\gamma},
\]

\[
\theta_i(q_1, q_2) = \frac{1}{\alpha^\gamma} \left( \frac{\beta}{\alpha} r \right)^{\alpha\gamma} D_i(q_1, q_2)^{\alpha\gamma} q_i^{\gamma},
\]

where \(\gamma = 1/(\alpha + \beta)\); since each \(w_i\) and \(\theta_i\) are determined by \(q_1\) and \(q_2\), hereafter we focus on \(q_1\) and \(q_2\) as strategic variables of the hospitals. Therefore, we obtain the cost function with respect to \(q_i\) as

\[
c_i(q_1, q_2) = AD_i(q_1, q_2)^{\alpha\gamma} q_i^{\gamma}, \quad \text{where } A = \frac{1}{\alpha^\gamma} \left[ \left( \frac{\beta}{\alpha} \right)^{\beta\gamma} + \left( \frac{\alpha}{\beta} \right)^{\beta\gamma} \right] r^{\alpha\gamma}. \quad (1)
\]

Note that \((pD_i/w_i)/\theta_i = \alpha/\beta\), which means that \(\alpha\gamma\) denotes the share of costs with improvement of waiting time in total costs and \(\beta\gamma\) denotes the share of costs without improvement of waiting time in total costs. The intuition behind this result is that if \(\alpha\) is larger (smaller) than \(\beta\), which implies that patients put more (less) value on waiting time than on other factors except for waiting.
time, then the costs with (without) improvement of waiting time required to attract patients will be more.

The average cost is given by

\[ c_i(q_i)/D_i = AD_i(q_1, q_2)^{-\beta\gamma}q_i^{\gamma}, \] (2)

which implies that there will be economies of scale if the comprehensive quality of hospital care is fixed. However, an expansion of demand will require an increase in the comprehensive quality level. If \( \gamma > 1 \), economies of scale appear only when the quantity of hospital care is small, which is consistent with the empirical results indicated by Aletras (1999). Equation (2) also implies that a decrease in \( \beta\gamma \) and increase in \( \gamma \) weaken economies of scale; in other words, economies of scale weaken as the share of the costs without improvement of waiting time in total costs decreases or as the cost function becomes more convex with regard to the comprehensive quality of hospital care. Accordingly, the profit function of the local public hospitals with quality level \( q_i \) is given by

\[ \pi(q_i) = pD_i(q_1, q_2) - c(q_i) + T_i, \]

where \( T_i \) is a lump-sum transfer received from the central government.

3 First-best outcome

Although an increase in capitation payment would provide an incentive to hospitals to compete for acquiring patients and lead them to improve their comprehensive quality of care and waiting time, the central government would not adjust a capitation payment to such a high value that its costs would be significantly higher than its benefits. In this section, in order to determine the reasonable amount of capitation payment, we derive the first-best outcome in a case where the two hospitals in the model are managed by the central government.

The central government decides the acceptable quality level of each hospital for maximizing social welfare, which is calculated as the summation of the surplus of the residents and profits of the local governments minus transfer payments from the central government. Social welfare \( S \) is estimated as follows:

\[ S = N_1u_1(0) + \int_0^X u_1(x)dx + \int_0^1 u_2(x)dx + N_2u_2(1) + \pi_1 + \pi_2 \]

\[ - (N_1 + N_2 + 1)(1 - r)p - T_1 - T_2 \]

\[ = (N_1 + N_2 + 1)v - t \left( \int_0^X xdx + \int_1^X (1 - x)dx \right) + q_1D_1(q_1, q_2) + q_2D_2(q_1, q_2) \]

\[ - AD_1^{\alpha\gamma}q_1^{\gamma} - AD_2^{\alpha\gamma}q_2^{\gamma}. \]

The optimal levels \( q_1 \) and \( q_2 \) necessary for the maximization of \( S \) are denoted by \( q_1^{OPT} \) and \( q_2^{OPT} \), respectively. When \( \gamma \leq 1 \) (in other words, \( \alpha + \beta \geq 1 \)), \( \lim_{\gamma \to \infty} S = \infty \), where finite \( q_1^{OPT} \) and \( q_2^{OPT} \) do not exist. Therefore, we assume that \( \gamma > 1 \) (in other words, \( \alpha + \beta < 1 \)). Since it is unrealistic that only one facility provides hospital care to all residents in the two cities and one village, we consider only those cases in which the demand on each local public hospital is more than 0 under \( q_1 = q_1^{OPT} \) and \( q_2 = q_2^{OPT} \); that is, \( 0 < X < 1 \) is satisfied.

The first-order conditions for the maximization of \( S \) are

\[ \frac{\partial S}{\partial q_1} = D_1 - A\gamma \left[ \alpha(D_1^{\alpha\gamma}q_1^{\gamma} - D_2^{\alpha\gamma}q_1^{\gamma})\frac{dD_1}{dq_1} + D_1^{\alpha\gamma}q_1^{\gamma-1} \right] = 0, \] (3)

\[ \frac{\partial S}{\partial q_2} = D_2 - A\gamma \left[ \alpha(D_2^{\alpha\gamma}q_2^{\gamma} - D_1^{\alpha\gamma}q_2^{\gamma})\frac{dD_2}{dq_2} + D_2^{\alpha\gamma}q_2^{\gamma-1} \right] = 0, \] (4)
which lead to the following lemma.

**Lemma 1.** \( q_1^{OPT} \) and \( q_2^{OPT} \) satisfy the condition \( q_1^{OPT} > q_2^{OPT} \), and \( w_1(q_1^{OPT}, q_2^{OPT}) < w_2(q_1^{OPT}, q_2^{OPT}) \).

**Proof:** See Appendix A.

This lemma implies that the hospital in city 1 (large city) is better than the hospital in city 2 (small city) in terms of quality as well as waiting time. The qualities based on resources that do not influence waiting time are independent of demand; for example, an increase in the number of patients does not influence the quality of the medical technology employed. Therefore, the costs per unit of demand with regard to the resources that do not influence waiting time decrease with an increase in the amount of demand; thus, it is cost-effective to invest additional resources in a hospital located in a large city, which has a large demand, rather than in a hospital located in a small city. In addition, in order to utilize these resources, it is better for a hospital located in a large city to perform more operations per amount of time by reducing its waiting time as compared to a hospital in a small city. Accordingly, the comprehensive quality of hospital care in city 1 is higher than that in city 2, and the waiting time for hospital care in city 1 is shorter than that in city 2.

### 4 Competition among local governments

In this section, we derive the equilibrium condition under which local public hospitals compete in terms of quality. For the same reason indicated in section 3, we assume that \( 0 < X < 1 \) is satisfied under equilibrium, and we further assume that \( \gamma > 1 \). Moreover, since the central government would not adjust a capitation payment to such a high level that it would cost significantly higher than its benefits, we only consider the equilibrium condition in which the quality of hospital care in each city does not differ significantly from the quality of hospital care in each city at the first-best outcome. Since we transform (3) and (4) to

\[
\frac{\partial S}{\partial q_1} = \left[ 1 - A \gamma \frac{1}{q_1} D_1^{-\beta \gamma} q_1^{\gamma} \right] D_1 - A \gamma \left[ \alpha (D_1^{-\beta \gamma} q_1^{\gamma} - D_2^{-\beta \gamma} q_2^{\gamma}) \frac{dD_1}{dq_1} \right]
\]

and

\[
\frac{\partial S}{\partial q_2} = \left[ 1 - A \gamma \frac{1}{q_2} D_2^{-\beta \gamma} q_2^{\gamma} \right] D_2 - A \gamma \left[ \alpha (D_2^{-\beta \gamma} q_2^{\gamma} - D_1^{-\beta \gamma} q_1^{\gamma}) \frac{dD_2}{dq_2} \right],
\]

respectively, the second-order conditions for the maximization of \( S \) would require \( \partial(D_i^{-\beta \gamma} q_i^{OPT})/\partial q_i > 0 \), which is accomplished for \((q_1^{OPT}, q_2^{OPT}) \in \{(q_1, q_2)|D_1 - \beta q_1 D_1/\partial q_1 > 0 \text{ and } D_2 - \beta q_2 D_2/\partial q_2 > 0 \}\). Accordingly, we consider only the equilibrium at which the quality of each hospital is indicated in the gray area of Figure 1.

Since a local public hospital owned by a city government must focus on the city’s social welfare, it determines the quality of its care in order to maximize the social welfare of the city, denoted by \( s_i \), in the following manner:

\[
s_1 = N_1 u_1(0) + \pi_1
= [v + q_1 + (1 - r)p] N_1 + p X(q_1, q_2) + T_1 - AD_1^{\alpha \gamma} q_1^{\gamma}, \tag{5}
\]

\[
s_2 = N_2 u_2(1) + \pi_2
= [v + q_2 + (1 - r)p] N_2 + p (1 - X(q_1, q_2)) + T_2 - AD_2^{\alpha \gamma} q_2^{\gamma}. \tag{6}
\]

---

7If \( \beta = 0 \), this cost-effectiveness disappears, and we obtain \( q_1^{OPT} = q_2^{OPT} \), which is consistent with the result of Aiura and Sanjo (2010).

8The central government, which considers the positive externalities of health care, would adjust a capitation payment whose costs reasonably outweigh its direct benefits to residents.
The first-order conditions for maximizing $s_1$ and $s_2$ are given by

\[
R_1 = \frac{\partial s_1}{\partial q_1} = N_1 + p \frac{\partial X}{\partial q_1} - A \gamma \left( \alpha D_1^{\beta_1} q_1^{\gamma} \frac{\partial D_1}{\partial q_1} + D_1^\alpha q_1^{-1} \right) = 0
\]

and

\[
R_2 = \frac{\partial s_2}{\partial q_2} = N_2 - p \frac{\partial X}{\partial q_2} - A \gamma \left( \alpha D_2^{\beta_2} q_2^{\gamma} \frac{\partial D_2}{\partial q_2} + D_2^\alpha q_2^{-1} \right) = 0,
\]

respectively.

We obtain the Nash equilibrium by identifying the intersection of the two response functions, $R_1(q_1, q_2) = 0$ and $R_2(q_1, q_2) = 0$. The slopes of $R_1(q_1, q_2) = 0$ and $R_2(q_1, q_2) = 0$ are derived in the following manner:

\[
\left. \frac{dq_2}{dq_1} \right|_{R_1(q_1, q_2)=0} = -\frac{\partial R_1/\partial q_1}{\partial R_1/\partial q_2} = -(A_k + A_l) \left[ 2\alpha^2 \frac{\partial D_2}{\partial q_2} (D_2 - \beta q_2 \frac{\partial D_2}{\partial q_2} D_2^{-(\beta + 1)} q_2^{-1} + \gamma (\beta \alpha q_2^2 (\frac{\partial D_2}{\partial q_2})^2 + (\gamma - 1) D_2^2 ) D_1^{-(\beta + 1)} q_1^{-1} \right]
\]

and

\[
\left. \frac{dq_2}{dq_1} \right|_{R_2(q_1, q_2)=0} = -\frac{\partial R_2/\partial q_1}{\partial R_2/\partial q_2} = -(A_k + A_l) \left[ 2\alpha^2 \frac{\partial D_2}{\partial q_2} (D_2 - \beta q_2 \frac{\partial D_2}{\partial q_2} D_2^{-(\beta + 1)} q_2^{-1} + \gamma (\beta \alpha q_2^2 (\frac{\partial D_2}{\partial q_2})^2 + (\gamma - 1) D_2^2 ) D_2^{-(\beta + 1)} q_2^{-1} \right].
\]

respectively. Since we confirm that $D_i - \beta q_i \frac{\partial D_i}{\partial q_i} > 0$ and $\partial D_1/\partial q_1 = \partial D_2/\partial q_2 = -\partial D_1/\partial q_2 = -\partial D_2/\partial q_1$, the slope of $R_1(q_1, q_2) = 0$ is greater than 2, and the slope of $R_2(q_1, q_2) = 0$ is between 0 and 1/2. Moreover, we can show that $q_1|_{R_1=0}$ and $D_1=D_2 > q_2|_{R_2=0}$ and $D_1=D_2$. Figures 2 and 3 illustrate the two response functions. We propose the following lemma:

**Lemma 2.** A unique Nash equilibrium exists, if the intersection of $R_1(q_1, q_2) = 0$ and $X(q_1, q_2) = 0$ is located at the upper right of the intersection of $R_2(q_1, q_2) = 0$ and $X(q_1, q_2) = 0$, and if the intersection of $R_1(q_1, q_2) = 0$ and $X(q_1, q_2) = 1$ is located at the lower left of the intersection of $R_2(q_1, q_2) = 0$ and $X(q_1, q_2) = 1$. At the unique Nash equilibrium, the demand for hospital care in city 1 is higher than that in city 2, that is, $D_1(q_1^*, q_2^*) > D_2(q_1^*, q_2^*)$, where $q_1^*$ and $q_2^*$ denote the comprehensive quality levels at the equilibrium.

Lemma 2 implies that an intuitively recognizable equilibrium exists in which the hospital in the large city is superior to that in the small city in terms of comprehensive quality. Section 5 shows such an equilibrium, for instance, in numerical analysis. The reason for such an equilibrium to exist is that there are resources that do not influence waiting time. As mentioned in section 3, the hospital in the large city rather than that in the small city is cost-effective in terms of investment of these resources; thus such an equilibrium exists. Further, the following additional lemma holds with regard to the two response functions.

**Lemma 3.** Both $R_1(q_1, q_2) = 0$ and $R_2(q_1, q_2) = 0$ move away from the origin as $p$ increases.

**Proof:** It is confirmed that $D_i^{\beta_i} q_i^\gamma$ and $D_i^\alpha q_i^{-1}$ are increasing functions.

---

9If $\beta = 0$, this cost-effectiveness disappears, and such equilibrium does not exist, which is consistent with the result of Aiura and Sanjo (2010).
The following propositions can be derived from Lemma 3:

**Proposition 1.** In the competitive equilibrium among local public hospitals, the comprehensive quality of hospital care in both the hospitals increases with an increase in capitation payment.

**Proposition 2.** In the competitive equilibrium among local public hospitals, waiting time shortens with an increase in capitation payment.

**Proof:**

\[
\frac{dw_i(q_i^*)}{dp} = -A\gamma \left( D_i - \beta q_i^* \frac{\partial D_i}{\partial q_i} \right) D_i^{-\alpha \gamma q_i^* - (\gamma + 1)} \frac{dq_i^*}{dp} + \beta \gamma D_i^{-\alpha \gamma q_i^* - \gamma} \frac{\partial D_i}{\partial q_i} \frac{dq_i^*}{dp}.
\]

Here, \( i \neq -i \in \{1, 2\} \). We confirm that \( D_i - \beta q_i(\partial D_i/\partial q_i) > 0 \) and \( \partial D_i/\partial q_i < 0 \). Moreover, Proposition 1 shows that \( dq_i^*/dp > 0 \) and \( dq_{-i}^*/dp > 0 \). Therefore, we derive that \( \frac{dw_i(q_i^*)}{dp} < 0 \). □

If capitation payment is sufficient, it would serve as an incentive for city hospitals to provide the residents of the village with hospital care. Therefore, as the capitation payment increases, the hospitals in each city increase their comprehensive quality of health care in order to acquire patients from the village. Further, Proposition 2 implies that the waiting time in each hospital improves with an increase in the capitation payment amount.

In order to determine the difference in the improvement of comprehensive quality of hospital care provided by the two hospitals as a result of an increase in capitation payment, we derive \( dq_1^*/dp \) and \( dq_2^*/dp \) by solving

\[
\begin{pmatrix}
\frac{\partial R_1}{\partial q_1} & \frac{\partial R_1}{\partial q_2} \\
\frac{\partial R_2}{\partial q_1} & \frac{\partial R_2}{\partial q_2}
\end{pmatrix}
\begin{pmatrix}
\frac{dq_1^*}{dp} \\
\frac{dq_2^*}{dp}
\end{pmatrix} = -\begin{pmatrix}
\frac{\partial R_1}{\partial p} \\
\frac{\partial R_2}{\partial p}
\end{pmatrix},
\]

which yields

\[
\frac{dq_1^*}{dp} - \frac{dq_2^*}{dp} = \frac{(R_1/\partial q_1 + R_1/\partial q_2)(\partial R_2/\partial p) - (R_2/\partial q_1 + R_2/\partial q_2)(\partial R_1/\partial p)}{(R_1/\partial q_1)(R_2/\partial q_2) - (R_1/\partial q_2)(R_2/\partial q_1)}
\]

\[
= \frac{1}{\Delta} \gamma \left[ \frac{\alpha \gamma}{2t} (D_2^{-\beta \gamma q_2^{-1}} - D_1^{-\beta \gamma q_1^{-1}}) + (\gamma - 1)(D_2^{\alpha \gamma q_2^{-2}} - D_1^{\alpha \gamma q_1^{-2}}) \right],
\]

where \( \Delta = (R_1/\partial q_1)(R_2/\partial q_2) - (R_1/\partial q_2)(R_2/\partial q_1) > 0 \).

The two hospitals determine the quantity of care at which marginal revenues equal marginal costs. Since owing to an increase in capitation payment the additional marginal revenues are identical for both the hospitals, the difference in the additional marginal costs between the two hospitals as a result of an increase in the capitation payment reflects the manner in which the difference in the comprehensive quality of health care between the two hospitals changes with an increase in capitation payment. Hospitals incur additional costs as a result of an increase in the quantity of care provided by them. Further, hospitals also incur additional costs as a result of an increase in the comprehensive quality of care, because an increase in the quantity of hospital care services requires an increase in the comprehensive quality of hospital care. Therefore, as the two terms on the right-hand side of (7) imply, the difference in these additional marginal costs is divided into the following two parts: the first term implies the difference in the additional marginal costs between the two hospitals as a result of an increase in the quantity of hospital care, and the second implies the difference in the additional marginal costs between the two hospitals as a result of an increase in the comprehensive quality of hospital care.
We focus on the case where \( q_1 > q_2 \) to check the sign of (7), because the hospital in the large city realistically seems to be superior to the hospital in the small city in terms of comprehensive quality.

When \( \beta \gamma \) is approximately 0 and \( q_1 > q_2 \), the sign of the first term on the right-hand side of (7), \( \alpha \gamma (D_2^{\gamma} q_2^{\gamma-1} - D_1^{\gamma} q_1^{\gamma-1})/(2t) \), is negative. In other words, when the costs with improvement of waiting time almost corresponds to the total costs and the comprehensive quality of hospital care in the large city is superior to that in the small city, the hospital in the small city has an advantage over the hospital in the large city in terms of increasing the quantity of hospital care. Since an improvement of waiting time would require resources in proportion to the demand, economies of scale do not work well in a situation where improvement of waiting time is widely claimed to attract patients. Thus, the costs required for increasing the quantity by one unit are higher in the hospital located in the large city than in the hospital located in the small city. On the other hand, when the costs without improvement of waiting time almost corresponds to the total costs, the hospital in the large city has an advantage over the hospital in the small city in terms of increasing the quantity of hospital care. Since the quality of hospital care on the basis of resources that do not influence waiting time is independent of demand, economies of scale work well in the situation where the other resources except waiting time are widely claimed to attract patients. Thus, as compared to the hospital located in the small city, the costs required for increasing the quantity by one unit are lower in the hospital located in the large city owing to the economies of scale.

However, note that if \( \beta \gamma \) is approximately 1, or if \( t \) is sufficient, the value of the first term on the right-hand side of (7) is approximately 0. When \( t \) is sufficient, residents ascribe greater importance to the distance between their residences and the hospitals than to the comprehensive quality of hospital care; thus, an increase in capitation payment does not affect a large change in the demand for hospital care, and consequently the impact on the first term on the right-hand side of (7) is small. On the other hand, when \( \beta \gamma \) is approximately 1, since economies of scale work effectively even with small quantities, the difference between the two hospitals is rather small in terms of economies of scale; thus, the impact on the first term on the right-hand side of (7) is small.

When \( q_1 > q_2 \), the sign of the second term on the right-hand side of (7), \( (\gamma - 1)(D_2^{\alpha \gamma} q_2^{\gamma - 2} - D_1^{\alpha \gamma} q_1^{\gamma - 2}) \), depends significantly on \( \gamma \). In other words, the curvature of the marginal costs with regard to quality contributes considerably to the difference in additional marginal costs with an increase in the comprehensive quality of hospital care between the two hospitals when the hospital in the large city offers higher quality than that in the small city. Since resources in proportion to the demand are required for improvement of waiting time, the hospital in the large city, which has a large demand, has to incur higher costs for improving its waiting time than the hospital in the small city. Thus, if the marginal costs of the two hospitals are the same, the additional marginal costs of the hospital in the large city are higher than that of the hospital in the small city. However, the marginal costs of hospitals differ according to the quality of hospital care offered by them. If the marginal costs are convex with regard to the comprehensive quality of hospital care, then the higher the comprehensive quality, the higher the additional marginal costs required for improving the comprehensive quality by a certain amount. Therefore, the additional marginal cost resulting from an increase in the comprehensive quality of hospital care is higher for the hospital in the large city, which offers a higher comprehensive quality of hospital care, than that for the hospital in the small city, which offers a lower comprehensive quality of hospital care. On the contrary, if marginal costs are concave with regard to the comprehensive quality of hospital care, the higher the comprehensive quality of hospital care, the lower the additional marginal costs that are sufficient for improving the comprehensive quality by a certain amount. Therefore, we cannot conclude which of these two hospitals incurs a higher amount of additional marginal costs as a result of an increase.
in the comprehensive quality of hospital care, but if \( \alpha \gamma \) is approximately 0, the hospital in the small city, which offers a lower comprehensive quality of hospital care, may incur higher additional marginal costs with an increase in the comprehensive quality of hospital care than the hospital in the large city, which offers a higher comprehensive quality of hospital care.

Conclusively, on the basis of (7), the following proposition holds:

**Proposition 3.** When an improvement of waiting time is widely claimed to attract patients and the cost function is sufficiently convex with regard to the comprehensive quality of hospital care, an increase in capitation payment reduces the difference in the comprehensive quality of hospital care between the two hospitals where the hospital in city 1 (large city) offers a better comprehensive quality than that offered by the hospital in city 2 (small city). On the contrary, when quality improvement except for waiting time is widely claimed to attract patients and the cost function is not sufficiently convex with regard to the comprehensive quality, an increase in capitation payment increases the difference in the comprehensive quality of hospital care provided by these two hospitals.

**Proof:** We assume that \( q_1 \geq q_2 \). Since \( \lim_{\alpha \gamma \to 1} \alpha \gamma = 0 \) and \( \lim_{\alpha \gamma \to 1} \beta \gamma = 0 \), when \( \gamma \geq 2 \),

\[
\lim_{\alpha \gamma \to 1} \text{sgn} \left[ \frac{dq_1^*}{dp} - \frac{dq_2^*}{dp} \right] = \text{sgn} \left[ \frac{1}{2t}(q_2^{\gamma-1} - q_1^{\gamma-1}) + (\gamma - 1)(D_2 q_2^{\gamma-2} - D_1 q_1^{\gamma-2}) \right] < 0.
\]

When \( 1 < \gamma < 2 \),

\[
\lim_{\beta \gamma \to 1} \text{sgn} \left[ \frac{dq_1^*}{dp} - \frac{dq_2^*}{dp} \right] = \text{sgn} \left[ (\gamma - 1)(q_2^{\gamma-2} - q_1^{\gamma-2}) \right] > 0.
\]

These equations confirm Proposition 3.

Economies of scale work well when quality improvement except for waiting time is widely claimed to attract patients and the cost function is not sufficiently convex with regard to the comprehensive quality of hospital care, whereas economies of scale do not work well when improvement of waiting time is widely claimed to attract patients and the cost function is sufficiently convex with regard to the comprehensive quality; thus, Proposition 3 implies that the strength of economies of scale influences the difference in comprehensive quality of hospital care between the two cities as a result of an increase in capitation payment. The hospital industry is generally categorized as a labor-intensive industry; thus, economies of scale do not work well. Further, the comprehensive quality of health care in a large city appears to be superior to that in a small city. Therefore, as indicated by Proposition 3, an increase in capitation payment reduces the difference in the comprehensive quality of hospital care between the two hospitals, where the hospital in city 1 (large city) offers better comprehensive quality hospital care than that offered by the hospital in city 2 (small city).

If \( D_1 > D_2 \), the cost that the local public hospital in city 1 incurs for improvement of waiting time is larger than that incurred by the local public hospital in city 2. Moreover, when the costs without improvement of waiting time almost corresponds to the total costs, the comprehensive quality of hospital care almost depends on waiting time. Therefore, as compared to the local public hospital in city 1, the local public hospital in city 2 can increase the comprehensive quality of its hospital care at a lower cost; thus, the local public hospital in city 2 has a cost advantage over the local public hospital in city 1 in terms of competing to acquire patients from the village. In addition, an increase in capitation payment makes the local public hospitals in both the cities more enthusiastic regarding acquiring patients from the village. As a result, an increase in capitation payment makes the local public hospital in city 2, which has a cost advantage, more enthusiastic than the local public hospital in city 1 with respect to acquiring patients from the village; thus, the difference in the comprehensive quality of hospital care between cities 1 and 2 decreases.
As described above, when the costs with improvement of waiting time almost corresponds to the total costs, the improvement of waiting time plays a major role in reducing the difference in the quality of hospital care between cities 1 and 2. However, the comprehensive quality of hospital care is partially influenced by the resources that do not influence waiting time, which creates economies of scale and weakens the reduction of the difference in the comprehensive quality of hospital care owing to improvement of waiting time. Accordingly, the reduction of the difference in the comprehensive quality of hospital care between the hospitals in cities 1 and 2 is less than that in waiting time between the hospitals in these two cities. On the contrary, the reduction in the difference in waiting time between the hospitals in cities 1 and 2 is more than that in the comprehensive quality of hospital care between the hospitals in these two cities. Proposition 4 presents its extreme case.

**Proposition 4.** The equilibrium at which the comprehensive quality of hospital care in city 1 is higher than that in city 2 (i.e., \( q_1^* > q_2^* \)) does not guarantee that waiting time for the hospital in city 1 is shorter than that for the hospital in city 2 (i.e., \( w_1(q_1^*) < w_2(q_2^*) \)).

**Proof:** If \( q_1 \geq q_2, D_1 > D_2 \) is satisfied. Therefore,

\[
\lim_{q_1 \to q_2} (w_1(q_1) - w_2(q_2)) = \frac{1}{\alpha^\gamma} \left( \frac{\beta}{\alpha} \right)^{\beta \gamma} \left( D_1^{\beta \gamma} - D_2^{\beta \gamma} \right) q_2^{-\gamma} > 0,
\]

which implies that \( w_1(q_1^*) > w_2(q_2^*) \) if \( q_1^* \) is slightly more than \( q_2^* \).

Proposition 4 implies that the hospital whose waiting time is shorter than that of the other hospital is not always superior to the other hospital in terms of comprehensive quality. Although we often focus on the ratio of doctors to population using public censuses in order to determine the difference in the health-care environment between regions, Proposition 4 cautions that we may make faulty decisions if we rely only on this metric. For example, OECD Regions at a Glance 2009 shows that Japan has a more balanced regional distribution of physicians than most European countries. Although this implies that access to health care in Japan is even between rural and urban regions, it does not imply that the comprehensive quality of health care in the rural regions of Japan is certainly comparable to that of its urban regions.

## 5 Numerical analysis

Proposition 3 implies that in the labor-intensive hospital industry, disparity of the comprehensive quality between large and small cities fades away with an increase in capitation payment, but Proposition 3 cites only the extreme case when \( \alpha \gamma \) or \( \beta \gamma \) is close to 1. Therefore, in this section, we show that this implication remains valid within realistic parameters by using numerical analysis.

Generally speaking, the share of labor costs in total costs of a hospital is between 50% and 70%\(^{10} \). Since health care is a service that doctors and nurses provide by spending time, labor is a resource that influences waiting time; thus, we set \( \alpha \gamma = 1 - \beta \gamma \) to either 0.55 or 0.65. In order to encourage the effect of economies of scale, we set \( 1/\gamma = \beta + \alpha = 0.9 \). Further, we assume two cases for \( t \), \( t = 0.5 \) and \( t = 1 \), in order to investigate the effect of transport improvements. Regarding the other parameters, we set \( N_1 = 2, N_2 = 1.5, v = 1, a = 1.25, w = 1, \) and \( T_1 = T_2 = 0 \).

Figure 4 depicts the change of \( q_1, 1/w_1, \) and \( S \) with \( p \) for four cases with different parameter values. In all cases, the hospital in city 1 is superior to the hospital in city 2 in terms of comprehensive quality (\( q \)) and shortness of waiting time (\( w \)) when the capitation payment (\( p \)) is small, and these

\(^{10}\)For example, the share of labor costs in total costs was 57.1% in 2008 in the United Kingdom (as shown by the UK Centre for the Measurement of Government Activity, 2010) and 67.9% between 1997 and 2001 in Switzerland (as shown by Farsi and Filippini, 2006).
disparities decrease with an increase in capitation payment. Further, in all cases, the waiting time for hospital care is more effective than the comprehensive quality of hospital care in terms of reduction of these disparities. These results correspond to and support the discussions presented in section 4.

When we compare the results between the cases of $\alpha \gamma = 0.55$ and $\alpha \gamma = 0.65$, the reduction of these disparities is faster when $\alpha \gamma = 0.65$ than when $\alpha \gamma = 0.55$. The economies of scale in hospitals weaken as hospitals become more labor-intensive. Therefore, the higher the share of costs with improvement of waiting time in the total costs, the faster is the reduction of disparities as a result of an increase in capitation payment between the two hospitals. Further, when we compare the results between the cases of $t = 0.5$ and $t = 1$, the reduction of disparities between the two hospitals is faster in the case of $t = 0.5$ than in the case of $t = 1$, because the reduction of $t$ encourages competition between hospitals.

Social welfare in all regions follows an inverse U-shaped relationship with capitation payment. An adequate capitation payment improves the quality of hospital care as well as social welfare. However, excessive amounts of capitation payments worsen social welfare.

6 Discussion

6.1 Effect of patient choice

We discuss the effect of patient choice in the subsection. We assume that residents are not permitted free choice in hospitals, that the patients living in $[0, 1/2]$ are compelled to visit the hospital in city 1, and that the other patients are compelled to visit the hospital in city 2.

The cities’ social welfare, $s_1$ and $s_2$, is reformed from (5) and (6) to

$$s_1 = [v + q_1 + (1 - r)p]N_1 + \frac{p}{2} + T_1 - A(N_1 + 1/2)^{\alpha \gamma q_1^\gamma}$$

and

$$s_2 = [v + q_2 + (1 - r)p]N_2 + \frac{p}{2} + T_2 - A(N_2 + 1/2)^{\alpha \gamma q_2^\gamma}.$$ 

The first-order conditions to maximize $s_1$ and $s_2$ are given by

$$\hat{R}_1 = \frac{\partial s_1}{\partial q_1} = N_1 - \gamma A(N_1 + 1/2)^{\alpha \gamma q_1^\gamma - 1} = 0$$

and

$$\hat{R}_2 = \frac{\partial s_2}{\partial q_2} = N_2 - \gamma A(N_2 + 1/2)^{\alpha \gamma q_2^\gamma - 1} = 0,$$

respectively, which imply that the quality of hospital care in both the cities is independent of the capitation payment from the central government. Patients from the village visit the specified hospitals irrespective of the quality of hospital care, and the hospitals obtain a certain amount of capitation payment; thus, an increase in a capitation payment by the central government is not an incentive for city hospitals to improve their comprehensive quality. The comprehensive quality of the city hospitals, which only consider maximizing their city’s social welfare, would be significantly lower than the comprehensive quality at the first-best outcome. Therefore, if the central government invokes an adequate amount of capitation payment, introducing a free choice of hospitals would raise the comprehensive quality levels of hospital care in both the hospitals. Further, drawing an analogy with the proof offered by Proposition 2, it can be said that the waiting time of the hospitals in both the cities would also improve.

Even if the location of this boundary is changed, the discussion in this subsection holds.
6.2 Extended province

In the previous sections we assumed that the two cities and the village (agricultural area) are governed by different local governments; consumer surplus in the village was not included in the objective function of the hospitals owned by the cities’ governments. However, local governments may govern not only an urban area but also a portion of an agricultural area. Under this assumption, in this subsection we consider the consumer surplus in the agricultural area governed by the cities’ governments in the objective function of hospitals.

In this subsection, we assume that the local government in city 1 governs the agricultural area in [0, 1/2] as well as city 1 and the local government in city 2 governs the agricultural area in (1/2, 0] as well as city 2. When \( q_1 > q_2 \), \( s_1 \) and \( s_2 \), are reformed from (5) and (6) to

\[
\tilde{s}_1 = [v + q_1 + (1 - r)p](N_1 + 1/2) + p(X(q_1, q_2) - 1/2) + T_1 - AD_1^{\alpha\gamma}q_1^\gamma t \int_0^{1/2} xdx
\]

and

\[
\tilde{s}_2 = [v + q_2 + (1 - r)p]D_2 + (v + q_1 - rp)(X(q_1, q_2) - 1/2) + T_2 - AD_2^{\alpha\gamma}q_2^\gamma t \left( \int_{1/2}^X xdx + \int_{X}^1 (1 - x)dx \right).
\]

The first-order conditions for maximizing \( s_1 \) and \( s_2 \) are given by

\[
\tilde{R}_1 = \frac{\partial \tilde{s}_1}{\partial q_1} = (N_1 + 1/2) + p \frac{\partial X}{\partial q_1} - A\gamma \left( \alpha D_1^{\beta\gamma}q_1^{\gamma} \frac{\partial D_1}{\partial q_1} + D_1^{\alpha\gamma}q_1^{-1} \right) = 0
\]  

\[\text{(8)}\]

and

\[
\tilde{R}_2 = \frac{\partial \tilde{s}_2}{\partial q_2} = (N_2 + 1/2) - p \frac{\partial X}{\partial q_2} + (q_1 - q_2) \frac{\partial X}{\partial q_2} - A\gamma \left( \alpha D_2^{\beta\gamma}q_2^{\gamma} \frac{\partial D_2}{\partial q_2} + D_2^{\alpha\gamma}q_2^{-1} \right) = 0,
\]  

\[\text{(9)}\]

respectively. From (8), the slope of \( \tilde{R}_1(q_1, q_2) = 0 \) is identical to the slope of \( R_1(q_1, q_2) = 0 \) and \( \tilde{R}_1(q_1, q_2) = 0 \) moves away from the origin as \( p \) increases. On the other hand, from (9),

\[
\left. \frac{dq_2}{dq_1} \right|_{\tilde{R}_2(q_1, q_2)=0} = - \left. \frac{\partial \tilde{R}_2/\partial q_1}{\partial \tilde{R}_2/\partial q_2} \right|_{\tilde{R}_2(q_1, q_2)=0} = - \left. \frac{\partial R_2/\partial q_1 + \partial X/\partial q_2}{\partial R_2/\partial q_2 - \partial X/\partial q_2} \right|_{R_2(q_1, q_2)=0}.
\]

Since \( \partial \tilde{R}_2/\partial q_2 < 0 \) according to the second-order condition for the maximization of \( \tilde{s}_2 \), the slope of \( R_2(q_1, q_2) = 0 \) is less than 1/2, and \( R_2(q_1, q_2) = 0 \) moves away from the origin as \( p \) increases. Note that the slope of \( R_2(q_1, q_2) = 0 \) may be negative; thus, the comprehensive quality and waiting time of the hospital in city 1 improves with an increase in capitation payment. However, we cannot conclude whether or not the comprehensive quality and waiting time of the hospital in city 2 improve (in other words, although Proposition 1 holds with respect to the hospital in city 1, it does not hold with respect to the hospital in city 2). If the difference in quality between the two hospitals decreases, the comprehensive quality and waiting time of the hospital in city 2 would improve with an increase in capitation payment, just as those of the hospital in city 1.

\[12\] This is because \( \frac{dq_2}{dq_1} \bigg|_{\tilde{R}_2(q_1, q_2)=0} = - \left. \frac{\partial \tilde{R}_1/\partial q_1}{\partial \tilde{R}_1/\partial q_2} \right|_{\tilde{R}_1(q_1, q_2)=0} = - \left. \frac{\partial R_1/\partial q_1}{\partial R_1/\partial q_2} \right|_{R_1(q_1, q_2)=0} = \frac{dq_2}{dq_1} \bigg|_{R_1(q_1, q_2)=0}. \]
Since the difference in the change in the comprehensive quality of hospital care between the two hospitals as a result of an increase in capitation payment is homogeneous to (7)\(^{13}\), Proposition 3 holds. Thus, when an improvement of waiting time is widely claimed to attract patients, the comprehensive quality of hospital care and waiting time in both the hospitals improve with an increase in capitation payment, and the difference in the comprehensive quality and waiting time between the two hospitals reduces. On the contrary, when improvement of waiting time is not claimed to attract patients and economies of scale work adequately, the difference in the comprehensive quality of hospital care and waiting time between the two hospitals may widen and the quality of hospital care and waiting time of the hospital in city 2 may worsen.

An increase in capitation payment will definitely incentivize hospitals to offer better hospital care to their residents. When the share of the costs to improve waiting time in total costs is sufficient, the local government that governs the small city experiences a cost advantage as described in section 4 and would increase the comprehensive quality of hospital care of its own hospital in order to receive additional capitation payment. However, if economies of scale work adequately, then it may be more desirable for the local government that governs the small city to entrust its residents to the hospital in the large city that offers better comprehensive quality of hospital care by efficiently exploiting a capitation payment from the economies of scale than to obtain more capitation payment by encouraging its residents to visit its own hospital; thus, the local government that governs the small city may reduce the comprehensive quality of its hospital care.

7 Conclusion

This paper illustrates the effect of capitation payment on the comprehensive quality of hospital care and waiting time between two hospitals that are located in two different cities and compete for acquiring patients. The hospitals, which are labor-intensive, improve their comprehensive quality of hospital care and waiting time with an increase in capitation payment. However, the extent of these improvements differs according to the population size of the cities: the hospital in the small city shows greater improvements in the comprehensive quality of hospital care and waiting time than the hospital in the large city. This result implies that regional disparity in the quality of hospital care decreases with an increase in capitation payment. Further, reduction in the disparity of waiting time between the two hospitals is greater than that of the comprehensive quality of hospital care.

In summary, the impact of incentives for hospitals to compete for acquiring patients on the basis of their comprehensive quality and waiting time depends on the population size of the city in which the hospitals are located. Therefore, we must pay attention to the population around the hospitals while theoretically and empirically analyzing the effect of policies that induce public hospitals to compete with each other.

Acknowledgements

We would like to thank William M. Doerner, Tomoya Ida, Toshiharu Ishikawa, Akio Kawasaki, Yuichi Kimura, Tatsuaki Kuroda, Noriaki Matsushima, Kenichi Mizuta, Toru Naito, Hikaru Ogawa, Hiroshi Ono, Sasuo Sanjo, Yasuhiro Sato, Norio Shimoda, Luigi Siciliani, and Hiroyuki Takami. Any mistakes herein are, of course, our own. I acknowledge the financial support from KAKENHI (22330095).

Appendix A: Proof of Lemma 1

\(^{13}\)This is because \(\bar{R}_1/\partial q_1 + \bar{R}_1/\partial q_2 = R_1/\partial q_1 + R_1/\partial q_2, \bar{R}_2/\partial q_1 + \bar{R}_2/\partial q_2 = R_2/\partial q_1 + R_2/\partial q_2,\) and \(\Delta = (\bar{R}_1/\partial q_1)(\bar{R}_2/\partial q_2) - (\bar{R}_1/\partial q_2)(\bar{R}_2/\partial q_1) > 0.\)
First, we confirm that \( D_1(q_1^{OPT}, q_2^{OPT}) \geq D_2(q_1^{OPT}, q_2^{OPT}) \). \( q_1 \) and \( q_2 \) that satisfies that \( D_1 = D_2 = (N_1 + N_2 + 1)/2 = \tilde{N} \) are denoted by \( \tilde{q}_1 \) and \( \tilde{q}_2 \) respectively. Since \( N_1 > N_2 \), \( \tilde{q}_1 < \tilde{q}_2 \) is satisfied. Moreover,

\[
S|_{q_1=\tilde{q}_1+\epsilon} \text{ and } q_2=\tilde{q}_2-\epsilon-S|_{q_1=\tilde{q}_1-\epsilon} \text{ and } q_2=\tilde{q}_2+\epsilon = (A_k + A_t) \left\{ (\tilde{N} + \epsilon)^{\alpha\gamma} [(\tilde{q}_2 + \epsilon)^{\gamma} - (\tilde{q}_1 + \epsilon)^{\gamma}] - (\tilde{N} - \epsilon)^{\alpha\gamma} [(\tilde{q}_2 - \epsilon)^{\gamma} - (\tilde{q}_1 - \epsilon)^{\gamma}] \right\}.
\]

(10)

Since \( \tilde{q}_1 < \tilde{q}_2, \alpha > 0, \) and \( \gamma \geq 1, \) the right side of (10) is more than 0 if \( \epsilon > 0; \) thus, \( S|_{q_1=\tilde{q}_1+\epsilon} \text{ and } q_2=\tilde{q}_2-\epsilon > S|_{q_1=\tilde{q}_1-\epsilon} \text{ and } q_2=\tilde{q}_2+\epsilon, \) which confirms that \( D_1(q_1^{OPT}, q_2^{OPT}) \geq D_2(q_1^{OPT}, q_2^{OPT}). \)

Further, using (1), we transform (3) and (4) to

\[
\begin{pmatrix}
\frac{\alpha\gamma}{D_1} \frac{dD_1}{dq_1} + \frac{\gamma}{q_1}
\frac{\alpha\gamma}{D_1} \frac{dD_1}{dq_1} \\
\frac{\alpha\gamma}{D_2} \frac{dD_2}{dq_2} + \frac{\gamma}{q_2}
\frac{\alpha\gamma}{D_2} \frac{dD_2}{dq_2}
\end{pmatrix}
\begin{pmatrix}
c_1 \\
c_2
\end{pmatrix}
= \begin{pmatrix}
D_1 \\
D_2
\end{pmatrix}.
\]

(11)

The solution of (11) with respect to \( c_1 \) and \( c_2 \) is

\[
\begin{pmatrix}
c_1 \\
c_2
\end{pmatrix}
= \frac{1}{A} \begin{pmatrix}
\left(\frac{\alpha\gamma}{D_2} \frac{dD_2}{dq_2} + \frac{\gamma}{q_2}\right) D_1 + \frac{\alpha\gamma}{D_2} \frac{dD_2}{dq_2} D_2 \\
\left(\frac{\alpha\gamma}{D_1} \frac{dD_1}{dq_1} + \frac{\gamma}{q_1}\right) D_1 + \frac{\alpha\gamma}{D_1} \frac{dD_1}{dq_1} D_2
\end{pmatrix},
\]

(12)

where

\[
A = \left(\frac{\alpha\gamma}{D_1} \frac{dD_1}{dq_1} + \frac{\gamma}{q_1}\right) \left(\frac{\alpha\gamma}{D_2} \frac{dD_2}{dq_2} + \frac{\gamma}{q_2}\right) - \left(\frac{\alpha\gamma}{D_1} \frac{dD_1}{dq_1} \frac{\alpha\gamma}{D_2} \frac{dD_2}{dq_2} \right) > 0.
\]

(12) is transformed to

\[
\frac{c_1}{c_2} = \frac{D_1 q_1 \left[\left(\alpha\gamma q_2 \frac{dD_2}{dq_2} + \gamma D_2\right) D_1 + \alpha\gamma q_2 D_2 \frac{dD_2}{dq_2}\right]}{D_2 q_2 \left[\left(\alpha\gamma q_1 \frac{dD_1}{dq_1} + \gamma D_1\right) D_2 + \alpha\gamma q_1 D_1 \frac{dD_1}{dq_1}\right]} = \frac{D_1 q_1 (N\alpha q_2 + tD_1 D_2)}{D_2 q_2 (N\alpha q_1 + tD_1 D_2)}.
\]

(13)

Substituting (1) into the left side of (13), we have

\[
\left(\frac{D_2}{D_1}\right)^{\beta\gamma} \left(\frac{q_1}{q_2}\right)^{\gamma-1} = \frac{\tilde{N} \alpha q_2 + tD_1 D_2}{\tilde{N} \alpha q_1 + tD_1 D_2}.
\]

(14)

To satisfy (14), \( q_1 > q_2 \) must hold, because \( \beta > 0, \gamma > 1, \) and \( D_1 \geq D_2. \) Therefore, \( q_1^{OPT} > q_2^{OPT} \) is satisfied.

Further,

\[
\frac{l_1(q_1^{OPT})/D_1(q_1^{OPT})}{l_2(q_2^{OPT})/D_2(q_1^{OPT}, q_2^{OPT})} = \frac{q_1 (\tilde{N} \alpha q_2 + tD_1 D_2)}{q_2 (\tilde{N} \alpha q_1 + tD_1 D_2)} > \frac{q_1 \tilde{N} \alpha q_2}{q_2 \tilde{N} \alpha q_1} = 1,
\]

which derive \( l_1(q_1^{OPT}, q_2^{OPT})/D_1(q_1^{OPT}, q_2^{OPT}) > l_2(q_2^{OPT})/D_2(q_1^{OPT}, q_2^{OPT}). \)
References


Figure 1: The area of analysis considered in section 4.
\[ D_1(q_1, q_2) = D_2(q_1, q_2) \quad R_1(q_1, q_2) = 0 \]
\[ X(q_1, q_2) = 0 \]
\[ q_1 = q_2 \]

\[ X(q_1, q_2) = 1 \]
\[ R_2(q_1, q_2) = 0 \]

Figure 2: The response functions if \( N_1 > N_2 + 1 \).
Figure 3: The response functions if $N_1 \leq N_2 + 1$. 
Case 1: $\alpha_g=0.55, t=0.5$

Case 2: $\alpha_g=0.65, t=0.5$

Case 3: $\alpha_g=0.55, t=1.0$

Case 4: $\alpha_g=0.65, t=1.0$

Note

Upper bold line: $q_i$; lower bold line: $1/w_i$; upper thin solid line: $q_2$; lower thin solid line: $1/w_2$; dashed line: $S$.

$q_i$ and $1/w_i$ are logarithmically scaled on the left axis. $S$ is linearly scaled on the right axis.

The other parameters are same for any case: $N_i=2, N_2=1.5, v=1, a=1.25, w=1, 1/\gamma=0.9$, and $T_1=T_2=0$.

Figure 4: Changes in $q_i$, $1/w_i$, and $s_i$ with $p$. 