A Note on the Distinction Between Two Market Conditions in Japanese Health Care Market: Excess Demand vs SID †

Ken Aoki*

Abstract

A positive correlation between physician density and per capita utilization of medical services gave rise to the supplier-induced demand hypothesis, SID. The SID, which says physicians can create demand for their own interests, appeals to health care research because physicians are not only agents for determining the necessary care for their patients but also suppliers of medical services.

However, the positive correlation can also be explained as a result of excess demand. In contrast, SID occurs when a market is in excess supply. Two mutually exclusive hypotheses can explain the same phenomenon, but they differ in welfare implications. It is important to identify the market condition to adopt appropriate policies toward competition, cost containment and so forth. In spite of their importance, conclusions about market conditions are still controversial. There is great diffculty in distinguishing the market condition because two models with different market conditions have almost the same comparative statics results.

This article shows that the market condition can be distinguished by investigating the stability of physicians' regional distribution. Contrary to the conventional wisdom, instability of an equally distributed equilibrium means markets are not in supplier-induced demand but in excess demand.

Keywords：Market Condition, Supplier-induced Demand, Excess Demand
JEL classification：I11, L11, J44

I Introduction

A positive correlation between physician density and per capita utilization of medical services gave rise to the supplier-induced demand hypothesis, SID. The SID, which says physicians can create demand for their own interests, appeals to health care research[1)] because physicians are not only agents for determining the necessary care for their patients but also suppliers of medical services.

However, the positive correlation can also be explained as a result of excess demand. In contrast, SID occurs when a market is in excess supply. Two mutually exclusive hypotheses can explain the same phenomenon, but they differ in welfare implications. It is important to identify the market

† The author would like to thank Ted Frech for helpful comments and UCSB for its hospitality. Financial supports from the Pfizer Health Research Foundation and KAKENHI (2253024) are gratefully acknowledged.

* Department of Economics, Sophia University
E-mail: ken-aoki@sophia.ac.jp

condition to adopt appropriate policies toward competition, cost containment and so forth. In spite of their importance, conclusions about market conditions are still controversial. There is great diffculty in distinguishing the market condition because two models with different market conditions have almost the same comparative statics results.

This article shows that the market condition can be distinguished by focusing on the stability of the regional distribution of physicians. If we only investigate suppliers' behavior in one market, it is hard to identify the underlying market condition. However, it can be distinguished by observing multiple markets at once. The market of interest in our study has fee-for-service payments with controlled prices, as in the Japanese health care system. For this reason we do not adopt the standard equilibrium model as another alternative.

II The Model

1. Physician Behavior

A representative physician's utility depends on income Y, hours of work t, and the demand creation s.

$$u = U(Y, t, s), \tag{1}$$

where U is continuously differentiable and $U(Y, t, s)$ is strictly quasi-concave in Y and t. It is also assumed that $U_Y > 0$, $U_t < 0$, and $s \gtreqless 0 \Rightarrow U_s \lesseqgtr 0$. The maximum hours of work are normalized to unity, $0 \leq t \leq 1$. $s(> 0)$ has been the key ingredient to endogenously determine the optimal level of demand creation in SID models, which was pioneered by Evans (1974). In this model, $s < 0$ is also possible in order to deal with excess demand.

The nature of markets under fee-for-service payments with the controlled price P is captured by

$$Y = PM(s)N \quad \text{and} \quad t = g(M(s)N), \tag{2}$$

where $M(s)$ is the intensity of care per patient as a function of s with $M' > 0$, N is the number of patients, g is the inverse function of the production with $g' > 0$ and $g'' > 0$.

First, we derive the individual supply as a benchmark. Physicians maximize their utilities with respect to N and s.

$$\max_{N, s} U(PM(s)N, g(M(s)N), s) \tag{3}$$

The first-order conditions for (3) are

$$\begin{aligned} M(s)(U_Y P + U_t g') &= 0 \\ M'N(U_Y P + U_t g') + U_s &= 0 \end{aligned} \tag{4}$$

Suppose that the maximization problem has an interior solution. $N^*(P)$ and $s = 0$ denote the solution. Then the individual supply function is written as $M(0)N^*(P)$.

Next, we consider the situation in which physicians cannot freely choose N. This corresponds to the actual market because physicians can neither generate additional patients even when the patient visits are less than $N^*(P)$, nor neglect patients when $N > N^*(P)$, by law. In reality, N must be determined between demand and supply as a result of non-price rationing. But, for simplicity, we assume that N is exogenous. $N < N^*(P)$ and $N > N^*(P)$ for each physician correspond to the market in the case of excess supply and excess demand, respectively.

In this case, physicians maximize their utilities only with respect to s.

$$\max_{s} \; U(PM(s)N, g(M(s)N), s) \tag{5}$$

The first-order condition of this problem is

$$M'N(U_Y P + U_t g') + U_s = 0. \tag{6}$$

Suppose that this problem has a unique interior solution $s^*(N, P)$. Let the value function of problem (5) be $v(N, P)$. For the following discussion, we derive the properties of these functions.

$s^*(N, P)$ is positive and negative for $N < N^*(P)$ and $N > N^*(P)$, respectively, in the neighborhood of $N^*(P)$ because of the second-order condition for (3). Lemma 1 demonstrates that this property is globally held.

Lemma 1. *For any price level P, the optimal solution* $s^*(N, P)$ *is*:

(i) $s^*(N, P) \gtreqqless 0$ *for* $N \lesseqqgtr N^*(P)$
(ii) *continuous in N and P*

Proof. (i) We show that $s^*(N, P) < 0$ for any $N > N^*(P)$ at an arbitrary P. Suppose that there is $N' > N^*(P)$ such that $s^*(N', P) \geq 0$, which is abbreviated to $s' = s^*(N', P)$. Since the first-order condition (6) is satisfied at s' and $U_s \leq 0$ if $s' \geq 0$, the first term of (6) is greater than or equal to 0, every equality is held only if $s' = 0$.

The following discussion is in terms of t–Y space, see **Fig. 1**. $MRS(t, Y; s)$ denotes the marginal rate of substitution of working hours for income dY/dt at (t, Y) with given s, and $MI(t)$ represents the marginal income at t with regard to the income function $Y = Pg^{-1}(t)$. By definition, these two have the following relationship at any point (t, Y) on the income function:

$$MRS(t, Y; s) \gtreqqless MI(t) \Leftrightarrow U_Y P + U_t g' \gtreqqless 0. \tag{7}$$

Note that the right hand side of the above equation and the first term of FOC (6) have the same sign. Since $U_s \leq 0$ for $s' \geq 0$, $MRS(t', Y'; s') \leq MI(t')$ is held, where $t' = g(M(s')N')$ and $Y' = PM(s')N'$. Hence, the slope of the indifference curve at (t', Y') is less than or equal to the marginal income, every equality is satisfied only if $s' = 0$.

Next, we introduce an auxiliary point (t_0, Y_0) such as $t_0 = g(M(0)N')$ and $Y_0 = PM(0)N'$. The indifference curve through (t_0, Y_0) with $s = 0$ is steeper than the income curve, $MRS(t_0, Y_0; 0) > MI(t_0)$, because $U(t, Y, 0)$ is decreasing in N for $N > N^*(P)$.

Since the slope of the indifference curve $U(t, Y, 0) = u_0 (= U(t_0, Y_0, 0))$ is greater than the slope of $U(t, Y, s') = u' (= U(t', Y', s'))$ in the interval $t_0 \leq t \leq t'$, and $Y_0 < Y'$, two indifference curves cross in the interval. Let the intersection be (t'', Y''). First, when $s' = 0$, it is immediately contradictory because there are two indifference curves with the same $s = 0$ intersect at (t_0, Y_0). Next, when $s' > 0$, $U(t'', Y'', 0) > U(t'', Y'', s')$ is satisfied because $U_s < 0$ if $s > 0$. (t_0, Y_0) is feasible when $N = N'$ and its utility level u_0 is greater than u'. This means that s' cannot be the maximizer at $N = N'$. Other cases can also be proven in a similar manner.

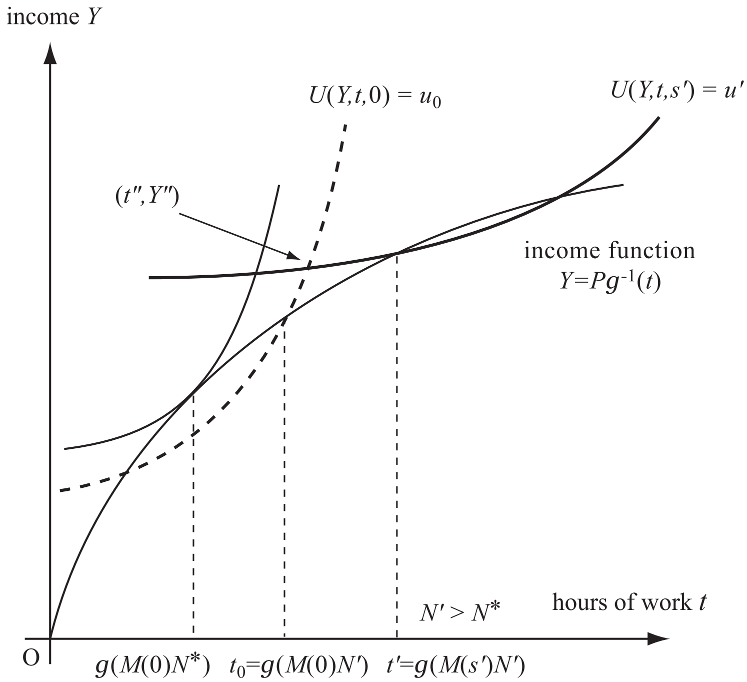

Figure 1: Indifference curves in t–Y space if s' were greater than 0 for $N' > N^*$.

(ii) Let $w(s, N; P) = U(PM(s)N, g(M(s)N), s)$. The problem (5) is the same as a problem which maximizes $w(s, N; P)$ on $\Gamma(N; P) = \{s \mid 0 \leq g(M(s)N) \leq 1\}$. Since $w(s, N; P)$ is continuous in s, N and P, and $\Gamma(N; P)$ is compact-valued correspondence and continuous in N and P, the continuity of the function $s^*(N, P)$ follows from Berge's maximum theorem. □

Proposition 1 says that the value function is hump-shaped in N. This property is essential to determine the stability of a regional distribution.

Proposition 1. *For any given price level P, the value function $v(N, P)$ is:*
(a) *maximized at $N^*(P)$*
(b) *strictly increasing in N for $N < N^*(P)$ and strictly decreasing in N for $N > N^*(P)$*
(c) *continuous in N and P*

Proof. (a) is an immediate consequence of $N^*(P)$ being the maximizer of the unconstrained maximization problem (3).

(b) We show that for any $N_1 > N^*(P)$ there exists an interval $(N_1, N_2]$ such that $v(N_1, P) > v(N', P)$ for all $N' \in (N_1, N_2]$. By Lemma 1, $M(s^*(N, P))N$ is continuous in N and $M(s^*(N_1, P))N_1 < M(0)N_1$. Then there exists an interval $(N_1, N_2]$ such that $M(s^*(N, P))N \leq M(0)N_1$ for all $N \in (N_1, N_2]$. Let s_1 and s' be the optimal solutions for N_1 and N', respectively. For all $N' \in (N_1, N_2]$,

$$\begin{aligned} v(N_1, P) &= U(PM(s_1)N_1, g(M(s_1)N_1), s_1) \\ &\geq U(PM(s'')N_1, g(M(s'')N_1), s'') \\ &> U(PM(s')N', g(M(s')N'), s') = v(N', P), \end{aligned} \tag{8}$$

where s'' satisfies $M(s'')N_1 = M(s')N'$ $(\leq M(0)N_1)$. The first inequality follows because s_1 is the optimal solution for N_1. The second inequality follows because $0 \geq s'' > s'$ under the same Y and t, and $U_s > 0$ when $s < 0$. Another case is proven in the same way. Note also that, when $v(N, P)$ is differentiable in N, (b) is easily proven by the envelope theorem.

(c) is also a consequence of Berge's maximum theorem. □

2. Location Choice

Suppose that there are two isolated markets, A and B, which have identical demand for health care services. Physicians can freely choose the location of their practices.

The equilibrium of the regional distribution is characterized by physicians in both markets A and B having equal utilities.

III Equilibrium Distribution and Its Stability

One obvious equilibrium distribution is that a half the physicians locate in market A and the other half in market B. This distribution pattern is of special significance because this is the ultimate goal for policy makers attempting to dissolve regional inequalities of medical resources. We therefore call this the equally distributed equilibrium.

The equally distributed equilibrium is a feasible equilibrium both when two markets are in excess demand and when two markets are in excess supply. However, the stability of this equilibrium depends on the underlying market condition.

Proposition 2. *Suppose that two markets, A and B, have identical demand under fixed price P and physicians can freely choose their practice locations. For a small change, the equally distributed equilibrium is stable if both markets are in excess supply. In contrast, the equally distributed equilibrium is unstable if both markets are in excess demand.*

Proof. This is an immediate consequence of Proposition 1. If both markets are in excess supply, a physician's utility is increasing in N since the value function is increasing in N for $N < N^*(P)$. Hence, a physician in the market with fewer physicians, which means larger N per physician, has higher utility. Thus, moving from the lower utility market to the higher one cancels out the change. Meanwhile, if both markets are in excess demand, a physician's utility is lower in the market with fewer physicians. Moving from a lower utility market to the higher would make the initial change worse. □

Strictly speaking, the reverse is not true, because the equally distributed equilibrium is either stable or unstable if both markets are in equilibrium, depending on the shape of value function. However, neither market is truly in equilibrium without price adjustment. Therefore, except for this rare case, the reverse of Proposition 2 is also true. We can distinguish the market condition by investigating the stability of the regional distribution.

IV Conclusion

It is difficult to identify which condition the market is in when we focus on a single market. However, we show that the market condition can be distinguished by investigating multiple markets at once; that is, the stability of regional distribution. Surprisingly, instability of the equally distributed equilibrium means markets are not in SID but in excess demand.

Our result is totally different from the conventional wisdom, which holds that regional maldistribution is evidence of SID. In our model, it is excess demand rather than SID that causes maldistribution. Some previous studies found that regional inequality of medical resources became even worse due to SID. However, these findings might be evidence of prevailing excess demand.

Finally, while markets are separated geographically in this article, other types of separation would be also possible, e.g., physicians' specialties.

Notes

1) Numerous numbers of empirical researches have been attempted to test the SID hypothesis in all over the world. For example, Kishida (2001) tests the SID in Japanese outpatient market. McGuire (2000) provides a good overview of these researches.

References

[1] Evans, R. G. "Supplier–Induced Demand: Some Empirical Evidence and Implications." in Perlman, M. ed., *The Economics of Health and Medical Care*, Macmillan, 1974.

[2] Kishida, K. "Supplier Induced Demand Hypothesis versus Decline of Access Cost Hypothesis: Analysis with Panel Data of Second Medical Areas and Cities. (in Japanese)" *The Quarterly of Social Security Research*, 37 (3): 246–258, 2001.

[3] McGuire, T. G. "Physician Agency." in Culyer, A. J. and Newhouse, J. P. eds., *Handbook of Health Economics* Vol. 1A, Elsevier, 2000.

[4] Newhouse, J. P., Williams, A. P., Bennett, B. W., and Schwartz, W. B., "Does the Geographic Distribution of Physicians Reflect Market Failure?" *Bell Journal of Economics*, 13 (2): 493–505, 1982.