Provider-Client Interactions and Quantity of Health Care Use[1]

Hsien-Ming Lien,[2] Ching-to Albert Ma,[3] and Thomas G. McGuire[4]

September 2002

[1]This research received support from NIDA-P50-DA10233, and AHRQ-PO1-HS10803. We are grateful to the Maine Office of Substance Abuse for cooperation in providing the data for the study, and to Victor Aguirregabiria, Chungrong Ai, Richard Frank, Kevin Lang, Mingshan Lu, and Joe Newhouse for helpful comments.

[2](Corresponding author) Department of Public Finance, National Cheng-Chi University, 64 Chih-nan Rd., Sec. 2, Wensha, Taipei 11623, Taiwan, hmlien@nccu.edu.tw, Tel: 886-2-29393091-51658, Fax: 886-2-29390074.

[3]Department of Economics, Boston University, 270 Bay State Road, Boston, MA 02215, ma@bu.edu; Department of Economics, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, albertma@ust.hk

[4]Department of Health Care Policy, Harvard Medical School, 180 Longwood Avenue, Boston, MA 02115, mcguire@hcp.med.harvard.edu

Abstract

This paper considers three types of provider-client interactions that influence quantity of health care use: rationing, effort, and persuasion. By rationing, we refer to a quantity limit set by a provider; effort, the productive inputs supplied by a provider to increase a client's demand; persuasion, the unproductive inputs used by a provider to induce a client's demand. We construct a theoretical model incorporating all three mechanisms as special cases. When the general model is specialized into one of three mechanisms, a set of empirical implications emerges. We test for the presence of each mechanism using data of patients receiving outpatient treatment for alcohol abuse in the Maine Addiction Treatment System. We find evidence for rationing and persuasion, but not effort.

Key words: Induced Demand, Physician Behavior.

JEL: I11, L15

1 Introduction

The theoretical and empirical literatures in health economics take different approaches to the issue of provider induced demand. In the empirical literature, the term "provider induced demand" (PID) is used broadly to account for a range of empirical findings. Many studies have shown that when common demand-side variables (such as demand price, income and clinical needs) are controlled for, consumer demands continue to be affected by physician or provider variables. These supply-side variables include the number of competitors in a market (Cromwell and Mitchell, 1986), supply price (Rice, 1983; Yip, 1998), incentives for self-referral because of ownership of complementary inputs (Hillman et al., 1992), physician attitudes toward earnings (Rizzo and Blumenthal, 1994), demand shocks due to demographic changes (Gruber and Owings, 1996), incentives due to malpractice liability (Kessler and McClellan, 1996), and partnership incentives (Gaynor and Gertler, 1995). Empirical papers generally attribute these effects to PID.

The theoretical literature uses the term PID more narrowly. Referring to Fuchs (1978), Pauly (1980), Eisenberg (1986), Culyer (1989) and Williams (1998), we say that PID exists when the physician influences a patient's demand for care not necessarily in the patient's best interest. PID may be regarded as a form of manipulation: the patient is persuaded to demand health services. Besides persuasion, other ways by which a physician influences utilization have also been identified. A provider can set quantity directly; this is possible because health care services are nonretradable (Farley, 1986; Gaynor, 1994). When a provider's quantity setting power is exercised, deliberate influence on the patient's demand is unnecessary; the patient simply accepts or rejects the provider's decision. Third, the provider can undertake an action ("effort" or "quality"), noncontractible but observable to the patient. The effort influences the patient's valuation of services, therefore affecting demand (Ma and McGuire, 1997). Unlike the inducement behind PID, the physician's action or effort is productive.

The three mechanisms are conceptually distinct (See McGuire (2000) for a review). Under PID, the key word is *induce*; the patient's use of health services may be unrelated to his true demand. In the quantity-setting mechanism, when the physician unilaterally picks quantity, the patient may only accept or refuse. In the action-effort mechanism, a patient reacts rationally against the physician's effort. These mechanisms have different implications for a normative view of how health care markets function. It is important, therefore, for both academic research and policy, to apply carefully the ideas of theoretical papers to guide empirical work. In this paper, we take a step in this direction by analyzing a health care setting in which providers may influence quantities. Specifically, we use the data to test for the presence of the three mechanisms of influence.

Section 2 reviews the three mechanisms of provider influence on patient decision making. The theoretical model presented there incorporates all three mechanisms as special cases. When the general model is specialized into one of the three mechanisms, a specific set of predictions emerges. These predictions allow us to test for the presence of each mechanism. The statistical method for identifying the relative roles of the three mechanisms is presented in Section 3. Our data set, described in Section 4, is from public sector drug and alcohol treatment facilities. It contains more information than common health care utilization studies. In particular we have information on whose decision (the client's or the provider's) it was to terminate a course of treatment, as well as on treatment outcomes. The first piece of information helps to distinguish the quantity setting mechanism from effort and persuasion; the second separates productive effort from unproductive persuasion. Our empirical results are presented in Section 5. We find evidence for two of three mechanisms: direct quantity setting and persuasion. The mix of these three mechanisms is likely to differ across health care settings, but the perspective and the empirical methods proposed here can be used elsewhere. We discuss these prospects in Section 6.

2 Mechanisms of Provider Influence on Health Care Use

The health economics literature places significant emphasis on the modelling of physician-patient interactions. Such interactions determine the quantities of care supplied to patients. Table 1, adapted from McGuire (2000), summarizes the three mechanisms in the literature. Common to almost all papers on this topic is a monopolistically competitive market structure, but different informational and contractibility assumptions have been made. Perhaps the most straightforward account of quantity determination is persuasion, associated with the writings of Dranove (1988), Evans (1974), Fuchs (1978), Rice (1983) and many others. (See the column under "Persuasion" in Table 1.) Persuasion is simply regarded as a form of inducement that shifts the entire demand curve. Inducement itself has no effect on the health outcome, though any quantity demanded following from inducement may affect outcome. To complete the persuasion theory, some cost to the provider of inducement must be specified. Dranove (1988) contains a model in which the profit-maximizing amount of inducement is limited by the patient's suspicion of the provider's aggressiveness; other approaches include the resource cost of the activity (Stano, 1987), and increasing professional discomfort (Evans, 1974; McGuire and Pauly, 1991).

The second mechanism regards the health care provider as a quantity setter; see the column under "Nonretradability" in Table 1. Health care providers supply a nonretradable service (Farley, 1986; Gaynor, 1994). A monopolistic competitor selling a nonretradable service sets a quantity to maximize profit. Recognition of physician quantity setting power stemming from nonretradability accords well with most patients' experiences with physicians who, quite simply, "tell them what to do." Of course, the theory also must recognize that consumers need not always comply with physicians' instructions. Generally, one can interpret the quantity setting model as one in which quantity restrictions are placed by physicians, and patients respond to these restrictions. Empirically, this theory says that observed quantity results from the physician's quantity restriction and the patient's response. Also, like the persuasion mechanism, the quantity setting model does not involve any other input into health care production.

The third mechanism posits that a physician supplies a noncontractible input in health production; this is additional to other measurable quantities of treatments such as days, visits, or tests. This noncontractible input can be regarded as quality, such as the time and effort with which the physician tailors the treatment to the patient's needs, or the physician's care in delivering services. For our purpose, the important aspect of quality is that it affects health outcomes as well as the patient's behavior. Knowing a physician's quality, a patient will respond by demanding more or less care, depending on the productive relationship between the two inputs. Ma and McGuire (1997) develop such a theory. That paper models patient-physician interaction symmetrically—one input (quality) is chosen by the physician, another (quantity) by the patient. Notably, effort, a productive input, affects health outcome.

In this section, we develop a general model which yields each of the above three mechanisms as special cases. Each special case generates an empirical hypothesis, which we will test with our data. Our model generalizes Ma and McGuire (1997) in several directions. First, the possibility of persuasion is added. Second, we explicitly model provider quantity setting by a limit on utilization. Third, the provider does not know patients' preferences with certainty.

We begin with the patient's utility function: $\theta U(q, e) - p_d q$, where θU is the patient's benefit of receiving treatment quantity q when either the physician's amount of persuasion or quality is e.

(So for the moment, e represents either persuasion or quality.) The parameter θ can be regarded as the patient's "type," its value being known only to the patient. The physician only knows that θ follows the distribution function F with density f. The fee that the patient pays per unit of q is p_d.

Given e, a patient responds by choosing the optimal quantity. A patient with parameter θ picks q to solve:

$$\max_q \theta U(q,e) - p_d q. \tag{1}$$

The optimal quantity satisfies the first-order condition:

$$\theta \frac{\partial U(q,e)}{\partial q} = p_d. \tag{2}$$

From (2), a patient's quantity choice is affected by θ and e. Let the optimal quantity be $q(\theta, e)$. The comparative statics with respect to q can be obtained by the total differentiation[1] of (2):

$$U_q(q,e)d\theta + \theta[U_{qq}(q,e)dq + U_{qe}(q,e)de] = dp_d.$$

Setting $dp_d = 0$, we obtain:

$$\frac{\partial q}{\partial \theta} = -\frac{U_q(q,e)}{\theta U_{qq}(q,e)} > 0, \qquad \frac{\partial q}{\partial e} = -\frac{U_{qe}(q,e)}{U_{qq}(q,e)}.$$

If $U_{qe}(q,e) > 0$ (input complementarity), then $\partial q/\partial e > 0$. This implies that a patient selects a higher treatment quantity when the physician uses the higher effort or persuasion level.

The physician is assumed to maximize her expected profit, the expectation being taken over the patient's preference parameter θ. The provider selects effort or persuasion to affect the patient's demand, along with a quantity ceiling that restricts the maximum quantity of care a patient can receive. The extensive form of the game is described as follows. In the first stage, knowing only the distribution function $F(\theta)$, the physician picks an effort or persuasion level e as well as a quantity ceiling Q. In the second stage, the patient observes Q, θ, and e, and picks quantity q, where $q \leq Q$.

[1]Subscripts of functions denote their partial derivatives.

Consider the second stage when the patient must choose a quantity. Given e, suppose the patient's choice is given by the maximization of (1). If that yields a solution higher than Q, then Q will be the actual quantity. Otherwise, it will be given by (2). From (2), given Q and e, there exists a Θ such that $q(\theta, e) \leq Q$ if and only if $\theta \leq \Theta$, or $q(\Theta, e) = Q$. Given e, a choice of a ceiling, Q, is equivalent to a choice of Θ.

Let p_s be the price that a physician receives per unit of quantity, and $C(q, e)$ the physician's cost function when q units of quantity and e units of effort or persuasion are used. The physician's expected profit is

$$\Pi(q, e, \Theta) \equiv \int_0^{\Theta} [p_s q(\theta, e) - C(q(\theta, e), e)] dF(\theta) + [1 - F(\Theta)][p_s q(\Theta, e) - C(q(\Theta, e), e)]. \tag{3}$$

For those patients with θ below Θ, the quantity ceiling does not bind, and the profit is given by the integrand of the first term of (3); for θ above Θ, the quantity ceiling binds, and the last term in (3) is the profit. The physician picks e and Θ to maximize (3) subject to (2) for $\theta \leq \Theta$. The first-order conditions are

$$\frac{\partial \Pi}{\partial \Theta} = (p_s - C_q(Q, e))[1 - F(\Theta)] \frac{\partial q(\Theta, e)}{\partial \theta} = 0 \tag{4}$$

$$\begin{aligned} \frac{\partial \Pi}{\partial e} &= \int_0^{\Theta} (p_s q_e - C_e - C_q q_e) dF(\theta) + [1 - F(\Theta)][p_s q_e(\Theta, e) - C_e - C_q q_e(\Theta, e)] \\ &= \int_0^{\Theta} (p_s q_e - C_e - C_q q_e) dF(\theta) - [1 - F(\Theta)] C_e = 0, \end{aligned} \tag{5}$$

where the second equality of (5) follows from (4). The first-order condition (4) shows why a physician would set a quantity limit. When the value of θ is high, the patient demands a large quantity of care. The marginal cost function, however, is increasing and convex. For some given price p_s the physician will make a loss for those patients who have high values of θ. Installing a quantity limit avoids these losses. In fact (4) simply says the marginal cost at the quantity limit Q is exactly equal to the physician's price p_s.

In the general model, the equilibrium quantity is a result of the patient's choice or the physician's quantity limit. In each case, the equilibrium quantity is a function of the patient's and physician's prices, p_d and p_s, as are the equilibrium effort e and the quantity limit Q. Two special cases of the general model will correspond respectively to pure quantity setting and pure effort-persuasion theories of health care utilization.

For pure quantity setting, the utility function U is assumed to be independent of e so that the demand function q only depends on θ and price p_d. Then the equilibrium quantity limit will be $p_s = C_q(Q)$. The equilibrium quantity limit Q will only depend on p_s, not on p_d. Also, when the quantity limit does not bind, for those consumers with low values of θ, the equilibrium quantity will be given by $\theta U_q(q) = p_d$. We summarize the dichotomy by the following result.

Proposition 1 (Absent Effort or Persuasion) *In the pure quantity setting model, (i) when the quantity of treatment is less than the ceiling set by the physician, quantity is a function of the patient's price p_d but not the physician's price p_s; (ii) the physician's quantity ceiling is a function of p_s but not p_d.*

Next, we consider pure effort or persuasion. Here, we set the quantity ceiling Q to an arbitrarily high level: the physician must accept any quantity demanded by the patient. In this case, the second term in (3) vanishes and the equilibrium effort or persuasion will be given by the following first-order condition:

$$\int_0^\infty (p_s q_e - C_e - C_q q_e) dF(\theta) = 0. \tag{6}$$

The equilibrium effort depends on p_s directly, and indirectly on p_d through the demand function (2). In turn, the equilibrium quantity depends on both p_s and p_d.

Proposition 2 (Absent Quantity Ceiling) *In the pure effort-persuasion model, quantity is a function of both patient's and physician's prices, p_s and p_d.*

The above propositions will be used for empirical tests of the pure quantity and pure effort-persuasion models. We now propose a method to distinguish between quality and persuasion. To accommodate the presence of both (unproductive) persuasion and (productive) effort, we expand the patient's subjective benefit function to $U(H(q, \hat{e}), \tilde{e})$, where H denotes a patient's health, and is a function of quantity of health care q and productive effort $\hat{e}$. Subjective benefit also depends on a physician action, $\tilde{e}$, the persuasive activity that affects patient's demand but not health. We observe measures of health in our data, allowing us to see if the physician action influencing demand works through or independently of a health effect. We test the special case of our model in which H is a function only of q, not $\hat{e}$. This corresponds to a pure persuasion model. We now state:

Proposition 3 (Effort versus Persuasion) *In a pure persuasion model, health outcomes are unaffected by inducement, which depends on patient's and physician's prices,* p_s *and* p_d*. After effects from quantities are controlled for, health outcomes do not depend on prices in the pure persuasion model.*

3 Estimation Strategy

In this section we use Propositions 1-3 for empirical tests. Our data contain several special features. First, as in many health care data sets, prices paid by clients are not necessarily equal to those received by providers, due to insurance and reimbursement practices. As a result, we can separately identify demand and supply prices. Second, our data contain information about whose decision it was to end treatment for alcohol abuse. The client may "quit," or the provider may stop the episode by calling it "complete." Using this information we say that the provider's quantity limit is binding if and only if an episode is complete. Finally, our data include information on health outcomes. This information enables us to test whether the provider's action is persuasion or effort.

3.1 The General Specification

We begin with an empirical specification of the general model. Let q denote the quantity demanded by a client; Q, the quantity limit set by the provider. Let q^* be the observed quantity. The reduced form model can be described by a system of equations:

$$q = X_d\beta_d + \delta_d p_d + \gamma_d p_s + \mu_d, \tag{7}$$

$$Q = X_s\beta_s + \delta_s p_d + \gamma_s p_s + \mu_s, \tag{8}$$

$$q^* = j^* q + (1 - j^*)Q, \qquad j^* = I(q \leq Q). \tag{9}$$

Equations (7) and (8) describe a client's demand and a provider's limit respectively. Both a client's demand (q) and the provider's limit (Q) are functions of the supply price and the demand price in the general model. The vector X_d represents factors influencing a client's demand, such as income, health and demographic variables, and μ_d is an error term denoting unobserved client characteristics. Likewise, the vector X_s represents factors affecting a provider's limit and μ_s is an error term capturing unmeasured provider features. Because the provider may set the limit based on a patient's observed characteristics, X_s may contain elements of X_d. In addition, μ_d and μ_s may be correlated if the provider's limit is chosen according to some unobserved demand characteristics. Finally, equation (9) says q^* can be either q or Q, depending on the observed indicator j^*, where j^* equals to one if the event in the indicator function I is true.

Equations such as (7)-(9) are known as a disequilibrium model in the literature; see Maddala (1983) for a review. Nevertheless, the model presents a static perspective of treatment: q and Q are both chosen independently at the beginning of a course of treatment. While this characterization is common in many studies, it may be inappropriate for our purposes since q and Q could affect one another. A client may choose to visit the doctor more if his provider suggests a higher limit. On the contrary, a lower provider's limit may induce the client to reduce his demand. To adapt the model for our setting, we allow q and Q to be related: the client chooses the doctor's limit Q when it is not much bigger than q. Specifically, we replace equation (9) with

$$q^* = j^* q + (1 - j^*)Q, \qquad j^* = I(q \leq Q - \alpha + \varepsilon). \tag{10}$$

The parameter α is positive, and ϵ a random error with zero mean. Equation (10) says that on average whenever the client's desired demand is within α units of the provider's limit, he revises his demand upward and accepts the provider's limit.

The system of equations ((7), (8) and (10)) results in a complex likelihood function. For tractability, we assume that μ_d, μ_s, and ε are normally distributed with zero means and standard deviations σ_d, σ_s and σ_ε respectively. We assume further that μ_d and μ_s have correlation ρ and zero correlation with ε. The likelihood function for a client with the observed health care use q^* and the completion indicator j^*can be simplified as

$$\begin{aligned} \ln f(q^*,\ j^*) \quad &= \quad j^* \times \left[\ln \phi(\frac{D}{\sigma_d}) - \ln \sigma_d + \ln \Phi \left(\frac{\frac{\rho\sigma_d}{\sigma_s} D - \alpha - S}{\sqrt{\sigma_s^2(1-\rho^2) + \sigma_\varepsilon^2}} \right) \right] \\ &\quad + (1 - j^*) \times \left[\ln \phi(\frac{S}{\sigma_s}) - \ln \sigma_s + \ln \Phi \left(\frac{\frac{\rho\sigma_s}{\sigma_d} S + \alpha - D}{\sqrt{\sigma_d^2(1-\rho^2) + \sigma_\varepsilon^2}} \right) \right], \end{aligned} \tag{11}$$

where $D \equiv q^* - X_d\beta_d - \delta_d p_d - \gamma_d p_s$ and $S \equiv q^* - X_s\beta_s - \delta_s p_d - \gamma_s p_s$. ϕ and Φ are the density and cumulative density functions of the standard normal distribution respectively.[2] The exact variables used in the estimation are specified below.

3.2 Do Providers Use Effort or Persuasion?

We use Proposition 1 to test for the presence of effort or persuasion (for now we do not distinguish between them). Under a quantity ceiling restriction, the system of equations reduces to

$$q = \beta_d X_d + \delta_d p_d + \mu_d, \quad (12)$$
$$Q = \beta_s X_s + \gamma_s p_s + \mu_s, \quad (13)$$
$$q^* = j^* q + (1 - j^*)Q, \qquad j^* = I(q \leq Q - \alpha + \varepsilon). \quad (14)$$

Recall that p_s and p_d are included in equations (7) and (8) because of the provider's endogenous effort or persuasion. Without effort or persuasion, q does not depend on p_s; nor does Q on p_d. The joint restriction $\delta_s = \gamma_d = 0$ is therefore used for testing the presence of effort or persuasion.

3.3 Do Providers Set a Quantity Ceiling?

We use Proposition 2 to test for the presence of a quantity limit. Under the pure effort-persuasion setting, a provider sets her limit above the client's demand. As a result, the observed quantity is always a client's demand ($q^* = q$ or $j = 1$) and can be fully captured by the regressors in (7). Formally, the system of equations reduces to

$$q^* = X_d \beta_d + \delta_d p_d + \gamma_d p_s + \mu_d. \quad (15)$$

Given that Proposition 2 restricts Q to be larger than q, the quantity ceiling restriction can be tested by comparing the likelihood values of the general model and one with an arbitrarily large Q. Nevertheless, such a method suffers from two potential flaws. First, it is not clear which Q should be used for the comparison; the statistical power of the test depends on the chosen value of

[2]See Appendix for details.

Q. More seriously, the method relies directly on the information of j—the proportion of completed episodes is included in the likelihood function. In case that the indicator is incorrectly interpreted or measured with errors, directly employing the information of j^* may be problematic.

To be conservative, we consider another test that does not rely directly on j. Let $\hat{q}_g$ be the predicted quantity from the general model. Rewrite (15) as

$$q^* = X_d \beta_d + \delta_d p_d + \gamma_d p_s + \tau \hat{q}_g + \mu_d. \tag{16}$$

Here τ represents the explanatory power of the general model. The coefficient τ is expected to be one under the general model, and zero under the pure effort model since regressors in (7) fully explain q^*. The restriction $\tau = 0$ is thus used for testing the presence of quantity ceiling.

3.4 Do Providers Use Effort?

We use Proposition 3 to test for the presence of effort. To derive a testable hypothesis, we specify the health production function. Let H_1 and H_0 denote a client's health at admission and discharge respectively. A client's health outcome H_1 is a linear combination of a client's health at admission (H_0), personal characteristics (X_d), health care use (q^*), and effort or persuasion (e), namely,[3]

$$H_1 = \alpha_0 H_0 + \alpha_{1d} X_d + \alpha_2 q^* + \alpha_3 e + \eta_h, \tag{17}$$

where η_h denotes a random error. The above specification allows e to affect H_1 directly or indirectly through q^*. Proposition 3 says that conditional on q^*, H_1 is unaffected by persuasion.

Neither effort or persuasion is observed. To complete the estimation, we use the supply price as a proxy for e. Under pure persuasion, Proposition 3 implies that $\alpha_3 = 0$.

---

[3]Our specification of treatment production function is similar to Lu and McGuire (2002), except that they exclude effort as an input in the production function.

4 The Data

Our data source is the Maine Addiction Treatment System (MATS). The data describe people receiving substance abuse treatment at any publicly funded facility in Maine. About one percent of Maine's 1.2 millions residents have contact with a public substance abuse agency in a year (Lu and McGuire, 2002). A large portion of the service costs of substance abuse clients are funded by the state regulatory agency,[4] the Office of Substance Abuse (OSA).[5] For evaluation purposes, each agency receiving OSA funds is required to report information for every treated episode. In total, MATS collected records of over thirty thousand treatment episodes from 82 agencies, spanning the time between October, 1990 to September, 1995.[6]

MATS collects treatment information through standardized forms completed during personal interviews at admission and discharge.[7] The admission form covers extensive questions about a client's background, such as demographics, income, employment and legal status, as well as an initial assessment of a client's health status. The discharge form contains comprehensive information on services provided in an episode, including providers, expected payment sources, delivered services (types of treatment, number of visits, etc), along with an assessment of a client's health condition at that time.[8]

More importantly for our research, MATS records which party decided to terminate the treat-

[4]The total budgeted expediture of the Office of Substance Abuse was $10,085,716, or about $8 per capita in 1995. Around 60% of total budget is used to finance treatment costs for indigent clients. Source: Office of Substance Abuse 1995 Data Book; State of Maine Budget Document, 1994-95.

[5]The Department of Human Service was the responsible agency prior to the creation of OSA. OSA was created in July 1990 as a branch of the State's Executive Department. After July 1, 1996, OSA was transferred to the Department of Mental Health, Mental Retardation, and Substance Abuse Service. OSA was responsible for allocating state and federal funds for sbustance abuse, and for contracting with agencies that provide substance abuse services.

[6]Our data actually contain MATS episodes over seven years, from Sepemter 1989 to October 1996 (FY90-96). However, the data for FY90 and FY96 are truncated, resulting a much smaller number of episodes compared to other sample years. In addition, many episodes in these two years are incomplete due to truncation. The analysis therefore uses only data of treatment episodes between FY91 and FY95.

[7]OSA requires that a client to be interviewed either by his counselor, or by an experienced staff member in the agency.

[8]Because of its abundant information, MATS is used by Lu and McGuire (2002), Machado (2001), Ackerberg, Machado and Riordan (2001) for the study of treatment productivity, and by Shen (1999) for the study of strategic risk selection by providers.

ment episode in the discharge form. A treatment episode is recorded as either complete or incomplete. For our empirical implementation, we regard quantity as the provider's limit if and only if a provider says that a client has completed a treatment episode. Consider first an incomplete treatment episode. Here, the quantity must have been the result of a conscious decision by the patient to terminate a treatment; we regard that as the client's demand. Now for a complete treatment episode, there seem to be two possibilities. Both the client and the provider would like to stop treatment, or it is the provider's decision. In the latter case, the classification of the quantity as the provider's limit is correct; in the former case, the classification is not so clear. Nevertheless, a physician often can successfully convince a patient that further treatment is unwarranted if indeed the physician does not want to continue. On balance, we think that it is more natural to regard the quantity as the provider's limit when a treatment episode is complete. In any case, we do not use a strict classification: see our discussion in equation (10).

Our analysis focuses on outpatient treatment for alcohol abuse, the most prevalent form of treatment in MATS.[9] We further restrict the sample in a number of ways. First, due to institutional reasons, some clients in the data could not choose quantity.[10] We ignore these clients because their treatment behaviors are different from our model. Second, our analysis relies on the client's payment source to construct proxies for prices. Clients with no recorded payment sources in the data are therefore dropped. Finally, in order to control for agency fixed effects in the estimation, we only keep those agencies having at least 20 MATS episodes over the sample years. The remaining sample consists of 7615 episodes. Because MATS is collected on an episode basis, there are only 6778 unique clients. In what follows, to simplify discussion, we disregard the distinction between episodes and clients.

Table 2 reports the variables included in X_d, starting with variables describing characteristics of clients. 69.9% of clients are male and 78.3% are single, although the majority of single clients (67.8%) were married before. The average client is around thirty years old, with a high school

[9]Of all alcohol abuse episodes in MATS, a half of them (54.5%) receive outpatient treatment for alcohol abuse.

[10]Explicitly, we exclude patients who are (1) enrolled in Driver's Education and Evaluation Program (DEEP) (2) currently in jail (3) referred to other agencies or programs (e.g. for inpatient service) (4) deceased or moved away. DEEP are programs designed to prevent future offenses caused by drivers with problems of substance abuse. Enrollees in these programs may be required to attend a certain number of visits to recover their licenses.

diploma and a monthly household income of about \$900.[11] Among the clients, 28.6% have full-time jobs, 9.8% work part time and the remaining 61.6% do not work at all, partly explaining the heterogeneity in the clients' monthly household incomes. Approximately 40% of clients have some legal involvement at admission, either being on probation or parole or waiting for a trial. Finally, 13.7% of clients have recognized psychiatric comorbidity at admission.

The second set of variables in X_d are related to a client's initial health. We employ two kinds of measures of a client's health at admission. First, we construct four dummies from the client's reported frequency of alcohol usage at admission, and denote the client as "light" if the client has been abstained from alcohol for over a month, "moderate" if his usage is less than once per week, "intermediate" if less than once per day, and "heavy" if the usage is even greater. Second, based on MATS's terminology, we construct dummies from the clinician's assessment of the severity of the client's health condition, indicating whether the client is a "casual," "life-style involved," "life-style dependent," or "dysfunctional" user, as assessed by the clinician. The summary statistics of these health variables are displayed in Table 2. In our sample, 18.8% of clients are moderate, 24.1% are intermediate and 9.4% are heavy users. By contrast, 23.5% are life-style involved, 47.5% are life-style dependent, and 22.2% are dysfunctional.

Table 3 presents the distribution of clients' payers. According to a client's primary expected payment source, we construct four different payer dummies: "OSA," "Medicaid," "Self-Paid," and "Third Party," and present its distribution in Table 3.[12] About 22% of clients are primarily supported by OSA funds (OSA); 29.4% are covered either by Medicaid or Medicare programs (Medicaid); 30.2% have private insurances (Third Party), and 18.5% pay treatment costs out of their own resources (Self-Paid). Many clients who are classified as Self-Paid are also partially supported by OSA.

The exact price paid by clients and recorded by providers are not available in MATS.[13] Nevertheless, the price a client pays and the fee a provider gets for treatment are mainly determined

[11]This is the average monthly income when income is known and stated by the patient and is less than \$9999.

[12]MATS reports a client's primary, secondary and tertiary expected source of treatment payment; the primary source of payment is applied to identify a client's payer group.

[13]Although MATS contains payment records of each provider at every fiscal year, they are not detailed enough to identify the exact payments (paid by the client and received by the provider) of a client. Thus, we are unable to derive the exact demand and supply prices from these payment records.

by the client's payer status. We thus use a client's payer status as a proxy for demand and supply prices. The ranking of supply and demand prices is also displayed in Table 3. Medicaid and OSA clients in general face a lower (zero) demand price. The provider receives a more generous reimbursement when clients are covered by Medicaid or private insurances. Because we only know relative prices, the demand or supply price is coded as 1 when it is "high" and 0 when it is "low".

Using payers as proxies of prices may introduce a selection problem because clients may choose payers according to their initial health. Although half of the clients in the sample cannot choose their payers (Medicaid or OSA), the other half (Third Party or Self-Paid) may have some choice.[14] In addition, the agency may attempt to match clients of worse initial health with generous payers. For instance, an agency may seek to enroll regular clients in Medicaid, and these clients may have larger health needs. To see if there is such a correlation between a client's payer and his health, we display in Table 4 the reported alcohol usage by payers. The severity indexes in Table 4 are almost identical across payers, though Medicaid and Third-Party clients have slightly worse severity; this implies that there is only a weak correlation between a client's health and his observed severity.[15]

A client's health care use is defined as the number of attended visits in an episode. Table 5 displays the mean and median of attended visits for complete (q) and incomplete episodes (Q). To show the relation between the health care use and demand or supply prices, we further separate them by the client's payer. The average number of attended visits is 20.7 for complete and 7.9 for incomplete episodes. Clearly, the number of attended visits is much higher for completers. Medicaid and Third Party clients, two groups with a high supply price, have more visits, especially for completed episodes. Finally, the sample median is consistently lower than the average for quitters, completers and the entire sample: health care use is positively skewed.

A client's health outcome is measured by the reported frequency of alcohol usage at discharge and categorized into the same usage groups as in Table 2. In MATS, an abstinent client falls in our "light" user category. As seen in Table 6, over 70% of clients are abstinent and less than 4% are still heavy users after treatment. Comparing with Table 4, we see at least a 20% increase in light users and a 6% decrease in heavy users, indicating an improvement in reported health.

---

[14]Clients need to meet certain critiria to be eligible for Medicaid programs. For OSA clients, their incomes also have to be below a certain threshold.

[15]We also check if there is a correlation between a client's clinical assessment and his payer. We find a similar pattern: Medicaid and Third Party clients have health slightly worse than the average.

5 Empirical Results

5.1 Basic Specification

We consider first a basic specification. The basic specification lets a client's demand depend on his personal characteristics, health and income status, as well as supply and demand prices.[16] Because a provider knows all the information in the data when setting her quantity limit, all the demand variables are also included in the provider's limit. In addition, we include provider dummies in the provider's limit to allow that each provider may have its specific practice style and select its preferred clients accordingly. Finally, due to the skewness in health care use, the client's health care use is measured by the log of the total number of visits.

5.1.1 General Model

Table 7 presents the estimated results of the client's demand (7) and the provider's limit (8) in the general model (Equations (7), (8) and (10)). Our estimates suggest that aged, educated or working clients demand more health care, as do clients with higher incomes; these clients may be more willing to attend treatment. Clients with legal involvements or severe clinical assessment demand less health care. On the other hand, the provider sets lower limits for male or older clients, but higher limits for those who have psychological problems. In addition, the provider sets higher limits for clients with worse health conditions; the worse is the client's health, the higher the limit.

The coefficients of supply and demand prices in (8) are .138 and -.186 respectively; both are significant. Given that the health use is measured by the log of the total number of visits, they suggest that a provider sets the limit 13.8% higher for clients having a high supply price, but 18.6% lower for clients having a high demand price. In terms of average number of visits (recall that the average number of visits for completers is 20), the estimates suggest that the limit is, relative to OSA clients, 2.7 visits higher for Medicaid clients; 1.0 lower for Third Party; 3.7 lower for Self Paid. By contrast, the estimated coefficient of demand and supply price in (7) are 0.014 (insignificant) and 0.153 (significant) respectively. The positive coefficient of supply price shows that a provider may employ effort or persuasion.

[16] A client's income is measured by the square root of his monthly income.

The last set of variables in the Table 7 include the standard deviations of all three random errors, the correlation between μ_d and μ_s, as well as α in equation (10). The random errors μ_d and μ_s are quite correlated (0.380); it indicates that unobserved characteristics are important determinants for clients to choose demands, and for providers to set limits. The provider, for instance, may set lower limits for client showing little interests in participating treatment. The constant α is estimated to be 0.621 and significant; it implies that if the client's demand is within a few visits of the provider's limit, they are willing to complete the treatment. The standard deviation of ε is 0.778 but insignificant.

Our model differs from the disequilibrium model as to how the completion indicator is observed (equation (10)). It is of interest to compare the estimated results of our model with that of the disequilibrium model (Equations (7), (8) and (9)). Table 8 presents the results of the disequilibrium model under the basic specification. A comparison between Table 7 and 8 reveals several significant differences. First, our general model has a larger likelihood value; this is not surprising since the general model permits a more flexible characterization of j^*. More interestingly, both σ_d and σ_d and ρ are found to be larger in the disequilibrium model, particular for the correlation coefficient; the unusually high correlation in the disequilibrium model to some extent suggests a relationship between q and Q.

5.1.2 Proposition 1: Absent Effort or Persuasion?

The goal of this paper is to use Propositions 1-3 to test for the presence of each mechanism. A pure quantity limit (Proposition 1) imposes a joint constraint: $\delta_s = \gamma_d = 0$ on the general model. Table 9 shows the values of the likelihood function (LR) of the unconstrained (general) and constrained (pure quantity) models. In Table 9, the value is -15442.3 for the general model and -15468.4 for the pure quantity model. Given that the LR value (52.2) is well above the critical value (9.2) at 1%, we reject the restriction and the hypothesis that a provider does not use effort or persuasion.

5.1.3 Proposition 2: Absent Quantity Ceiling?

Table 10 presents the estimated results of equation (16). From Table 10, the coefficient of predicted quantity is quite different from the expected value of pure effort-persuasion model (0), and even larger than that of the general model (1). The hypothesis that a provider does not set a quantity limit is rejected.

At this point we have evidence that providers use a quantity ceiling, and that they use *either* effort or persuasion. We now employ data on outcomes to differentiate effort from persuasion.

5.1.4 Proposition 3: Effort versus Persuasion

Proposition 3 tests if a client's health outcome is affected by the supply price, a proxy for effort. Given the ordinal health outcome measure, we estimate the health production function by the ordered logit, where better outcome measures are assigned for smaller numbers.[17] Furthermore, we stratify (17) according to the client's reported alcohol usages at admission because treatment may work differently for clients with different initial health status.[18] The estimated coefficients of the health production for each usage group are presented in Table 11.

According to Table 11, light users have quite different treatment effects than other users.[19] Attending more visits, for instance, contributes much less to a client's health outcome for light users than other clients. Our coefficient of interest, supply price, is insignificant across all usage groups. Since our health outcome is actually a severity measure (negative health), the effort would imply that the coefficient on the supply price is negative. It is positive in all the models. The findings fail to reject the hypothesis that a provider's action is entirely persuasion.

5.2 Robustness Checks

In this subsection, we examine if the results are robust to a number of diagnostic checks. We begin by checking if the results are robust to various model specifications. An endogeneity bias may arise if the client's severity has not been adequately controlled. This is especially important since our estimation strategy relies on the prices derived from the payer status. The supply price, for instance, may be overstated if severe clients choose more generous payers. Although the unobserved client's severity cannot be directly identified, its magnitude can be assessed by checking how the

[17]Specifically, the health outcome is denoted as 4 if a client is a heavy user at discharge; 3 for intermediate user; 2 for moderate user, and 1 for light user.

[18]Lu and McGuire (2002) finds that the treatment effects may depend on a client's severity at admission.

[19]Our results are consistent with the findings of Lu and McGuire (2002).

results respond to "observed" client severity. For this purpose, in Column (1) of Table 12, the estimation drops variables of clinical assessment in the client's demand and provider's limit.

As noted before, each provider may have its own practice style. Moreover, the specific practice style may be influenced by the provider's financial conditions. While this effect to some extent is controlled by the provider's dummies, to examine its possible impact, in Column (2) we further controls this effect by adding dummies of a provider's income.[20]

So far the specification assumes that clients do not perceive the specific practice style of each provider. This may seem unrealistic since such information is unlikely to be contained. To allow for the possibility that clients may select providers based on their needs, in Column (3), we include both the provider and income dummies in the client's demand.

In Columns (4)-(6), we check if the findings are robust to other distributional assumptions and sample compositions. In the basic specification, q is measured by the log number of attended visits. As a result, all the random errors follow the log-normal distribution. In Column (4), q is measured by the total number of visits and μ_d, μ_d, and ε are hence normally distributed. In Column (5) and (6), we re-estimate the model under the basic specification, with a more restricted sample. Clients who are legally involved may behave differently from those who are not. In Column (5), we use only clients that are not legally involved at admission. Our sample consists of clients with more than one episode. In Column (6), we include only their first episodes in the sample.

As shown in Table 12, all the estimated results strongly reject the pure effort-persuasion model, supporting the presence of quantity limits. We also reject the pure quantity ceiling model in every case, suggesting the presence of effort or persuasion as well. Lastly, the t value of the supply price is insignificant in almost all cases. Where it is not, the coefficient is positive and significant, and this is consistent with the presence of persuasion but not effort in the provider's action.

[20]Because a provider's income source is mainly determined by the payer mix of its client, we construct four dummies from income sources: "OSA Above," "Medicaid Above," "Third Party Above," and "Self Paid Above" to represent a provider's financial condition. The dummy "OSA Above" is set at 1 if the proportion of OSA clients for a provider in a fiscal year is higher than the sample average; likewise for other dummies. Our dummies are constructed in the same way as in Lu and McGuire (2002), where they found that a provider's financial condition has a large impact on the client's health care use.

6 Conclusions

Health care use is the result of an interaction between a client's health demand and a provider's health supply. To affect the utilization, a provider can either change his own supply, or influence a client's demand for health care. Three types of provider-patient interactions are discussed in the paper: rationing sets a quantity limit; effort and persuasion influence a client's health demand by productive (but noncontractible) and unproductive inputs respectively. Each mechanism is distinct and has different implications.

While previous research has emphasized the distinctions among these mechanisms in theory, there has been very little work that studies them empirically.[21] Instead, most empirical studies regard persuasion as the only available mechanism for providers, and interpret all the utilization change affected by supply factors as PID. In fact, utilization change may be affected by other mechanisms such as provider rationing and effort; each of these has welfare implications different from persuasion. For instance, while persuasion is not productive, effort is.

The distinction is particularly important in markets with competing managed care plans, where quantity limits and persuasion have very different implications for normative interpretation of a quantity effect of managed care. If plans ration by quantity limits, consumers' demand remains valid, and their choices among plans could be regarded as being in their best interest.[22] By contrast, to the extent that consumers' demand is manipulated by providers, consumers accede to the rationing, persuaded by the provider that quantities are in their best interest. Under this mechanism, consumers' choices among plans will not reliably signal their true self-interest.

In this paper, we construct a theoretical model that incorporates all three interactions as special cases. When the general model is specialized, a set of empirical implications emerges. We employ these implications to test for the presence of each mechanism, using data on substance abuse treatment in the Maine Addiction Treatment System. We find evidence for rationing, confirming the view that health providers ration services to prevent high use of some clients. The presence of

[21]See, however, Grytten et. al (2001) who distinguish an availability effect (rationing) from inducement effect (persuasion). Their work uses incomes of physicians who practice under various levels of competitions (high or low physician density) to separate these two effects. They do not consider effort as a mechanism affecting utilization.

[22]The choices do not necessarily lead to an efficient outcome in the market due to adverse selection and other potential market failures.

persuasion is also quite evident; it says that PID exists, even when accounting for mechanisms of effort and rationing. We do not find evidence supporting the presence of effort.

We now discuss some limits of our analysis. First, as we said earlier, using payers as proxies of prices of supply and demand may lead to some bias, since clients with high severity or treatment values are inclined to choose generous insurance. Although the sample shows a weak correlation between payer and severity, it is still possible that clients who benefit from treatment more are in plans with high supply prices. In that case, the effects of rationing and persuasion are likely to be overestimated.

Second, we use a static model of provider-client interaction. We have assumed that the quantities demanded by clients and supplied by providers are determined at the start of an episode. In practice, a provider's limit and a client's demand may be determined sequentially. It is likely that providers and clients modify their decisions using the information gathered during a course of treatment. A richer, sequential model could extend and refine our results.

This paper develops a framework for distinguishing mechanisms of health care provider-client interactions. While our findings may be specific to substance abuse treatment, our methods can be generalized to other health services. We expect other conclusions to be drawn, based on the features of each health service. The majority of health data sets do not contain information about who decides to end treatment, a key variable in the estimation. Our paper has demonstrated the importance of this key variable, particularly when providers might influence health care use.

References

[1] Ackerberg, D. Machado, M. P. and Riordan, M. H. (2001) "Measuring the Relative Performance of Providers of a Health Service," NBER Working Paper No. W8385.

[2] Cromwell, J. and Mitchell, J. B. (1986) "Physician-Induced Demand for Surgery," *Journal of Health Economics*, 293-313.

[3] Cuyler, A.J. (1989), "The Normative Economics of Health Care Finance and Provision," *Oxford Review of Economic Policy*, 5: 34-58.

[4] Dranove, D. (1988) "Demand Inducement and the Physician/Patient Relationship," *Economic Inquiry*, 26: 251-298.

[5] Eisenberg, J. M. (1986), *Doctors' Decisions and the Cost of Medical Care*, Health Administration Press, Ann Arbor, Michigan.

[6] Evans, R.C. (1974) "Supplier-induced Demand: Some Empirical Evidence and Implications," in Mark Perlmann ed., *The Economics of Health and Medical Care*, London, Macmillan, 162-173.

[7] Farley, P.J. (1986) "Theories of the Price and Quantity of Physician Services," *Journal of Health Economics*, 315-333.

[8] Fuchs, V.R. (1978) "The Supply of Surgeons and The Demand for Operations," *Journal of Human Resources*: 32-56.

[9] Gaynor, M. (1994) "Issues in the Industrial Organization of the Market for Physician Services," *The Journal of Economics and Management Strategy*, 211-255.

[10] Gaynor, M., and Gertler, P. (1995) "Moral Hazard and Risk Spreading in Partnerships," *RAND Journal of Economics*, 26: 591-614.

[11] Grytten, J.et. al. (2001) "The Income Effect and Supplier Induced Demand. Evidence from Primary Physician Services in Norway," *Applied Economics*, 33 (11): 1455-1467.

[12] Gruber, J. and Owings, M. (1996) "Physician Financial Incentives and Cesarean Section Delivery," *RAND Journal of Economics*, 27: 99-123.

[13] Hillman, A., et al. (1992) "Physicians' utilization and charges for outpatient diagnostic imaging in a Medicare Population," *Journal of the American Medical Association*, 268: 2050-2054.

[14] Kessler, D. P. and McClellan, M (1996) "Do Doctors Practice Defensive Medicine?," Quarterly *Journal of Economics*, 111: 353-390.

[15] Lu, M. and McGuire, T.G. (2002) "The Productivity of Outpatient Treatment for Substance Abuse," *Journal of Human Resources, 27 (2): 309-335.*

[16] Ma, C.A. and McGuire, T.G. (1997) "Optimal Health Insurance and Provider Payment," *American Economic Review*, 87(4): 685-704.

[17] Machado, M.P (2001) "Dollars and Performance: Treating Alcohol Misuse in Maine," *Journal of Health Economics,* 20(4): 639-67.

[18] Maddala, G. S. (1983) *Limited-Dependent and Qualitative Variables in Econometrics*, Economic Society Monographs in Quantitative Economics.

[19] McGuire, T.G. and Pauly, M.V. (1991) "Physician Response to Fee Changes With Multiple Payers," *Journal of Health Economics*, (10) 4: 385-410.

[20] McGuire, T.G. (2000) "Physician Agency," in Culyer and Newhouse (eds) *Handbook of Health Economics* , North Holland.

[21] Pauly, M.V. (1980) *Doctors and Their Workshops: Economic Models of Physician Behavior*, University of Chicago Press.

[22] Rice, T. (1983) "The Impact of Changing Medicare Reimbursement Rates on Physician- Induced Demand," *Medical Care*, 21: 803-815.

[23] Rizzo, J.A. and Blumenthal, D. (1994) "Physician Labor Supply: Do Income Effects Matter," *Journal of Health Economics*, 13:433-453.

[24] Shen, Yujing "Selection Incentives in a Performance-Based Contracting System," *Health Service Research*, forthcoming.

[25] Stano, M. (1987) "A Further Analysis of the Physician Inducement Controversy," *Journal of Health Economics*, 6: 229-238

[26] Yip, W. (1998), "Physician Responses to Medical Fee Reductions: Changes in the Volume and Intensity of Supply of Coronary, Artery Bypass Graft (CABG) Surgeries in the Medicare and Private Sectors," *Journal of Health Economics, (*17) 6: 675-699

[27] Williams, A. (1998), "Medicine, Economics, Ethics, and the NHS: A Clash of Cultures?" *Health Economics*, 7:565-568.

Appendix

Let q^* and j^* be a client's observed quantity and completion indicator. The likelihood function of the general model (equations (7), (8) and (10)) can be written as

$$\begin{aligned}
\ln f(q^*, j^*) &= j^* \times \ln f(q = q^*, j^* = 1) \\
&\quad +(1 - j^*) \times \ln f(Q = q^*,\ j^* = 0) \\
&= j^* \times [\ln f(q = q^*) + \ln f(j^* = 1|q = q^*)] \\
&\quad +(1 - j^*) \times [\ln f(Q = q^*) + \ln f(j^* = 0|Q = q^*)]
\end{aligned}$$

where the second line applies the conditional expectation.

The values of $\ln f(q = q^*)$ and $\ln f(Q = q^*)$ can be derived straightforward since μ_s and μ_d are both normally distributed. Rewriting the conditional probability of an episode being incomplete conditional on quantity q^* as

$$\begin{aligned}
f(j^* = 1|q = q^*) &= f(q^* < Q - \alpha + \varepsilon | q = q^*) \\
&= f(S + \alpha - \mu_s^{'} | \mu_d = D)
\end{aligned}$$

where $D \equiv q^* - X_d\beta_d - \delta_d p_d - \gamma_d p_s$, $S \equiv q^* - X_s\beta_s - \delta_s p_d - \gamma_s p_s$, and $\mu_s^{'} = \mu_s + \varepsilon$. Since $\mu_s^{'}$ or μ_d are jointly normally distributed[23], we have[24]

$$f(j^* = 1|q = q^*) = \Phi\left(\frac{\frac{\rho\sigma_d}{\sigma_s} D - \alpha - S}{\sqrt{\sigma_s^2(1-\rho^2) + \sigma_\varepsilon^2}}\right).$$

Likewise,

$$f(j^* = 0|Q = q^*) = \Phi\left(\frac{\frac{\rho\sigma_s}{\sigma_d} S + \alpha - D}{\sqrt{\sigma_d^2(1-\rho^2) + \sigma_\varepsilon^2}}\right).$$

---

[23]Because μ_s or μ_d are bivariate normal and ε is uncorrelated with either μ_s or μ_d, $\mu_s^{'}$ are μ_d are also jointly normal.

[24]Let X_1 and X_2 be bivariately normally distributed: $(X_1, X_2) \sim BVN(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, where μ_1 and μ_2 are means of X_1 and X_2 respectively, σ_1^2 and σ_2^2, variances, and ρ is the correlation coefficient between X_1 and X_2. Then $X_1| \ X_2 \sim N(\alpha + \beta X_1, \sigma^2)$, where $\beta = \frac{\rho\sigma_2}{\sigma_1}$, $\alpha = \mu_2 - \beta\mu_1$, and $\sigma^2 = \sigma_2^2 - \beta^2\sigma_1^2$.

Thus,

$$\begin{aligned}\ln f(q^*,\ j^*) &= j^* \times \left[\ln \phi(\frac{D}{\sigma_d}) - \ln \sigma_d + \ln \Phi\left(\frac{\frac{\rho\sigma_d}{\sigma_s} D - \alpha - S}{\sqrt{\sigma_s^2(1-\rho^2)+\sigma_\varepsilon^2}}\right)\right] \\ &\quad +(1-j^*) \times \left[\ln \phi(\frac{S}{\sigma_s}) - \ln \sigma_s + \ln \Phi\left(\frac{\frac{\rho\sigma_s}{\sigma_d} S + \alpha - D}{\sqrt{\sigma_d^2(1-\rho^2)+\sigma_\varepsilon^2}}\right)\right].\end{aligned}$$

Table 1: Determinants of Provider-Patient Interactions on Quantity of Health Care

	Persuasion	Nonretradability Allows Quantity Setting	Choice of Non-Contractible Input
Main Features	Physicians take actions to persuade; constrained by demand response or ethics	Supply determination within demand constraints	Demand response to "quality" or other costly physician inputs
Illustrative Paper	Dranove (1988)	Farley (1986)	Ma and McGuire (1997)
Information	Asymmetric	Complete	Complete
Physician s action in influencing use	Unobservable	N/A	Not Contractible
Physician s action in influencing Outcome	No	N/A	Yes

Table 2: Basic Characteristics for Clients at Admission

		Mean Percent	Standard Deviation	Median
Age (Years)		33.2	10.7	32.2
Education (Years)		11.5	2.2	12.0
Income ($ Per Month)		928.3	875.8	500.0
Male		69.9%		
Married		21.7%		
Psych problem		13.7%		
Legal involvement		40.8%		
Employment				
	Full time	28.6%		
	Part time	9.8%		
	Unemployed	61.6%		
Reported frequency				
	Light	47.7%		
	Moderate	18.8%		
	Intermediate	24.1%		
	Heavy	9.4%		
Clinical assessment				
	Causal	6.8%		
	Life-style involved	23.5%		
	Life-style dependent	47.5%		
	Dysfunctional	22.2%		
Total Obs		7615		

Table 3: Payer Status and Prices of Demand and Supply

Payer Status	Demand Price	Supply Price	
OSA	Low	Low	21.9%
Medicaid	Low	High	29.4%
Self-Paid	High	Low	18.5%
Third Party	High	High	30.2%

Table 4: Payer Status and Reported Frequency of Usage at Admission

	Light	Moderate	Intermediate	Heavy
OSA	48.0%	18.6%	24.2%	9.2%
Medicaid	45.9%	22.1%	23.3%	8.7%
Self-Paid	52.7%	15.9%	22.5%	8.9%
Third Party	42.0%	18.5%	28.1%	11.4%

Table 5: Payer Status and Health Care Use (Visits)

	Completer (Q)		Quitter (q)		Total	
	Mean	Median	Mean	Median	Mean	Median
OSA	17.9	12	6.6	4	10.2	6
Medicaid	27.4	18	9.5	5	15.2	7
Self-Paid	17.8	14	7.5	4	12.2	7
Third Party	20.1	15	7.4	4	13.4	8
Total	20.7	14	7.9	4	12.9	7

Table 6: Payer Status and Reported Frequency of Usage at Discharge

	Light	Moderate	Intermediate	Heavy
OSA	73.2%	12.5%	10.7%	3.6%
Medicaid	68.4%	16.9%	11.1%	3.6%
Self-Paid	78.0%	9.6%	9.0%	3.4%
Third Party	73.3%	12.6%	10.8%	3.3%
Total	73.3%	12.9%	10.3%	3.5%

Table 7: Estimated Coefficients of the General Model

	A Client s Demand	A Provider s Supply
Provider dummies	No	Yes
Age (years)	0.010	-0.003
	[0.002]**	[0.002]*
Male	-0.130	-0.151
	[0.036]**	[0.034]**
Married	-0.143	-0.052
	[0.039]**	[0.035]
Education (years)	0.026	-0.005
	[0.008]**	[0.007]
No legal involvement	-0.254	0.042
	[0.035]**	[0.036]
Full time	0.140	-0.085
	[0.043]**	[0.038]*
Part time	0.161	-0.052
	[0.054]**	[0.049]
Income (square root)	0.003	-0.002
	[0.001]*	[0.001]
Psych problem	-0.024	0.111
	[0.045]	[0.046]*
Life-style involved	-0.217	0.092
	[0.043]**	[0.042]*
Life-style dependent	-0.346	0.167
	[0.041]**	[0.047]**
Dysfunctional	-0.360	0.276
	[0.057]**	[0.068]**
Moderate user	-0.040	0.313
	[0.073]	[0.059]**
Intermediate user	-0.141	0.573
	[0.074]	[0.069]**
Heavy user	-0.149	0.514
	[0.076]	[0.070]**
Supply price	0.153	0.138
	[0.034]**	[0.031]**
Demand price	0.014	-0.186
	[0.035]	[0.036]**
Constant	1.712	
	[0.126]**	
σ_d	1.181	[0.081]**
σ_s	0.885	[0.140]**
ρ	0.380	[0.151]**
α	0.621	[0.052]**
σ_ε	0.778	[0.501]
Likelihood Value	-15442.3	
Observations	7615	

1. * significant at 5%; ** significant at 1%; the number in the parenthesis denotes the standard deviation.
2. the constant in the provider limit equation is omitted due to providers' dummies.

Table 8: Estimated Coefficients of the Disequilibrium Model

	A Client s Demand	A Provider s Supply
Provider dummies	No	Yes
Age (years)	0.01	0.003
	[0.002]**	[0.001]*
Male	-0.14	-0.147
	[0.037]**	[0.027]**
Married	-0.148	-0.093
	[0.040]**	[0.029]**
Education (years)	0.029	0.01
	[0.008]**	[0.006]
No legal involvement	-0.279	-0.105
	[0.035]**	[0.028]**
Full time	0.153	0.019
	[0.042]**	[0.031]
Part time	0.17	0.054
	[0.054]**	[0.040]
Income (square root)	0.003	0
	[0.001]**	[0.001]
Psych problem	-0.032	0.045
	[0.046]	[0.036]
Life-style involved	-0.213	-0.047
	[0.043]**	[0.032]
Life-style dependent	-0.344	-0.074
	[0.040]**	[0.033]*
Dysfunctional	-0.374	-0.034
	[0.057]**	[0.048]
Moderate user	-0.004	0.163
	[0.069]	[0.049]**
Intermediate user	-0.103	0.26
	[0.068]	[0.051]**
Heavy user	-0.118	0.223
	[0.072]	[0.053]**
Supply price	0.149	0.142
	[0.034]**	[0.025]**
Demand price	0.021	-0.097
	[0.036]	[0.028]**
Constant	1.822	
	[0.112]**	
σ_d	1.311	[0.071]**
σ_s	0.905	[0.150]**
ρ	0.872	[0.181]**
Likelihood Value	-15482.3	
Observations	7615	

1. * significant at 5%; ** significant at 1%; the number in the parenthesis denotes the standard deviation.
2. the constant in the provider limit equation is omitted due to providers' dummies.

Table 9: Hypotheses Tests for Propositions 1 (Basic Specification)

	Likelihood Value	LR Value
General model	-15442.3	
Pure quantity	-15468.4	52.2**

. * significant at 5%; ** significant at 1%

Table 10: Estimated Coefficients of the Pure Effort-Persuasion Model

	Coefficient	Standard Deviation
Age (years)	-0.001	[0.001]
Male	0.020	[0.028]
Married	0.011	[0.030]
Education (years)	-0.006	[0.006]
No legal involvement	0.053	[0.026]*
Full time	-0.031	[0.031]
Part time	-0.034	[0.040]
Income (square root)	-0.001	[0.001]
Psych problem	0.021	[0.035]
Life-style involved	0.028	[0.032]
Life-style dependent	0.056	[0.029]
Dysfunctional	0.085	[0.043]
Moderate user	0.005	[0.050]
Intermediate user	0.028	[0.048]
Heavy user	0.025	[0.052]*
Supply price	-0.013	[0.025]
Demand price	-0.011	[0.026]
Predicted Quantity (General Model)	1.080	[0.022]**
Constant	-0.001	[0.001]
Observations	7615	

Note:. * significant at 5%; ** significant at 1%

Table 11: Estimation of Health Outcomes with Clinical Assessment

	Light	Moderate	Intermediate	Heavy
Age (years)	0.001	-0.026	0	-0.007
	[0.007]	[0.007]**	[0.005]	[0.008]
Male	-0.327	-0.118	-0.186	0.041
	[0.166]*	[0.132]	[0.118]	[0.198]
Married	0.136	0.146	-0.056	-0.145
	[0.182]	[0.151]	[0.132]	[0.207]
Education (years)	-0.13	0.038	-0.034	-0.032
	[0.034]**	[0.031]	[0.024]	[0.037]
No legal involvement	-0.124	0.357	0.353	0.339
	[0.161]	[0.128]**	[0.117]**	[0.192]
Full time	-0.215	0.078	-0.176	-0.593
	[0.202]	[0.154]	[0.131]	[0.219]**
Part time	-0.157	0.247	-0.04	0.272
	[0.262]	[0.195]	[0.181]	[0.335]
Income (square root)	-0.012	-0.004	-0.003	-0.003
	[0.006]*	[0.004]	[0.004]	[0.005]
Life-style involved	0.365	0.499	0.654	0.066
	[0.194]	[0.165]**	[0.149]**	[0.236]
Life-style dependent	0.168	0.201	-0.376	-1.071
	[0.343]	[0.193]	[0.302]	[1.315]
Dysfunctional	0.475	0.427	-0.409	-0.756
	[0.320]	[0.194]*	[0.293]	[1.281]
Psych problem	0.256	0.473	-0.147	-0.628
	[0.341]	[0.217]*	[0.307]	[1.284]
Total visits (log)	-0.028	-0.722	-1.054	-1.075
	[0.063]	[0.055]**	[0.053]**	[0.085]**
Supply price	0.232	0.208	0.135	0.3
	[0.156]	[0.125]	[0.109]	[0.176]
Observations	3629	1436	1831	691

1. significant at 5%; ** significant at 1% ; the number in the parenthesis denotes standard deviation.
2. the health outcome is denoted as 4 if a client is a heavy user at discharge; 3 for intermediate user; 2 for moderate user, and 1 for light user.
3. the cut off values in the ordered logit are omitted.

Table 12: Hypotheses Tests for Propositions 1-3 (Basic and Other Specifications)

			(1)	(2)	(3)	(4)	(5)	(6)
Sample			Full	Full	Full	Full	Legally Involved	First Episode
Specification	Client s Demand (q)	Income Dummies			x			
		Provider Dummies			x			
	Provider s Limit (Q)	Income Dummies		x	x			
		Provider Dummies	x	x	x	x	x	x
Prop 1	Likelihood value	General	-15524.6	-15439.4	-15036.2	-38419.3	-8920.2	-13024.8
		Effort/Persuasion	-15554.6	-15465.0	-15098.0	-38433.4	-8939.4	-13049.8
	LR value		60.0**	51.2**	50.0**	28.2**	38.4**	50.0**
Prop 2	T value		47.04**	49.24**	43.53*	42.28**	48.28**	49.20**
Prop 3 (light)			1.57	1.48	1.97*	1.62	0.95	1.67
Prop 3 (moderate)			1.63	1.66	0.88	0.93	1.59	1.23
Prop 3 (intermediate)			1.34	1.24	0.34	1.09	0.82	1.59
Prop 3 (heavy)			1.70	1.70	1.40	0.74	1.88	1.48
Total Obs			7615	7615	7615	7615	4356	6778

*: significant at 5%; **: significant at 1%