

1 Model

1.1 The Physician-Patient Market

We now present the model itself before proceeding with further discussion. We consider $i = 1, \dots, I$ patients and $j = 1, \dots, J$ physicians.¹ Patients are characterized by their “medical necessity” $\kappa_i \in \mathbb{R}^+$ and their “taste for sick leave” $\gamma_i \in \mathbb{R}$, respectively following the ex-ante cumulative distributions $F(\kappa)$ and $G(\gamma)$, which are public knowledge. Physicians are described by their “service quality” $V_j \in \mathbb{R}^+$ and cost of visit $\tau_j \in \mathbb{R}^+$, also known both to all patients and other physicians, as well as their revenue function $R_j(\cdot)$.

A patient i can visit a physician for treatment and may be also granted sick leave. After the patient is assigned a physician j , his utility function—implicitly dependent on his characteristic (κ_i, γ_i) tuple—is defined piecewise as follows:

$$U_i(V_j) = \begin{cases} \gamma_i + V_j \kappa_i - \tau_j & \text{if he's granted sick leave,} \\ V_j \kappa_i - \tau_j & \text{if he only visits the physician,} \\ 0 & \text{if no visit takes place} \end{cases}$$

As we see, there are three components to patient utility: an interaction between the patient’s medical need κ_i and the physician’s service quality V_j which implies their complementarity, his “taste” for sick leave γ_i in the case he’s granted one, and τ_j , which we define as the cost of visit for physician j . As well as being complementary, κ_j as multiplied by V_j would give it the interpretation of being the marginal willingness to pay for physician quality.²

Whereas patients may visit at most one physician, a physician may see several patients. We define Q_j as the expected number of patients for physician j , the demand for her services. We say ‘expected’ because, as we’ll see later on, patients may opt for a mixed strategy, assigning a certain *probability* to visiting j , and we define Q_j as the sum of the ex-ante probability of visit of all patients, not their ex-post realization.

As physician j has the option to issue sick leave to a given patient i which visits her, we likewise define X_j as the *expected* number of sick leaves granted by physician j , given her ex-ante patient demand and how many of them would be granted one.

We define the physician’s utility function as follows:

$$U_j(Q_j, X_j) = R_j(Q_j) - P(X_j)$$

where $R_j(\cdot)$ is an individual, growing, concave *revenue* function defined over expected total clients, and $P(\cdot)$ is a growing, convex *punishment* function on X_j , composed

¹Note on language: For ease of reference, we refer to any patient as “he”, and to any physician as “she” (as we have done already). At this point it shall be noted that the first person plural (we) employed over the course of this article is to be read as including the reader or perhaps as a royal *we*.

²Conceptually, willingness to pay would also depend on patient i ’s wealth level w_i . To steer away from discussions on wealth disparity, we interpret κ_i as the *average* willingness to pay for V_j .

of the probability of being fined for a certain number of sick leaves issued and the magnitude of the fine. The implication is that after a given number of patients, the disutility of an additional (expected) sick leave issued would outweigh physician j 's financial incentive for further clientele. The j subindex indicates that we will allow $R_j(\cdot)$ (revenue by visit) to vary across physicians. We assume both $R_j(\cdot)$ and $P(\cdot)$ to be twice differentiable.

We stress again that we define physician utility in terms of the *expected* realization of patient demand and granted sick leaves, as befits the logic of this game, where doctors take action before patients (we will specify the timing of our game below).

Following ?, we focus on *threshold equilibria*, wherein each physician's strategy is the choice of a value $\bar{\kappa}_j \in \mathbb{R}_0^+$, such that of the patients who visit j , those with a κ_i value above or at that threshold will be granted sick leave, and those strictly under it won't. As a result of this, each physician will be identified by their given 'service quality', revenue function, cost of visit to the patient and choice of threshold: $\{V_j, R_j(\cdot), \tau_j, \bar{\kappa}_j\}$. The set $\{\bar{\kappa}_j\}_{j=1}^J$ will be known to patients when deciding on their behavior strategy.

Our models are different iterations of a general game with the following timing:

- First stage: Physicians simultaneously choose $\bar{\kappa}_j$.
- Second stage: Observing the set $\{V_j, \tau_j, \bar{\kappa}_j\}_{j=1}^J$, each patient chooses (or is assigned to) some doctor j .
- Third stage: Each patient can choose to see their physician and incur a visit cost τ_j , or refrain from doing so.

Conditional upon his visit (as not to visit renders null utility), the utility of patient i from seeing physician j is:

$$U_i(V_j, \bar{\kappa}_j) = \begin{cases} \gamma_i + V_j \kappa_i - \tau_j & \text{if } \kappa_i \geq \bar{\kappa}_j, \\ V_j \kappa_i - \tau_j & \text{if } \kappa_i < \bar{\kappa}_j \end{cases}$$

We will usually use u_{ij} as a shorthand.

1.2 Non-Search Equilibrium

For illustration purposes, we first devote attention to a non-search baseline, where each patient is randomly assigned to a physician, with an equal probability of being matched to any of the J physicians. Their only say in the matter is whether they'll then visit physician j , which is to say, the second stage of the game is out of their hands, and they only make choices in the third stage after assignment.

A patient won't visit his assigned physician if his expected utility from such a visit is negative, we call this the *free disposal* requirement. As such, a physician j 's expected patient demand, as a function of her threshold $\bar{\kappa}_j$ (and given the parameter V_j) will be the following:

$$Q_j(\bar{\kappa}_j) = \frac{I}{J} \left[\int_{\tau_j/V_j}^{\infty} dF(\kappa) + \int_{\min\{\bar{\kappa}_j, \tau_j/V_j\}}^{\tau_j/V_j} \int_{\tilde{\gamma}(\kappa)}^{\infty} dG(\gamma) dF(\kappa) \right] \quad (1)$$

where the left term consists of the mass of patients who just by virtue of physician j 's service quality V_j would be willing to pay a visit (i.e. $\kappa_i \geq \tau_j/V_j$), and the right term would be patients who only see physician j solely out of the expectation of getting sick leave ($\kappa_i \geq \bar{\kappa}_j$ & $\gamma_i \geq \tau_j - V_j\kappa_i$), and wouldn't visit otherwise. We define $\tilde{\gamma}(\kappa) := \tau_j - V_j\kappa$ as the lower limit of the inner integral over γ .

Given that each patient with a κ_i higher or equal to $\bar{\kappa}_j$ is granted sick leave, the expected total number of such certificates granted by j , as a function of $\bar{\kappa}_j$, is:

$$X_j(\bar{\kappa}_j) = \frac{I}{J} \int_{\bar{\kappa}_j}^{\infty} \int_{\tilde{\gamma}(k)}^{\infty} dG(\gamma) dF(k) \quad (2)$$

which includes in theory patients from both the left and right term of equation (1).

For given formulations of the functions $Q_j(\bar{\kappa}_j)$ and $X_j(\bar{\kappa}_j)$, each physician solves for the following constrained optimization:

$$\bar{\kappa}_j^* \equiv \arg \max_{\bar{\kappa}_j} R_j(Q_j) - P(X_j) \quad \text{s.a.} \quad 0 \leq \bar{\kappa}_j \leq \bar{\kappa}_{\max} \quad (3)$$

where $\bar{\kappa}_{\max} < \infty$ is the maximum value the choice of threshold of any physician may take, as we mentioned earlier.??

An inner solution to problem (3) takes on the following form because of the first order condition:

$$R'_j(Q_j) \frac{\partial Q_j}{\partial \bar{\kappa}_j} = P'(X_j) \frac{\partial X_j}{\partial \bar{\kappa}_j} \quad (4)$$

Proposition 1.1. *In the non-search model as described thus far, when the solution $\bar{\kappa}_j^*$ of problem (3) $\in (0, \frac{\tau_j}{V_j}]$, this $\bar{\kappa}_j^*$ is described by:*

$$R'_j(Q_j(\bar{\kappa}_j)) = P'(X_j(\bar{\kappa}_j))$$

Proof. For $Q_j(\bar{\kappa}_j)$ and $X_j(\bar{\kappa}_j)$ as defined by (1) and (2), when $\bar{\kappa}_j \in (0, \frac{\tau_j}{V_j}]$:

$$\frac{\partial Q_j}{\partial \bar{\kappa}_j} = -\frac{I}{J} \int_{\tilde{\gamma}(\bar{\kappa}_j)}^{\infty} dG(\gamma) f(\bar{\kappa}_j) = \frac{\partial X_j}{\partial \bar{\kappa}_j}$$

such that both derivatives are cancelled in equation (4). As a result, if the optimal $\bar{\kappa}_j^* \in (0, \frac{\tau_j}{V_j}]$, the proposition follows. \square

This proposition implies that when for low enough values of κ_i , any new patients gained by physician j are those she entices through the expectation of getting sick leave, meaning that in the vicinity of $\bar{\kappa}_j \in (0, \frac{\tau_j}{V_j}]$ the change in expected patients $\Delta Q_j / \Delta \bar{\kappa}_j$ through a variation in $\bar{\kappa}_j$ is the same as the change in expected sick leaves issued $\Delta X_j / \Delta \bar{\kappa}_j$.

We can see that (3) offers three possibilities by way of solution: two corner solutions at 0 and $\bar{\kappa}_{\max}$, and an inner solution fulfilling equation (4). The proposition below proves such an inner solution can't hold for $\bar{\kappa}_j \in (\frac{\tau_j}{V_j}, \bar{\kappa}_{\max})$.

Proposition 1.2. *In the non-search model as described thus far, no $\bar{\kappa}_j$ in $(\frac{\tau_j}{V_j}, \bar{\kappa}_{\max})$ can be the solution of problem (3).*

Proof. For $\bar{\kappa}_j \in (\frac{\tau_j}{V_j}, \bar{\kappa}_{\max})$,

$$Q_j(\bar{\kappa}_j) = \frac{I}{J} \left[\int_{\tau_j/V_j}^{\infty} dF(\kappa) \right] \text{ and thus } \frac{\partial Q_j}{\partial \bar{\kappa}_j} = 0$$

whereas

$$\frac{\partial X_j}{\partial \bar{\kappa}_j} = -\frac{I}{J} \int_{\bar{\gamma}(\bar{\kappa}_j)}^{\infty} dG(\gamma) f(\bar{\kappa}_j) < 0$$

so equation (4) can never hold, which would be required of $\bar{\kappa}_j$ as an inner solution to problem (3). \square

The intuition for this proposition is that $\kappa_i = \tau_j/V_j$ defines the threshold above which patients are willing to visit their assigned physician j regardless of whether they get sick leave, so choosing such a $\bar{\kappa}_j$ is always either too strict or inefficiently lenient. We can speak of a “captured clientele”. A similar behavior will be observed in the models with patient search, only that, since patients freely choose physicians, a “captured” patient won’t be one which always visits j , rather one that always assigns *positive probability* to visiting physician j , no matter her choice of threshold $\bar{\kappa}_j$.

Equilibrium in the non-search model. *Given a physician-patient market specified by $\{(\kappa_i, \gamma_i)\}_{i=1}^I, \{(V_j, R_j(\cdot), \tau_j)\}_{j=1}^J, P(\cdot)\}$, we define equilibrium as a set of physician thresholds $\{\bar{\kappa}_j\}_{j=1}^J$ satisfying problem (3) and their corresponding equilibrium (expected) patient demand and issued sick leaves $\{(Q_j, X_j)\}_{j=1}^J$ as described by (1) and (2), respectively.*

This non-search baseline model, based off?, already presents some elements which will be present in our more complex models later, but completely neglects patient behavior, who are reduced to being assigned a physician and choosing to visit her or not. Patient search will allow us to formalize their free choice of physician as a function of the model parameters and physician strategies.

1.3 General Search Framework

We introduce patient ‘search’ as a general framework in which patients can choose freely among all physicians.

We define for each patient i a vector $S_i \in \Delta(\mathcal{J})$, where \mathcal{J} is the the J -dimensional vector composed of all $1, \dots, J$ physicians. S_i represents the *strategy* of patient i in this game, specifying the probability of visiting each physician j . Each component s_{i1}, \dots, s_{iJ} within S_i denotes the probability that patient i will visit physicians $1, \dots, J$, respectively.

In order to play the role of a proper subprobability measure, for each patient i the components $s_{ij} \in S_i$ must fulfill the following criteria:

$$\forall j, s_{ij} \geq 0 \tag{i}$$

$$\sum_{j=1}^J s_{ij} \leq 1 \quad (\text{ii})$$

We will allow the sum of all components for each patient i to be less than one (hence *sub*-probability), implying the presence of an *outside option* for patients, that is, to not visit any physicians. To have this option is important, as patient rationality in our models entails a “*free disposal*” property, meaning that a patient will never visit a physician if his expected utility from such a visit is less than 0, that is, $s_{ij} = 0$ if $U_i(V_j, \bar{\kappa}_j) = 0$. This makes the third stage of the game as described in Section 1.1 trivial, as under free choice patient i will only be “assigned” with positive probability to physicians he will be willing to visit.

We can re-interpret the non-search model as each patient being made to play by the strategy $\{S_i : s_{i1} = s_{i2} = \dots = s_{iJ} = \frac{1}{J}\}$, and it is the lack of a free choice which makes the third stage non-trivial.

We shall further specify the form S_i will take. Consider now the J -dimensional vector U_i , where each component u_{i1}, \dots, u_{iJ} indicates the utility patient i expects from a visit to physicians $1, \dots, J$ respectively. As a shorthand, we shall write $u_{i,-j}$ to indicate the $J - 1$ components of U_i excluding u_{ij} .

In the models considered, each component s_{ij} of S_i will be defined as:

$$s_{ij} \equiv g_i(u_{ij}, u_{i,-j})$$

where $g_i(\cdot)$ is a continuous function weakly increasing in the first argument u_{ij} , and weakly decreasing in the remaining arguments given by $u_{i,-j}$. Our model specifications will consist in giving this function $g_i(u_{ij}, u_{i,-j})$ a specific form.

Defining s_{ij} over u_{ij} makes it dependent upon patient i 's characteristic tuple (κ_i, γ_i) . We will use the $s_{ij}(\kappa_i, \gamma_i)$ formulation to define physician expected demand and sick leaves issued as functions over patient strategies:

$$Q_j(\bar{\kappa}_j, \bar{\kappa}_{-j}) = I \int_0^\infty \int_0^\infty s_{ij}(\kappa, \gamma) dG(\gamma) dF(\kappa) \quad (5)$$

$$X_j(\bar{\kappa}_j, \bar{\kappa}_{-j}) = I \int_{\bar{\kappa}_j}^\infty \int_0^\infty s_{ij}(\kappa, \gamma) dG(\gamma) dF(\kappa) \quad (6)$$

At face value, these definitions appear less informative than their non-search equivalents in (1) and (2). This is because s_{ij} is more general than the patient ‘strategy’ of the non-search model (of being assigned to any physician with equal probability), so we can’t yet make claims about e.g. sections in the distribution of γ where s_{ij} is strictly 0. The two models we will introduce, the *implicit* and *explicit* search models, will build upon the general framework we’ve delineated and we will highlight their different qualitative implications.

For now though, notice that present we Q_j and X_j not only in terms of the threshold $\bar{\kappa}_j$ chosen by physician j , but also in terms of the thresholds of the other $J - 1$ physicians, which we abbreviate as $\bar{\kappa}_{-j}$. Whereas before in ‘non-search’ physicians

were simply allotted a given number of patients, now they will *compete* for them, as patients' S_i behavioral strategy will consider the whole of $\{\bar{\kappa}_j\}_{j=1}^J$ when considering which physician(s) they will visit with positive probability.

As a consequence of physician competition for patients, equilibrium $\{\bar{\kappa}_j^*\}_{j=1}^J$ will need to take on the form of a Nash equilibrium, as physicians take into consideration the strategies of the other $J - 1$ physicians when choosing their own, such that they adapt to the influence they expect other physicians will have on their own “client” base, which we call the *strategic effect*. We explore in the Appendix ?? the reasons for the presence of this effect, which is absent from ??’s models, the basis for our own. We attribute it to the lack of *additive separability* across patients in the physician’s utility function when one of $R_j(\cdot)$ or $P(\cdot)$ is *strictly* convex/concave.

Before moving on to our first model specification, we note that this general formulation of s_{ij} is sufficient to make the next proposition.

Proposition 1.3. *In the patient search framework as described, for two physicians j, l , such that $\bar{\kappa}_j \neq \bar{\kappa}_l$, $\bar{\kappa}_j$ and $\bar{\kappa}_l$ are strategic complements.*

Proof. The proof is in the Appendix ??.

□

1.4 The “Implicit” Search Model

The “implicit” search model is a McFadden Logit choice model modified to give null probability visits to physicians which afford patient i non-positive utility, so that the free disposal requirement is fulfilled. We also call it the Logit model.

The reason for calling it the “implicit” search model is that strictly speaking patient search is not formally included, yet one arrives at results qualitatively similar to specifications that do (like the “explicit” search model). The choice strategy noisily assigns positive probability to physicians who render i high utility, and the higher this u_{ij} expected utility is, the higher the probability of visit.

Instead of *explicitly* defining search and subsequent choice, probability of visit depends upon expected utility from physician j , u_{ij} , and a weighting parameter $\lambda \in \mathbb{R}^+$. The lower the value of λ , the noisier patient “search” is: they give high assignment probability to sub-optimal—yet feasible—physician choices. On the other hand, as $\lambda \rightarrow \infty$, patient choice is concentrated with probability approaching 1 at the physician with the highest V_j .

We define the components s_{ij} of the patient’s strategy vector S_i as follows:

$$s_{ij} = \frac{\alpha_{ij}}{\sum_{k=1}^J \alpha_{ik}}, \text{ where } \alpha_{ij} = \begin{cases} e^{\lambda u_{ij}}, & \text{if } u_{ij} > 0 \\ 0, & \text{if } u_{ij} \leq 0 \end{cases} \quad (7)$$

Unlike the “explicit” search model, the probability that patient i visits physician j is *strictly* growing in κ_i , rather than being a piecewise constant function with different levels. There is a discrete jump in probability at $\bar{\kappa}_j$, but elsewhere above 0 the function is smoothly increasing, rewarding physicians with high V_j .

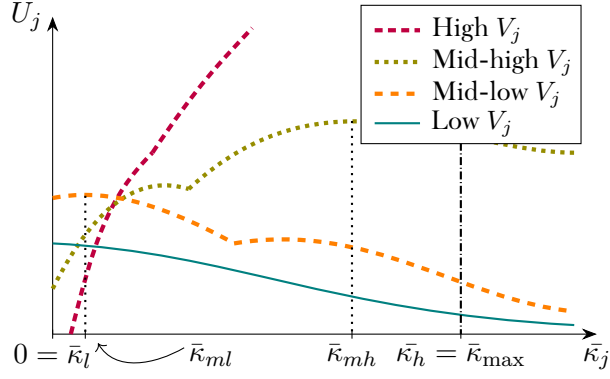


Figure 1: Physician utility U_j in the Logit model for different levels of V_j , as it varies across their choice of $\bar{\kappa}_j$

Figure 1 is a stylized representation of four “types” of physician, as ordered by their level of service quality V_j . The ‘low’ and ‘high’ types are driven to a corner solution at 0 and $\bar{\kappa}_{\max}$, respectively. We differentiate between two ‘middle’ types who choose inner solutions of $\bar{\kappa}_j$, as many physicians will present two local maxima across their curve, and their “type” is determined by which of the two is highest.

The reason curves present this piecewise behavior, as though composed of two individual concave curves joint together, is because for each physician j in this model there exists the threshold $\kappa_i = \tau_j/V_j$, as in the non-search model, above which patients always give positive probability to visiting j . Visiting chances may be *improved* upon by also granting sick leave to patients above that threshold (in terms of their κ_i), but taking it away doesn’t drive their s_{ij} towards 0, as it would with patients whose $\kappa_i < \tau_j/V_j$.

This difference in incremental patient loss through a raise in $\bar{\kappa}_j$ between the $[0, \frac{\tau_j}{V_j}]$ and $[\frac{\tau_j}{V_j}, \bar{\kappa}_{\max}]$ sections of the curve accounts for the presence of two peaks for some values of V_j . We call ‘mid-low’ those physicians whose first peak is higher than the second, the others we call ‘mid-high’ (as they’d have a higher V_j for the parametrizations we use later in this paper).

The fact that the breaking point of the curve occurs earlier the higher the V_j results from the fact that τ_j grows at a lower rate than 1 for increases in V_j in our simulations in Section ?? on which we based this graph, but this need not be the case. Notice also that $\bar{\kappa}_j$ is growing in V_j , which isn’t an analytical property from our model formulation, but an ex-post proclivity, which comes from using *well-behaved* (read: single-peaked, continuous) distributions for $F(\kappa)$ and $G(\gamma)$.

Equilibrium for the Logit model is formalized below.

Equilibrium in the “implicit” search model. *Given a physician-patient market specified by $\{(\kappa_i, \gamma_i)\}_{i=1}^I, \{(V_j, R_j(\cdot), \tau_j)\}_{j=1}^J, P(\cdot), \lambda\}$, we define equilibrium as a set of patient strategies $\{S_i : s_{i1} \dots s_{iJ}\}_{i=1}^I$ as defined by (7) and physician thresholds $\{\bar{\kappa}_j\}_{j=1}^J$ satisfying problem (3) with their corresponding equilibrium (expected) patient demand and issued sick leaves $\{(Q_j, X_j)\}_{j=1}^J$ as described by (5) and (6), respectively.*

1.5 The “Explicit” Search Model

The “explicit” search model, which we also call the sequential model, has patient strategies S_i be the *explicit* result of formalized physician search on part of the patient. This specification is more in line with ?, although made trickier by physician heterogeneity in *quality* (V_j) as well as “prescription” *leniency* — in our case, of sick leave issuance.

Given the two-fold source of patient utility from physicians, we will speak in general of patients’ ‘reserve utility’. In a similar vain to sequential job search models, we envision patients as searching for a physician in discrete time, and when assigned a physician j , they have the option to go visit them, which is equivalent to receiving utility u_{ij} in perpetuity, or look for a physician again, “drawing” another physician (with replacement) next period. We assume a temporal discount factor $\beta \in (0, 1)$, and no-recall, meaning a patient may only visit the current physician assigned.

This modelation can be defined in terms of dynamic programming. For each patient i , the solution to this problem will be an optimal threshold \bar{U}_i , henceforth their “reserve utility”, such that each patient is only willing to visit some physician j if the utility he would obtain from that visit, u_{ij} , is greater or equal to \bar{U}_i . We demonstrate in the Appendix ?? that the optimal characteristic equation for \bar{U}_i is the following:

$$\bar{U}_i = \frac{\beta}{1 - \beta} \sum_{j=1}^J \left\{ \frac{\mathbb{1}[u_{ij} \geq \bar{U}_i] \cdot (u_{ij} - \bar{U}_i)}{\mathbb{1}[u_{ij} \geq \bar{U}_i]} \right\} \quad (8)$$

The components s_{ij} of each patient’s strategy vector S_i are then defined in terms of patient i ’s ‘reserve utility’ \bar{U}_i :

$$s_{ij} = \frac{\mathbb{1}[u_{ij} \geq \bar{U}_i]}{\sum_{k=1}^J \mathbb{1}[u_{ik} \geq \bar{U}_i]} \quad (9)$$

Equation (9) makes the rather simplifying but reasonable assumption that patient i ‘draws’ physicians randomly and thus has equal probability to visit any of the physicians who give him utility above his threshold \bar{U}_i .

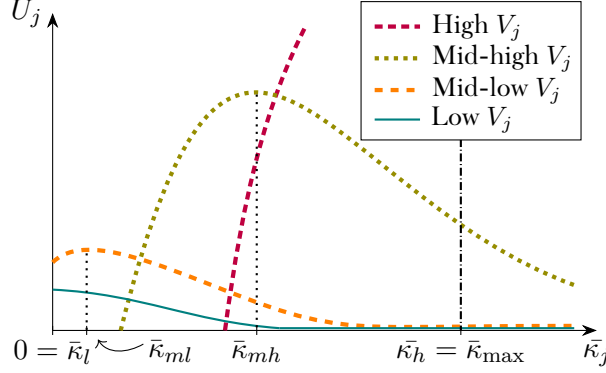


Figure 2: Physician utility U_j in the sequential model for different levels of V_j , as it varies across their choice of $\bar{\kappa}_j$

Figure 2 presents the stylized utility curves of physicians of different V_j in the manner of Figure 1 in the previous section. The curves are more *well-behaved* (read: quasi-concavity) than in the “implicit” search model for the simulations which form the basis of this figure, and could perhaps analytically be proven to have a unique pure strategy equilibrium under certain conditions.

One thing of note is that more often than in the Logit model, for (relatively) low enough values of V_j , physician utility will be strictly 0 for high choices of $\bar{\kappa}_j$, and for high values of V_j , utility will be *negative* at low thresholds $\bar{\kappa}_j$. This effect is more pronounced the higher the β parameter is, which plays the same role as λ in the previous model: the higher it is, the more patient demand will be concentrated at high quality physicians. A patient’s ‘patience’ β determines “how long” they’re willing to look for the ideal physician, i.e. of highest V_j . As $\beta \rightarrow 1$, demand for the physician with the highest V_j tends to be 1, and for anyone else to be 0.

As before, we formalize the notion of equilibrium in the sequential model.

Equilibrium in the “explicit” search model. *Given a physician-patient market specified by $\{(\kappa_i, \gamma_i)\}_{i=1}^I, \{(V_j, R_j(\cdot), \tau_j)\}_{j=1}^J, P(\cdot), \beta\}$, we define equilibrium as a set of patient thresholds $\{\bar{U}_i\}_{i=1}^I$ fulfilling (8), patient strategies $\{S_i : s_{i1} \dots s_{iJ}\}_{i=1}^I$ as defined by (9) and physician thresholds $\{\bar{\kappa}_j\}_{j=1}^J$ satisfying problem (3) with their corresponding equilibrium (expected) patient demand and issued sick leaves $\{(Q_j, X_j)\}_{j=1}^J$ as described by (5) and (6), respectively.*