

Choice of Urgent Care By Strategic But Uninformed Patients

Özden Engin Çakıcı

Kogod School of Business, American University, Washington, D.C., 20016, cakici@american.edu

Rajiv M. Dewan, Marshall Freimer

Simon Business School, University of Rochester, Rochester, NY, 14627, {dewan, freimer}@simon.rochester.edu,

Many patients have a choice of a provider for acute care, typically choosing between an Urgent Care facility (UC) or an Emergency Department of a hospital (ED). This choice is made by a patient who is generally not herself a trained medical professional. She is strategic but imperfectly informed in that she weighs all the costs and benefits of this choice, including quality of care, waiting time, co-payments and fees, but can only imperfectly assess the appropriateness of a facility for the perceived needed care. For instance, she might deem an ED as necessary even though an UC might suffice or she might deem UC as adequate when she really needs the full services of an ED. In the latter case, there will be a additional delay in getting the appropriate treatment, possibly affecting the outcome and other costs. Other factors affecting the choice are the expected waiting time at each facility and the co-payments and fees. UC are typically less congested than ED, and the waiting times for care are lower. The patients form expectations on the waiting time at each facility. The UC manager sets the fee for service in competition with an ED. We show that the errors in self-classification make UC less desirable, i.e, the UC has to discount its co-pay/fees further as the error rates increase. We also study the long run version of the problem which the UC manager sets the capacity of the center. We determine the impact of classification errors on the capacity choice.

Key words: service operations; healthcare operations management; urgent care; emergency care; classification errors; co-pay; capacity

History: This version: **March 6, 2015**. First version May 2014

1. Introduction

Since the early 1980s in US, some patients needing acute care have had a choice between a full fledged Emergency Department in a hospital or an Urgent Care facility that “delivers ambulatory medical care outside of a hospital emergency department on a walk-in basis, without a scheduled appointment¹.” The ED remains the the most comprehensive acute care treatment facility, dealing with cases that require immediate care to non-urgent cases. UC centers typically only provide

¹This definition and the statistics on Urgent Care centers in this and the next paragraph come from the industry trade association - Urgent Care Association of America (UCAOA). See especially their results from a survey in 2013. [UCAOA(2014)]

non-emergent care. In Table 1 we show the difference in the levels of care provided by EDs and UCs.

Emergency Severity Index [AHRQ(2014)]	Example of Condition	Percent [CDC(2011)]	Providers
Level 1 - Immediate	resuscitation, cardiac arrest, critically injured trauma patient	1.2%	ED only
Level 2 - Emergent	suicidal or homicidal patient, chemotherapy patient with fever, active chest pain	10.7%	ED only
Level 3 - Urgent	fractured ankle, abdominal pain, migraine	42.3%	ED and a few UC
Level 4 - Semi-urgent	sore throat and fever, stubbed toe, sprained ankle, minor laceration	35.5%	ED and UC
Level 5 - Nonurgent	poison ivy, toothache	8.0%	ED and UC

Table 1 Levels of Care

Urgent Care Centers have grown tremendously in popularity for providing acute care for non-emergent cases, i.e., for cases in levels 4 and 5 of Emergency Severity Index. The number of urgent care centers has grown from approximately 8,000 in 2008 to 9,300 in 2014, with 4 or more urgent care centers in cities with population of 100,000 or more².

Part of the reason for the rise of UC has been the congestion and wait times at emergency departments. Even as the patient demand for acute care has been increasing, the number of EDs have been decreasing in the U.S. (see Figure 1). The net effect is that the wait times have been increasing explosively. US Centers for Disease Control reports that between 2003 and 2009, the mean wait time to see a provider increased by 25% to almost an hour [Hing and Bhuiya(2012)]. The wait times are even worse for patients with non-emergent conditions. These patients who make up a significant proportion of patients at an ED, 43% in US in 2011 [CDC(2011)], are triaged into lower priority queues that have even longer wait times. These patients can be treated elsewhere such as in a UC [Weinick et al.(2010)Weinick, Burns, and Mehrotra] where the wait times are much lower. The wait time distribution at EDs and UCs are shown in Tables 2 and 3. We note that only 6% of patients at an UC waited more than 45 minutes where as 33.3% of the patients and ED waited an hour or more. Average wait times at each ED in US is available publicly, for instance

² Ibid.

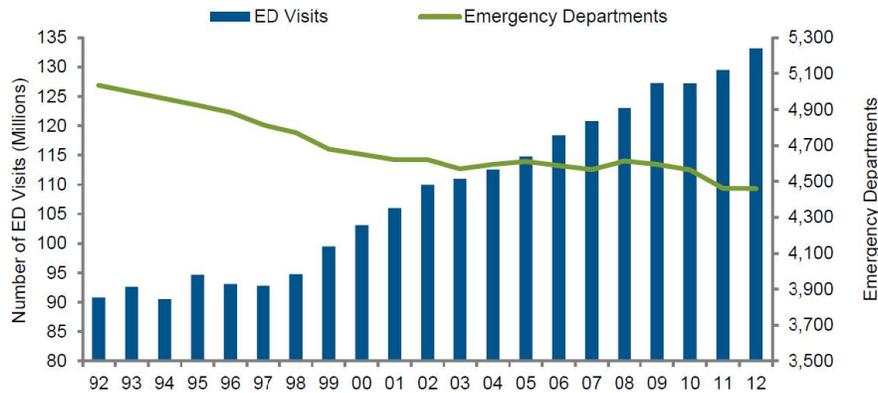


Figure 1 Emergency Department Visits and Emergency Departments in Community Hospitals, 1991 - 2012

Source: Avalere Health analysis of American Hospital Association Annual Survey data, 2012, for community hospitals. Hospitals are the ones reporting ED visits in the AHA Annual Survey. Available at: <http://www.aha.org/research/reports/tw/chartbook/2014/14chartbook.pdf>

at *ER Wait Watcher*³ which provides the average delay and travel time to near by EDs from any address in US. A screen-shot is shown in Figure 2. Diversion of some non-emergent patients from the ED to the UC can help all patients, but especially those who have non-emergent conditions.

Time to see a professional	% of visits
Fewer than 15 minutes	27.0%
15-59 minutes	40.7%
1 to 2 hours	13.5%
2 to 3 hours	4.8%
3 to 4 hours	1.6%
4 to 6 hours	1.0%
6 hours or more	0.7%

Table 2 Wait times at emergency departments

While most Emergency Departments are affiliated with hospitals and are typically located in urban environments, Urgent Care centers are located in suburbs and in locations convenient for

³ Published by Pro Publica at projects.propublica.org/emergency/ It uses annual wait time data gathered by Medicare

Time to see a professional	% of visits
Fewer than 15 minutes	57%
15-45 minutes	36%
45 minutes or more	6%

Table 3 Wait time at urgent care centers

Source: Urgent Care Association of America Foundation <http://www.urgentcarecenter.org/press.html>

ER Wait Watcher

Which Emergency Room Will See You the Fastest?



By Lena Groeger, Mike Tigas and Sisi Wei, ProPublica, Dec. 19, 2014. Updated: Jan. 14, 2015

Heading to the emergency room? You may wait a while before a doctor or other treating professional sees you — and the hospital nearest to you might not be the one that sees you the fastest. Use this tool to look up average ER wait times, as reported by hospitals to the federal government, as well as the time it takes to get there in current traffic, as reported by Google. [Find out more >](#)

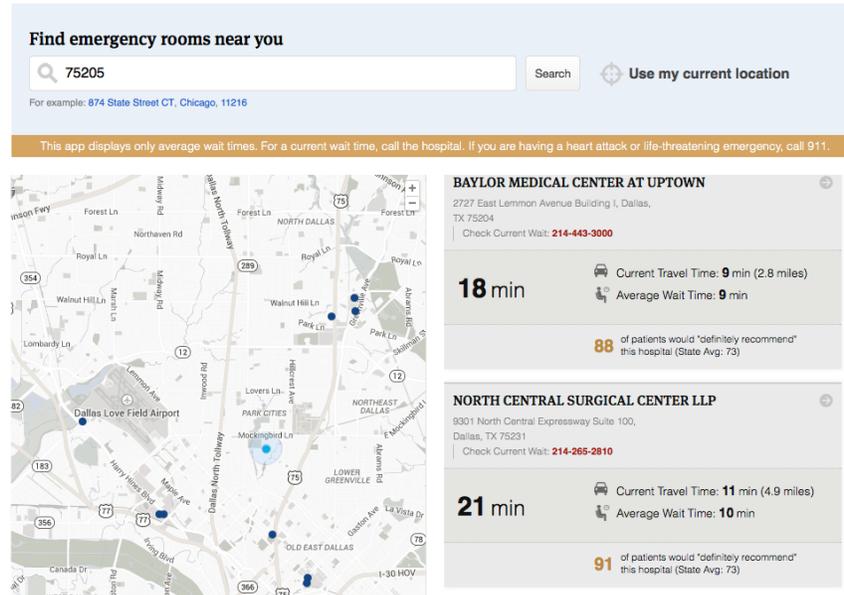


Figure 2 Wait and travel times to near by Emergency Departments provided by *ER Wait Watcher*

<http://projects.propublica.org/emergency/>

patients. The most common location for an UC is a shopping strip or a mall⁴.

The co-payments and fees charged by UC are typically much lower than that charged by an ED. *Consumer Reports* in their April 2009 Health Report [CR(2009)] reported that the co-pays for UC are comparable to that for a doctor visit, about \$120 on average, and much lower than at an ED, \$400 on average.

While UCs are more convenient and cheaper overall for non-emergent care, EDs still get non-emergent patients for a variety of reasons. Ability to pay is one reason. UCs only accept patients who have insurance or are willing to pay themselves. They do not accept medically indigent patients. These patients go to an ED which is required by law to accept them⁵. Another reason is availability. UC are typically open during business hours from 8AM to 10PM, whereas the ED is open 24 hours a day. Yet another reason is mis-classification by the patient who thinks that she has an emergent condition but actually does not. We consider the quality of self-classification in greater

⁴ According to a recent survey by the Urgent Care Association of America (UCAOA), the largest proportion, 38% of Urgent care centers are located in a shopping center or strip mall.

⁵ In 1986 the US enacted a law 'Emergency Medical Treatment and Active Labor Act (EMTALA)' that requires hospitals accepting medicare payments to provide emergency care treatment irrespective of ability to pay or legal status.

detail below.

So how does a patient choose an acute care facility? From an awareness standpoint, health insurance companies have a strong financial interest in making their insured patient know about the UC option because it is much cheaper for them. Using the 2008 Medical Expenditures Panel Survey (MEPS), the Government Accountability Office reports that the average charge for a non-emergency visit to an ED is \$2,101 versus \$203 at a health care center [GAO(2011)]. Many of the health insurance companies maintain websites and use other channels including mailings to inform the insureds of the UC option. Additionally, many potential patients go by Urgent Care centers in their usual day to day activities since many UC are located in malls, and other popular places.

The hard part in picking between an ED and an UC is deciding whether the patient's condition is emergent or not since UC do not offer care for patients with emergent conditions (Emergency Severity Index 1 to 3). This is because the judgement is made by the patient or her family who typically are not medically trained. Again, most insurance companies provide guidelines as a part of their Consumer health information (CHI) initiative. An example is provided in Figure 3. Most insurance companies have similar CHI.

Urgent Care Guidelines

Urgent Care treatment is best for these types of injuries:

- **Head:** No loss of consciousness, cuts less than one inch
- **Ears:** Earaches, infection, foreign object, severe dizziness or drainage
- **Eyes:** Scrapes, bruises, infections or a sty, objects in the eye, swelling around the eye
- **Nose:** Infection
- **Throat:** Sore throat
- **Chest:** Cough with or without fever, moderate asthma
- **Abdomen:** Persistent nausea and/or vomiting
- **Genital/Urinary:** Frequent trips to the bathroom, burning/pain with urination, vaginal/ penis discharge, bleeding or discomfort with intercourse
- **Back:** Minor strains or backache
- **Limbs/Skin:** Sprains without deformity, shallow or short cuts, stitch removal, puncture wounds to hands or feet, minor scrapes or burns, rash, insect or animal bites

Why you should choose Urgent Care instead of the Emergency Center for the listed conditions:

- You may have to wait longer in the Emergency Center, as the most serious injuries/cases are seen first.
- Co-pays are often more expensive in the Emergency Center than at Urgent Care.
- Your health plan may not pay for care if it is not a true emergency.

Figure 3 A simple example for CHI

Source: Mercy Health Cadillac <http://www.mercyhealthcadillac.com/urgent-care-guidelines>

Self-triage into emergent and non-emergent conditions by the patient or their family is fraught with costly errors. Patients who go to an ED with non-emergent conditions will have longer waits and much larger co-pays. Indeed, Medicare, a US health insurance system for people over the age of 65, allows an ED to charge a larger copay for non-emergent patients than for those with emergent conditions. Patients who go to an UC with emergent conditions also have serious consequences that include delay of care for a serious life threatening condition, extra co-pays, and delay from

additional travel. In this paper we model the self-triage activity as a process of classification with errors of both kinds. We consider a strategic patient who weighs all the costs, including waiting, co-pays, and the that of error in self-triage, in deciding whether to visit a UC or an ED.

2. Literature Review

The effect of patients with different priorities for treatment on waiting time and capacity has been studied from the 1970s in the context of elective and acute care [Shonick and Jackson(1973), Fries(1976)]. This situation is similar to our setting in which critical patients have priority over non-critical patients. They find the the proportion of low and high priority mix is very important determinant of hospital performance.

More recently, Dobson et al [Dobson et al.(2011)Dobson, Hasija, and Pinker] study the capacity reservation for urgent patients in primary care when both urgent and routine (nonurgent) patients exist. Their use of urgent though means that the patient should be seen in the same day rather than in a very short time unlike our definition for critical patients. They suggest that *for some cases the patient's symptoms would lead the physician to categorize him as urgent. In other cases, it is more the patient's own sense of urgency that drives the categorization.* According to their definition, some urgent patient's cannot be seen and they are considered as urgent overflow patients are either seen in overtime or sent to another facility or physician. Our paper specifically focuses on the patients who either deemed choose to go to an acute care facility because they find these facilities more convenient or because they do not want to wait for a delayed appointment.

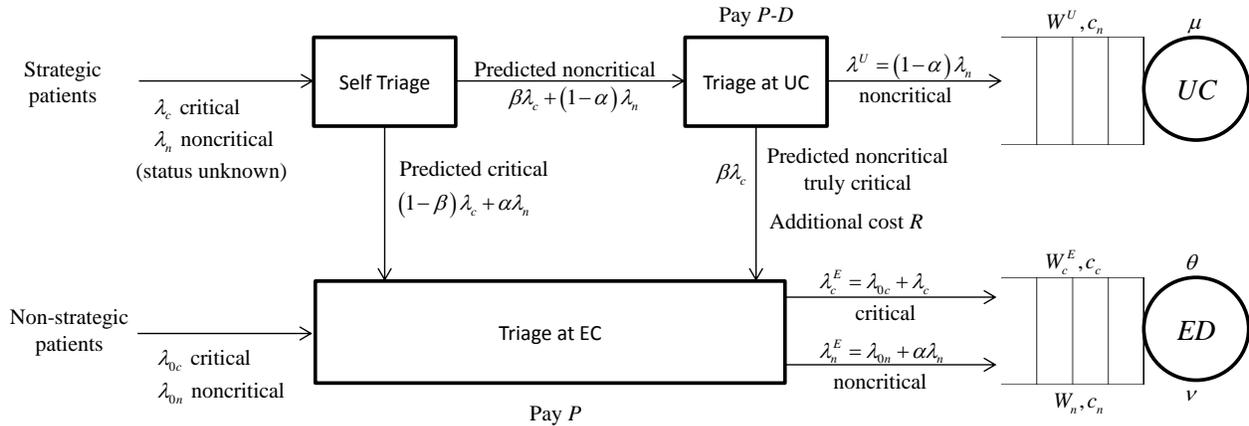
Patients, armed with information from the web, are increasing involved in making healthcare decisions [Baker et al.(2003)Baker, Wagner, Singer, and Bundorf]. There is empirical evidence that the patients strategically weigh the different factors such as quality of care, waiting times, co-payments, in deciding what kind of care to get [Nyman(1989), Padgett and Brodsky(1992)]. To these factors we add the cost of errors from self-classification.

3. The Model

3.1. Preliminaries

Patients may or may not have a choice of acute care facility, an Emergency Department (ED) at a hospital or an Urgent Care Center (UC), to visit. We label the patients who have the choice as *strategic* and those who do not as *non-strategic*. Though we include the nonstrategic patients in the workflow, we focus on strategic patients in the rest of this paper.

Some patients in either group may have an emergent condition. We call them *critical* patients. These patients need the full spectrum care offered by an ED. Others, who do not have an emergent



Notation:

λ_c	Arrival rate of strategic patients who are truly critical
λ_n	Arrival rate of strategic patients who are truly non-critical
λ_{0c}	Arrival rate of non-strategic patients who are truly critical
λ_{0n}	Arrival rate of non-strategic patients who are truly non-critical
α	The error rate of patients who are classified critical but are truly non-critical
β	The error rate of patients who are classified non-critical but are truly critical
P	Co-payment for ED services
D	Discount in co-payment for UC services
$P - D$	Co-payment net of discount for UC services
μ	The service rate at UC
θ	The service rate for critical patients at ED
ν	The service rate for non-critical patients at ED
W^U	The expected time in system for patients at UC
W_c^E	The expected time in system for critical patients at ED
W_n	The expected time in system for non-critical patients at ED
c_c	Waiting cost rate for critical patients
c_n	Waiting cost rate for non-critical patients

Figure 4 Preliminaries of the Model

condition are labeled as *noncritical*. The care-map is shown in Figure 4. Let λ_c and λ_n be the arrival rates of strategic patients who are truly critical and noncritical, respectively. A strategic patient’s choice starts with self-triage or classification based on perceived symptoms and their prior medical knowledge including that obtained from consumer health information (CHI) systems.

The self-triage or classification done by a non-medical professional is fraught with errors. On one hand, patient who is truly critical, such as a heart attack patient thinking that the angina is upset stomach, may self-classify as noncritical. On the other hand, patient who is truly noncritical, such as a patient with an upset stomach thinks that she has appendicitis, may self classify as critical. In either case, the patient is not fully informed or trained and may make an error. As we see from

the examples above, two types of errors are possible: (1) A truly noncritical patient predicts that she is critical (Type I error) or (2) a truly critical patient predicts that she is noncritical (Type II error)⁶. We define α and β as the error rates for Type I and Type II errors, respectively. Without loss of generality we take $\alpha + \beta \leq 1$ for if this is not the case then the labels in the guidelines can be switched to make it so. The confusion matrix, in Table 4 exhibits the disposition of the patients between true and predicted classes.

	Truly Critical	Truly Noncritical
Predicts Critical	$(1 - \beta)\lambda_c$	$\alpha\lambda_n$
Predicts Noncritical	$\beta\lambda_c$	$(1 - \alpha)\lambda_n$

Table 4 Confusion matrix (normalized by total number of strategic patients)

In this paper we explore a separating equilibrium in which strategic patients who self-triage as noncritical choose to go to UC instead of the ED. However, because of error in self triage, only a $(1 - \alpha)$ fraction of non-critical patients go the UC. The remainder, $\alpha\lambda_n$, go to the ED thinking they are critical. Similarly, $(1 - \beta)\lambda_c$ of the critical patients got to the ED directly, and $\beta\lambda_c$ erroneously thinking that they are non-critical go first to the UC and then are re-routed back to the ED.

Including the nonstrategic patient flows of λ_{0c} and λ_{0n} for critical and noncritical condition, the total net flows into the different facilities are:

$$\text{Net flow rate of critical patients into ED: } \lambda_c^E = \lambda_{0c} + (1 - \beta)\lambda_c + \beta\lambda_c = \lambda_{0c} + \lambda_c \quad (1)$$

$$\text{Net flow rate of noncritical patients into ED: } \lambda_n^E = \lambda_{0n} + \alpha\lambda_n \quad (2)$$

$$\text{Net flow rate of noncritical patients into UC: } \lambda^U = (1 - \alpha)\lambda_n. \quad (3)$$

We also define service rates in both acute care facilities. We define μ as the service rate for noncritical patients in UC, and θ and ν as the service rate for critical and noncritical patients in ED, respectively. Further, we assume that $\nu > \theta$, i.e., it takes longer to provide healthcare service to a critical patient than a non-critical patient.

3.2. Expected Cost for the Patients

Patients incur a variety of costs in using the medical facility. Some costs, such as co-pays, are pecuniary while others, such as those associated with waiting and transfer, and effects on outcome because of delays, are non-pecuniary but important to consider. At the time of arrival patients

⁶ In statistics Type I and Type II errors refer to the false positive and the false negative errors, respectively. We treat the critical class as the positive class for purposes of assigning Type I and Type II error names.

incur co-pay P in ED and $P - D$ in UC where D is the discount offered by UC to patients. Patients who choose to go to UC and triaged critical incurs the rerouting cost R from UC to ED. These patients incur both co-pays.

Patients who arrive at a facility may wait before they are treated, and treatment itself may take a while. We define the expected system times as W_c^E and W_n^E for critical and noncritical patients in ED, respectively and W^U for noncritical patients in UC given that predicted noncritical patients go to UC. We also assign patient waiting cost rate c_c and c_n for critical and noncritical patients, respectively. Putting these together we determine the expected disutility that are shown in the Table 5 below.

	Critical	Noncritical
Go to ED	$P + c_c W_c^E$	$P + c_n W_n^E$
Go to UC	$2P - D + R + c_c W_c^E$	$P - D + c_n W^U$

Table 5 Expected Cost incurred by a Strategic Patient

4. Separating Equilibrium for Patient Choice

A separating equilibrium, in which strategic patients who are predicted to be noncritical use the UC, is characterized next. Let C and NC be events where patient is truly critical and non critical, respectively and C' and NC' be the events that they are predicted critical and predicted noncritical, respectively. Let π denote the precision of the non-critical classification, i.e., $\pi = P[NC|NC']$ and π' denote the precision of the critical classification, i.e., $\pi' = P[C|C']$.

THEOREM 1. *A separating equilibrium, in which predicted noncritical patients go to UC and predicted critical patients go to ED, is obtained if and only if*

$$(1 - \pi)(P + R) - \pi c_n (W_n^E - W^U) \leq D \leq \pi'(P + R) - (1 - \pi')c_n (W_n^E - W^U). \quad (4)$$

Further if $\alpha + \beta \leq 1$ and $W_n - W^U \geq 0$, then there exists a D that satisfies both inequalities in (4).

Examining the confusion matrix in Table 4, we get the precisions referred to in the theorem:

$$\pi = \frac{(1 - \alpha)(1 - \chi)}{\beta\chi + (1 - \alpha)(1 - \chi)} \text{ and } \pi' = \frac{(1 - \beta)\chi}{(1 - \beta)\chi + \alpha(1 - \chi)} \quad (5)$$

where $\chi = \frac{\lambda_c}{\lambda_c + \lambda_n}$ is the prior probability of critical condition for strategic patients.

Intuitively, the conditions in Theorem 1 requires that the expected cost of picking a UC over an ED be less than the reduction in co-pays, which is D . The first term is the expected penalty from

rerouting the patients who are predicted noncritical but are truly critical. The second term is the benefit arising from a reduction in expected waiting cost of going to a UC over an ED. The details of the derivation are in the proof in the Appendix.

Any discount D that satisfies the inequalities in Theorem 1 will result in a separating equilibrium. However, the UC administration prefers lower discounts. Hence, the lower bound is binding for undominated separating equilibrium.

COROLLARY 1. *The lowest discount, D^* , offered by a profit maximizing UC is*

$$D^* = (1 - \pi)(P + R) - \pi c_n (W_n^E - W^U) \quad (6)$$

5. Analysis

For tractability and care of exposition, we now assume that the expected system times at the ED arise from an M/M/1 queuing discipline with non-preemptive priority for critical patients. Following [Kleinrock(1976)] and using the arrival and service rates defined in section 3.1, the expected system times are exhibited below.

$$W_c^E = \frac{\frac{\lambda_c^E}{\theta^2} + \frac{\lambda_n^E}{\nu^2}}{1 - \frac{\lambda_c^E}{\theta}} + \frac{1}{\theta} = \frac{\frac{\lambda_{0c} + \lambda_c}{\theta^2} + \frac{\lambda_{0n} + \alpha\lambda_n}{\nu^2}}{1 - \frac{\lambda_{0c} + \lambda_c}{\theta}} + \frac{1}{\theta} \quad (7)$$

$$W_n^E = \frac{\frac{\lambda_c^E}{\theta^2} + \frac{\lambda_n^E}{\nu^2}}{\left(1 - \frac{\lambda_c^E}{\theta} - \frac{\lambda_n^E}{\nu}\right) \left(1 - \frac{\lambda_c^E}{\theta}\right)} + \frac{1}{\nu} = \frac{\frac{\lambda_{0c} + \lambda_c}{\theta^2} + \frac{\lambda_{0n} + \alpha\lambda_n}{\nu^2}}{\left(1 - \frac{\lambda_{0c} + \lambda_c}{\theta} - \frac{\lambda_{0n} + \alpha\lambda_n}{\nu}\right) \left(1 - \frac{\lambda_{0c} + \lambda_c}{\theta}\right)} + \frac{1}{\nu} \quad (8)$$

Similarly, we assume an M/M/1 queuing discipline for the UC. Then the expected (system) waiting time in UC is exhibited below.

$$W^U = \frac{\frac{1}{\mu}}{1 - \frac{\lambda_n^U}{\mu}} = \frac{\frac{1}{\mu}}{1 - \frac{(1-\alpha)\lambda_n}{\mu}} = \frac{1}{\mu - (1-\alpha)\lambda_n} \quad (9)$$

We also assume that the ED is more congested than UC, i.e., $W_n \geq W^U$.

5.1. Effect of Error Rates

Type I error, in which a non-critical patient mistakenly classifies herself as critical, has two effects: one is shifting of traffic from UC to ED and the other is reducing the precision of patients who are classified as noncritical. We will examine each of these in turn.

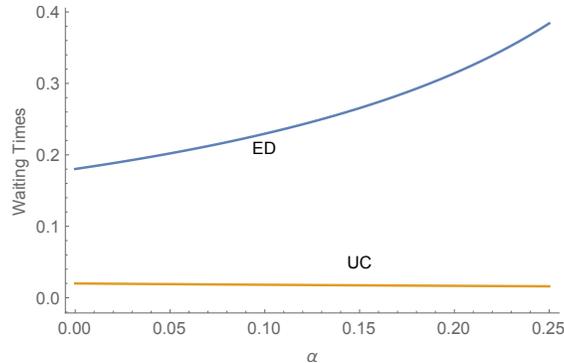
The flow of strategic noncritical patients is split between the ED at rate $\alpha\lambda_n$ and UC at the rate $(1 - \alpha)\lambda_n$. As the Type I error rate α increases, the ED gets more congested, UC gets less congested and the waiting time difference between the two increases. This is shown in Figure 5.

The second effect arises from reducing the precision of noncritical self-classification as the Type I error rate increases, i.e., $\pi = \frac{(1-\alpha)(1-\chi)}{\beta\chi + (1-\alpha)(1-\chi)}$ is decreasing in α . So a patient who finds herself at a

UC is less sure that it is the correct facility and has a higher likelihood of being rerouted to an ED, pay an additional co-pay at P and rerouting cost R .

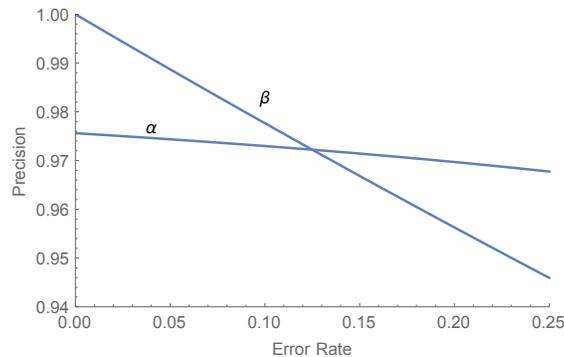
While a change in Type II error rate β does not have a direct impact on the traffic flow, it does reduce the precision of noncritical classification. The effect of α and β on the precision of noncritical classification is shown in Figure 6.

Figure 5 Expected System Times in ED and UC vs. Type I Error Rate



Parameters values: $\theta = 30$, $\nu = 120$, $\mu = 100$, $\lambda_{0c} = 4$, $\lambda_{0n} = 40$, $\lambda_c = 10$, $\lambda_n = 50$

Figure 6 Precision, $\Pr[NC|NC']$, vs. Type I and II Error Rates



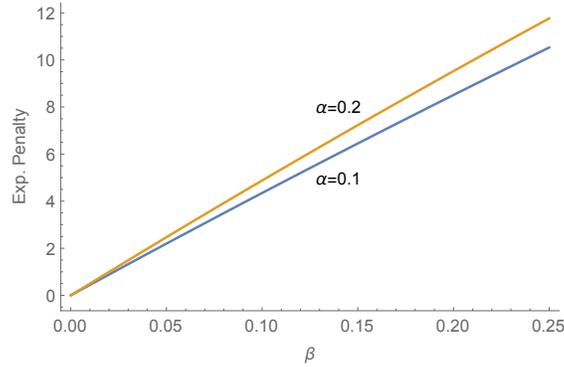
Parameters values: Type I α varying, $\beta=0.125$; Type II β varying, $\alpha=0.125$, $\chi=0.5$.

5.2. Short Run Analysis of Discount

In the short run, we take the capacity as fixed and focus our attention on the discount in co-pay offered by the UC, which is a key decision variable for the manager of a UC.

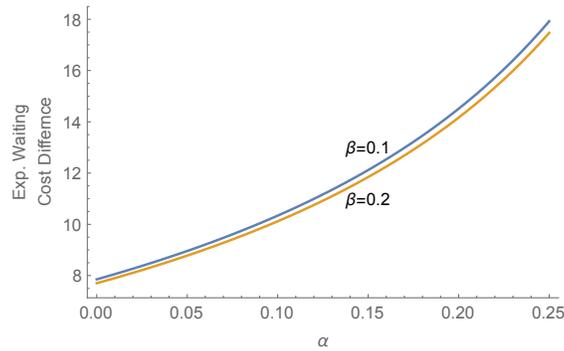
The discount, shown in (6) is set to compensate the patient for two differences between experience of the patient in the ED versus UC:

- (1) The potential expected penalty from rerouting critical patients who are incorrectly classified as non-critical and go to UC, i.e., $(1 - \pi)(P + R)$. This is strongly affected by the Type II error rate β as illustrated in Figure 7.

Figure 7 Expected Penalty vs. Type I and II Error Rates

Parameters values: $\theta = 30$, $\nu = 120$, $\mu = 100$, $\lambda_{0c} = 4$, $\lambda_{0n} = 40$, $\lambda_c = 10$, $\lambda_n = 50$, $P=100$, $R=100$

(2) The expected reduction in waiting cost from going to the UC instead of ED, i.e., $\pi(W_n - W^U)$. This is strongly affected by Type I error rate α . As we noted before, increasing α results in a shift of traffic from less congested UC to more congested ED, making the expected difference in waiting cost larger. This is shown in Figure 8.

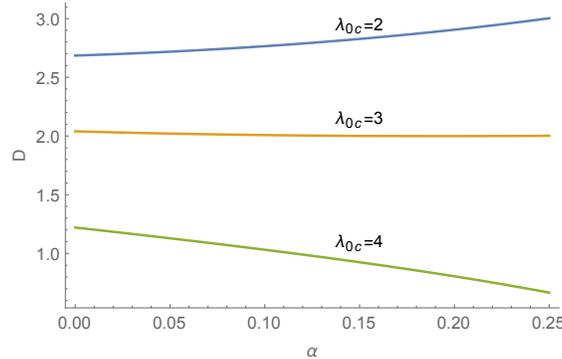
Figure 8 Expected System Time Cost Difference vs. Type I and II Error Rates

Parameters values: $\theta = 30$, $\nu = 120$, $\mu = 100$, $\lambda_{0c} = 4$, $\lambda_{0n} = 40$, $\lambda_c = 10$, $\lambda_n = 50$, $c_n=50$

The effect of these parameters on the discount is explained next.

The effect of Type I error rate α depends very much on the level of congestion in the ED. In the numerical experiments exhibited in Figure 9, we adjusted the congestion of the ED by varying the flow of nonstrategic critical patients, λ_{0c} . When $\lambda_{0c}=2$ and ED is less congested, an increase in α results in an increase in the discount through the penalty term. When $\lambda_{0c}=4$, i.e., when the ED is much more congested, an increase in α results in a decrease in the discount as the expected difference in waiting cost term dominates. Although the discount is decreasing with α when $\lambda_{0c}=4$, this is not a desirable scenario for the patients or the healthcare system as the ED is very congested and the Type I error is making it worse.

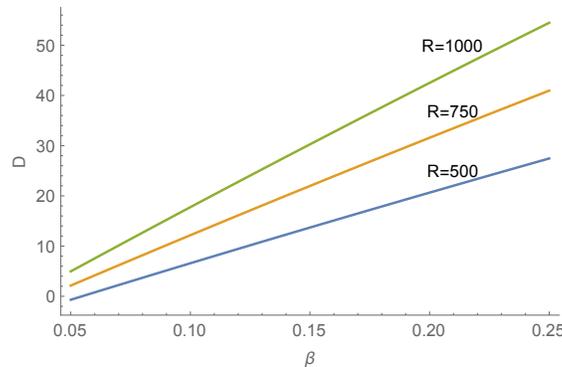
Figure 9 Discount vs. Type I Error Rate and ED Congestion



Parameters values: $\theta = 30$, $\nu = 120$, $\mu = 100$, $\lambda_{0n} = 20$, $\lambda_c = 10$, $\lambda_n = 50$, $c_n=50$, $P=100$, $R=100$, $\beta=0.125$

The effect of Type II error rate β is shown in Figure 10. The main effect of this error on the discount is via the penalty term, which is increasing in β and the rerouting cost R . So to achieve a lower discount UC administration can either focus on reducing the Type II error rate or the rerouting cost. Rerouting cost can be largely reduced by choosing the location of UC closer to an ED. Nevertheless, UCs create several service dimensions for predicted noncritical but truly critical patients who wrongly ended up in UC. For instance, UC finishes the triage eliminating the need for another triage in ED, calls ED and assigns the required doctors to the patient and also transport the patient to ED. On the other hand, UC can reduce Type II error rate by increasing the quality of CHI. *references are needed for reducing R and β from real life.*

Figure 10 Discount vs. Type II Error Rate and Rerouting Cost



Parameters values: $\theta = 30$, $\nu = 120$, $\mu = 100$, $\lambda_{0c} = 3$, $\lambda_{0n} = 40$, $\lambda_c = 10$, $\lambda_n = 50$, $c_n=50$, $P=150$, $\alpha=0.125$

These results are formally presented below in Proposition 1.

PROPOSITION 1. *The discount is increasing in R and β . If the ED is not too congested so that*

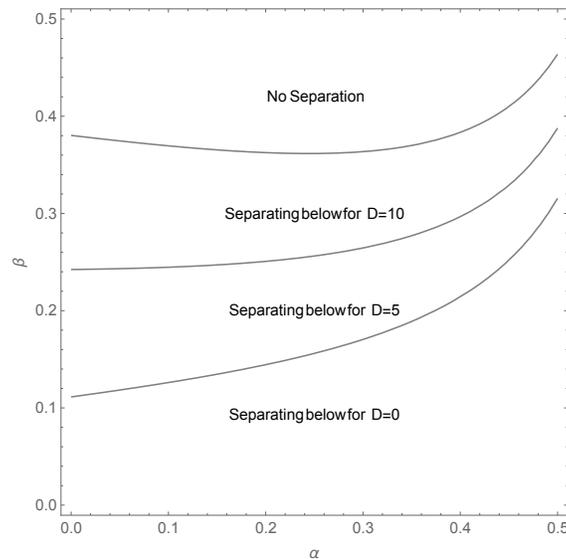
$$\frac{\rho_n}{(1 - \rho_n - \rho_c)}(W_n^E + W_c^E) \leq \frac{\beta\chi}{(1 - \alpha)^2}(P + R) - \lambda_n W^U{}^2 \quad (10)$$

then the discount is also increasing α .

Recall that the optimal discount is expected penalty from rerouting minus the savings in waiting time cost in picking UC over ED. An increase in R or β increases the expected penalty and thus increases the discount. The effect of α is more complicated. In general an increase in α , which shifts traffic from the UC to the ED, results in an increase of wait time at the ED wait time and a decrease in the wait time at the UC. However, the expected penalty from rerouting increases with reduced precision resulting from increased α . These two effects are competing. When the ED is not as congested, the effect of α on expected penalty is greater than that on changes in expected wait times and so the discount is increasing in α . However, this changes when the ED is very congested. In this case the expected waiting time cost overwhelms the change in the expected penalty and the discount may well decrease with increasing α .

5.2.1. Analysis of Separating Equilibrium We illustrate the change in the regions in which separating equilibrium exists for different discount values as a function of error rates in Figure 11. As expected, as discount increases, separating equilibrium exists for a larger set of error rate values. This is because as the discount increases, patients are more incentivized to go to UC.

Figure 11 Separating Equilibrium vs. Error Rates and Discount



Parameters values: $\theta = 30$, $\nu = 120$, $\mu = 100$, $\lambda_{0c} = 2$, $\lambda_{0n} = 40$, $\lambda_c = 10$, $\lambda_n = 50$, $c_n = 50$, $P=150$, $R = 50$

As Type II error rate increases the discount required to satisfy a separating equilibrium increases as shown in Proposition 1(b). The discount decreases in Type I error when ED is congested enough to satisfy the conditions in Proposition 1(c). As Type II error rate increases these conditions are satisfied hard as ED gets less congested. Hence, an increase in Type I error would enlarge the set of Type II error values that can satisfy the separating equilibrium at the same low discount level.

As discount increases the opposite starts to occur. In overall, unless ED has heavy traffic Type I error does not have a very high impact on the separating equilibrium as Type II error does.

We next examine the impact of insurance payments, fixed and variable costs and capacity choice on profitability in section 6.

6. Long Run Analysis

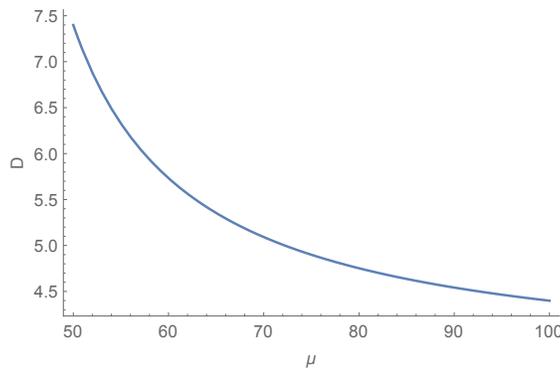
In the long run, the capacity of UC, μ , is a key decision variable. Let $D^*(\mu)$ and $W^U(\mu)$ be the optimal short run discount and the waiting time at UC for any capacity μ as in (6) and (9). In this section, we characterize the optimal capacity choice of UC to maximize its expected profit.

Let k be the capacity cost rate at the UC. This includes the facilities and consumables needed to treat patients. Let I be average payment from the insurance provider received by the UC for each noncritical patient treated by the UC. We assume that this is independent of the error rates α and β . Putting these together, the revenue per patient is $P - D^*(\mu) + I$.

As we see in Figure 12, the discount is decreasing in UC capacity. This sets up a trade-off between the revenue and the cost of capacity. The UC manager picks the profit maximizing capacity shown below:

$$\max_{\{\mu \geq 0, W_n \geq W^U(\mu)\}} [(P - D^*(\mu) + I)\lambda_n(1 - \alpha) - k\mu] \quad (11)$$

Figure 12 Discount vs. UC Capacity



Parameters values: $\theta = 30$, $\nu = 120$, $\mu = 100$, $\lambda_{0c} = 3$, $\lambda_{0n} = 40$, $\lambda_c = 10$, $\lambda_n = 50$, $c_n = 50$, $P=150$, $R = 100$, $\alpha = \beta=0.125$

The optimal capacity choice and the maximum profit for UC are formally presented in Proposition 2 below.

PROPOSITION 2. *If $k < c_n \lambda_n (1 - \alpha)(1 - \chi) W_n^2$, then a profit maximizing UC administration sets the optimal capacity at μ^* as:*

$$\mu^* = (1 - \alpha)\lambda_n \left(1 + \sqrt{\frac{\pi c_n}{(1 - \alpha)\lambda_n k}} \right) \quad (12)$$

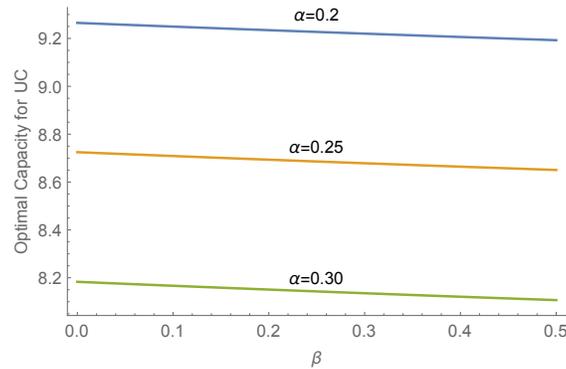
and in return UC makes the following maximum profit:

$$profit^* = (1 - \alpha)\lambda_n \left((P - D^*(\mu^*) + I - k) - \sqrt{\frac{\pi k c_n}{(1 - \alpha)\lambda_n}} \right). \quad (13)$$

6.1. Effect of Error Rates on the Optimal Capacity

As both error rates increase, precision (probability of noncritical patients conditioned on predicted noncritical patients) decreases. Moreover, an increase in Type I error rate decreases the net flow rate of noncritical patient to UC, i.e., viz. (3). Both impacts lead to a reduced expected waiting time of a patient so UC can pick a lower capacity as illustrated in Figure 13. We formally present these results in Proposition 3.

Figure 13 Optimal Capacity for UC vs. Error Rates



Parameters values: $\lambda_n = 10$, $\lambda_c = 2$, $c_n = 5$, $k = 25$, $\beta = 0.125$;

PROPOSITION 3. *The optimal capacity, μ^* , is decreasing in α and β .*

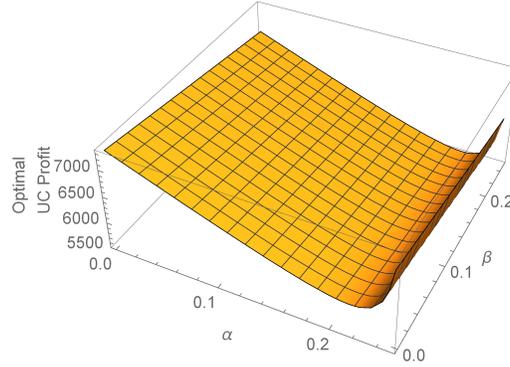
6.2. Error Rates and the Optimal UC Profit

The effect of error rates on UC profit depends on the level of congestion in the ED. In our analysis, we have assumed that the ED is more congested than the UC, i.e., $W_n > W^U$. If the congestion level is not too extreme, the limit is characterized in the proposition below, then the UC manager has an incentive to minimize both types of error. The situation is different if ED is extremely congested. This can set up an adverse situation in which increasing the Type I error rate to further congest the ED may result in a such a large reduction in the discount that it more than makes up for revenue lost from misclassification. In this case, the UC manager may well prefer a higher Type I error. But this comes at the cost of extreme congestion in the ED.

These situations are formally presented below in Proposition 4.

PROPOSITION 4. *If the congestion condition (10) in Proposition 1 is met and the UC is making a positive profit then the optimal profit is decreasing in α and β .*

Figure 14 Optimal Profit vs. Error Rates



Parameters values: $\theta = 30$, $\nu = 110$, $\lambda_{0c} = 2$, $\lambda_{0n} = 40$, $\lambda_c = 10$, $\lambda_n = 100$, $c_n = 6$, $P=100$, $R=100$, $I = 0$, $k=25$.

Appendix. Proof of Theorems and Propositions

Proof of Theorem 1 : There are two cases to analyze to prove that separating equilibrium exists.

Case 1: When all predicted noncritical patients go to UC, no single patient would want to change his decision and go to ED. We define EC^{1U} as expected cost of predicted noncritical patients in UC and EC_s^{1E} as the expected cost of a single predicted noncritical patient who changes his decision against others and goes to EC instead of UC, then we have

$$\begin{aligned} EC^{1U} &= \Pr(C|NC')(2P - D + c_c W_c^E + R) + \Pr(NC|NC')(P - D + c_n W_n^U), \\ EC_s^{1E} &= \Pr(C|NC')(P + c_c W_c^E) + \Pr(NC|NC')(P + c_n W_n^E). \end{aligned}$$

If $EC^{1U} \leq EC_s^{1E}$ then that all predicted noncritical patients go to UC is a Nash Equilibrium. Then we must have

$$\Pr(C|NC')(P + R) - \Pr(NC|NC')c_n (W_n^E - W_n^U) \leq D.$$

Case 2: When all predicted critical patients go to ED, no single patient would want to change his decision and go to UC. We define EC^{1E} as expected cost of predicted critical patients in ED and EC_s^{1U} as the expected cost of a single predicted critical patient who changes his decision against others and goes to UC instead of ED, then we have

$$\begin{aligned} EC^{1E} &= \Pr(C|C')(P + c_c W_c^E) + \Pr(NC|C')(P + c_n W_n^E) \\ EC_s^{1U} &= \Pr(C|C')(2P - D + c_c W_c^E + R) + \Pr(NC|C')(P - D + c_n W_n^U). \end{aligned}$$

If $EC^{1E} \leq EC_s^{1U}$ then that all predicted critical patients go to UC is a Nash Equilibrium. Then we must have

$$D \leq \Pr(C|C')(P + R) - \Pr(NC|C')c_n (W_n^E - W_n^U).$$

So including both cases we should have

$$P[C|NC'](P + R) - P[NC|NC']c_n (W_n^E - W_n^U) \leq D \leq P[C|C'](P + R) - P[NC|C']c_n (W_n^E - W_n^U). \quad (14)$$

Further if $W_n^E \geq W^U$, the upper bound is higher than the lower bound as long as $\Pr(NC|NC') \geq \Pr(NC|C')$ and $\Pr(C|C') \geq \Pr(C|NC')$ hold. But these two conditions are equivalent. So, we must have

$$\Pr[C|C'] \geq \Pr[C|NC'] \Leftrightarrow \frac{(1-\beta)\chi}{(1-\beta)\chi + \alpha(1-\chi)} \geq \frac{\beta\chi}{\beta\chi + (1-\alpha)(1-\chi)} \Leftrightarrow (1-\alpha-\beta)(1-\chi) \geq 0.$$

Adding the assumption that $\alpha + \beta \leq 1$, there exists a D that satisfies the separating equilibrium.

Proof of Proposition 1 : Define $\Delta W = W_n - W^U$. Note that we have assumed that $\Delta W > 0$.

$$\frac{dD}{dR} = \frac{d}{dR} \left(\frac{\beta\chi}{\beta\chi + (1-\alpha)(1-\chi)}(P+R) - \frac{(1-\alpha)(1-\chi)}{\beta\chi + (1-\alpha)(1-\chi)}c_n\Delta W \right) = \frac{\beta\chi}{\beta\chi + (1-\alpha)(1-\chi)} \geq 0.$$

Hence, the discount is increasing in R .

$$\frac{dD}{d\beta} = \frac{d}{d\beta} \left(\frac{\beta\chi}{\beta\chi + (1-\alpha)(1-\chi)}(P+R) - \frac{(1-\alpha)(1-\chi)}{\beta\chi + (1-\alpha)(1-\chi)}c_n\Delta W \right) = \frac{(1-\alpha)(1-\chi)\chi}{(\beta\chi + (1-\alpha)(1-\chi))^2}(P+R+c_n\Delta W) \geq 0.$$

Hence, the discount is increasing in β .

$$\begin{aligned} D_\alpha &= \frac{d}{d\alpha} \left(\frac{\beta\chi}{\beta\chi + (1-\alpha)(1-\chi)}(P+R) - \frac{(1-\alpha)(1-\chi)}{\beta\chi + (1-\alpha)(1-\chi)}c_n\Delta W \right) \\ &= \frac{(1-\chi)}{\beta\chi + (1-\alpha)(1-\chi)} \left(\frac{\beta\chi(P+R+c_n\Delta W)}{\beta\chi + (1-\alpha)(1-\chi)} - (1-\alpha)c_n\Delta W_\alpha \right) \\ &= \frac{(1-\chi)}{\beta\chi + (1-\alpha)(1-\chi)} (P[C|NC'](P+R+c_n\Delta W) - (1-\alpha)c_n\Delta W_\alpha) \end{aligned}$$

Hence we note that D_α is non-negative if:

$$P[C|NC'](P+R+c_n\Delta W) - (1-\alpha)c_n\Delta W_\alpha \geq 0 \quad (15)$$

Now we will look at ΔW_α more closely. Noting that $\lambda_n^E = \lambda_{0n} + \alpha\lambda_n$ and $\lambda^U = (1-\alpha)\lambda_n$, we get:

$$\frac{d\Delta W}{d\alpha} = \frac{dW_n}{d\lambda_n^E} \frac{d\lambda_n^E}{d\alpha} - \frac{dW^U}{d\lambda_n^U} \frac{d\lambda_n^U}{d\alpha} = \lambda_n \left(\frac{dW_n}{d\lambda_n^E} + \frac{dW^U}{d\lambda^U} \right). \quad (16)$$

$$\frac{dW^U}{d\lambda^U} = \frac{1}{(\mu - (1-\alpha)\lambda^U)^2} = W^{U2} \quad (17)$$

$$\begin{aligned} \frac{dW_n^E}{d\lambda_n^E} &= \frac{d}{d\lambda_n^E} \left[\frac{\rho_c/\theta + \lambda_n^E/\nu^2}{(1-\rho_c - \lambda_n^E/\nu)(1-\rho_c)} + 1/\nu \right] \\ &= \frac{(1-\rho_c - \rho_n)/\nu + (\rho_c/\theta + \rho_n/\nu)}{\nu(1-\rho_c - \rho_n)^2(1-\rho_c)} \\ &= \frac{\frac{1}{\nu(1-\rho_c)} + W_n^E - 1/\nu}{\nu(1-\rho_c - \rho_n)} \\ &= \frac{\frac{1}{\nu} \frac{\rho_c}{1-\rho_c} + W_n^E}{\nu(1-\rho_c - \rho_n)} \end{aligned}$$

$$\begin{aligned} &\leq \frac{\frac{1}{\theta} \frac{\rho_c}{1-\rho_c} + W_n^E}{\nu(1-\rho_c-\rho_n)} \text{ since } \theta < \nu \\ &\leq \frac{W_c^E + W_n^E}{\nu(1-\rho_c-\rho_n)} \end{aligned} \quad (18)$$

The last inequality arises from the fact that W_c^E is the expected waiting time for critical patients who have non-preemptive priority over non-critical patients and $\frac{1}{\theta} \frac{\rho_c}{1-\rho_c}$ is the expected waiting time for critical patients when there are no non-critical patients.

Substituting (17) and (18) into (16) we get:

$$\frac{d\Delta W}{d\alpha} \leq \frac{W_c^E + W_n^E}{\nu(1-\rho_c-\rho_n)} + W^U{}^2 \quad (19)$$

From equations (15) and (19) we get that $D_\alpha \geq 0$ if:

$$\frac{\rho_n(W_n^E + W_c^E)}{1-\rho_n-\rho_c} \leq \frac{(P+R)\beta\chi}{(1-\alpha)^2} - \lambda_n W^U{}^2$$

Proof of Proposition 2 : Solving for the necessary condition of the maximization problem in (11) we have

$$\begin{aligned} \frac{d\pi}{d\mu} &= -\frac{dD}{d\mu} \lambda_n(1-\alpha) - k \\ &= -\frac{d}{d\mu} \left(\frac{\beta\chi}{\beta\chi + (1-\alpha)(1-\chi)} (P+R) - \frac{(1-\alpha)(1-\chi)}{\beta\chi + (1-\alpha)(1-\chi)} c_n(W_n - W^U) \right) \lambda_n(1-\alpha) - k \\ &= -\left(\frac{(1-\alpha)(1-\chi)}{\beta\chi + (1-\alpha)(1-\chi)} c_n \frac{dW^U}{d\mu} \right) \lambda_n(1-\alpha) - k = \frac{\lambda_n(1-\alpha)^2(1-\chi)}{\beta\chi + (1-\alpha)(1-\chi)} \frac{c_n}{(\mu - \lambda_n(1-\alpha))^2} - k = 0. \end{aligned}$$

Solving the above for μ gives us

$$\mu^* = (1-\alpha)\lambda_n + \sqrt{\frac{c_n\lambda_n(1-\alpha)^2(1-\chi)}{k(\beta\chi + (1-\alpha)(1-\chi))}}. \quad (20)$$

Note that $\mu^* \geq 0$. Also, if $k \leq c_n\lambda_n(1-\alpha)(1-\chi)W_n^2$ then since $\alpha + \beta \leq 1$ we have $k \leq \frac{c_n\lambda_n(1-\alpha)^2(1-\chi)W_n^2}{\beta\chi + (1-\alpha)(1-\chi)}$ which guarantees that $W_n - W^U \geq 0$.

Further, for part b, embedding μ^* into (11) the optimal profit, π^* becomes

$$\pi^* = (P - D^*(\mu^*) + I - k)\lambda_n(1-\alpha) - \sqrt{\frac{k c_n \lambda_n (1-\alpha)^2 (1-\chi)}{(\beta\chi + (1-\alpha)(1-\chi))}}. \quad (21)$$

Proof of Proposition 3 : Taking the derivative of (12) with respect to α and β respectively give us the results as follows:

$$\begin{aligned} \frac{d\mu^*}{d\alpha} &= \frac{d}{d\alpha} \left((1-\alpha)\lambda_n + \sqrt{\frac{c_n\lambda_n(1-\alpha)^2(1-\chi)}{k(\beta\chi + (1-\alpha)(1-\chi))}} \right) = -\lambda_n - \frac{\sqrt{\frac{c_n\lambda_n(1-\alpha)^2(1-\chi)}{k(\beta\chi + (1-\alpha)(1-\chi))}} (2\beta\chi + (1-\alpha)(1-\chi))}{2(1-\alpha)(\beta\chi + (1-\alpha)(1-\chi))} < 0, \\ \frac{d\mu^*}{d\beta} &= \frac{d}{d\beta} \left((1-\alpha)\lambda_n + \sqrt{\frac{c_n\lambda_n(1-\alpha)^2(1-\chi)}{k(\beta\chi + (1-\alpha)(1-\chi))}} \right) = -\frac{\chi \sqrt{\frac{c_n\lambda_n(1-\alpha)^2(1-\chi)}{k(\beta\chi + (1-\alpha)(1-\chi))}}}{2(\beta\chi + (1-\alpha)(1-\chi))} < 0. \end{aligned}$$

Proof of Theorem ?? : Under full information, all patients know their true types. So we have $P[C|NC'] = 0$, $P[NC|NC'] = 1$, $P[C|C'] = 1$, $P[NC|C'] = 0$ in (14). Then, we must have

$$-c_n(W_n - W^U) \leq D_F \leq P + R$$

Further if $W_n \geq W^U$, the upper bound is higher than the lower bound, so there exists a D_F that satisfies the separating equilibrium under full information.

Proof of Proposition 4 : We will prove the results in turn.

First let us look at the effect of α . We require that the optimal profit be positive. Using (13), this requires that

$$P - D + I - k > \sqrt{\frac{kc_n\pi}{(1-\alpha)\lambda_n}} \quad (22)$$

Utilizing the envelope theorem we get

$$\frac{d\text{profit}^*}{d\alpha} = -(1-\alpha)\lambda_n D_\alpha - (P - D + I - k)\lambda_n - \frac{1}{2}\sqrt{kc_n\lambda_n} \left(\frac{-\pi + (1-\alpha)\pi_\alpha}{\sqrt{(1-\alpha)\pi}} \right) \quad (23)$$

Note that

$$(1-\alpha)\pi_\alpha = (1-\alpha) \frac{-\beta(1-\chi)\chi}{(\beta\chi + (1-\alpha)(1-\chi))^2} = -\pi(1-\pi)$$

Substituting above into (23) we get

$$\begin{aligned} \frac{d\text{profit}^*}{d\alpha} &= -(1-\alpha)\lambda_n D_\alpha - \lambda_n \left((P - D + I - k) - \frac{2-\pi}{2} \sqrt{\frac{kc_n\pi}{(1-\alpha)\lambda_n}} \right) \\ &\leq -(1-\alpha)\lambda_n D_\alpha - \lambda_n \left((P - D + I - k) - \sqrt{\frac{kc_n\pi}{(1-\alpha)\lambda_n}} \right) \\ &\leq 0 \text{ (Using equation (22) and Proposition 1)} \end{aligned}$$

Using envelope theorem we have

$$\begin{aligned} \frac{d\text{profit}^*}{d\beta} &= -\lambda_n(1-\alpha)D_\beta^* \\ &= -\frac{(1-\alpha)^2\lambda_n(1-\chi)\chi}{(\beta\chi + (1-\alpha)(1-\chi))^2} (P + R + c_n(W_n - W^U)) \leq 0. \end{aligned}$$

The inequality above follows since $W_n \geq W^U$.

References

- [AHRQ(2014)] AHRQ. 2014. *Emergency Severity Index (ESI): A Triage Tool for Emergency Department: Implementation Handbook*. Agency for Healthcare Research and Quality.
- [Baker et al.(2003)Baker, Wagner, Singer, and Bundorf] Baker, L., T.H. Wagner, S. Singer, M. Bundorf. 2003. Use of the internet and e-mail for health care information: Results from a national survey. *JAMA* **289**(18) 2400–2406.
- [CDC(2011)] CDC. 2011. *National Hospital Ambulatory Medical Care Survey: 2011 Emergency Department Summary Tables*. Centers for Disease Control/National Center for Health Statistics. URL http://www.cdc.gov/nchs/data/ahcd/nhamcs_emergency/2011_ed_web_tables.pdf.
- [CR(2009)] CR. 2009. Urgent care: What you need to know. *Consumer Reports on Health* URL <http://www.consumerreports.org/cro/2013/01/urgent-care-what-you-need-to-know/index.htm>.
- [Dobson et al.(2011)Dobson, Hasija, and Pinker] Dobson, Gregory, Sameer Hasija, Edieal J Pinker. 2011. Reserving capacity for urgent patients in primary care. *Production and Operations Management* **20**(3) 456–473.
- [Fries(1976)] Fries, B.E. 1976. Bibliography of operations research in health-care systems. *Operations Research* **24**(5) 801.
- [GAO(2011)] GAO. 2011. Hospital emergency departments: Health center strategies that may help reduce their use. *GAO Report GAO-11-643T*.
- [Hing and Bhuiya(2012)] Hing, Esther, Farida Bhuiya. 2012. Wait time for treatment in hospital emergency departments: 2009. *NCHS Data Brief, No. 102*.
- [Kleinrock(1976)] Kleinrock, Leonard. 1976. *Queueing Systems*, vol. II: Computer Applications. Wiley Interscience.
- [Nyman(1989)] Nyman, J.A. 1989. Excess demand, consumer rationality, and the quality of care in regulated nursing homes. *Health Services Research* **24** 105–127.

- [Padgett and Brodsky(1992)] Padgett, D.K., B. Brodsky. 1992. Psychosocial factors influencing non-urgent use of the emergency room: A review of the literature and recommendations for research and improved service delivery. *Social Science & Medicine* **35**(9) 1189 – 1197.
- [Shonick and Jackson(1973)] Shonick, William, James R. Jackson. 1973. An improved stochastic model for occupancy-related random variables in general acute hospitals. *Operations Research* **21**(4) 952–965.
- [UCAOA(2014)] UCAOA. 2014. Benchmarking survey headlines summary. *Urgent Care Association of America* .
- [Weinick et al.(2010)Weinick, Burns, and Mehrotra] Weinick, R. M., R. M. Burns, A. Mehrotra. 2010. Many Emergency Department Visits Could Be Managed At Urgent Care Centers And Retail Clinics. *Health Affairs* **9** 1630–1636.