Feel Blue so Go Online: An Empirical Study of Social Support among Patients

Abstract

This paper investigates whether an online healthcare community benefits patients' mental health. We propose an inhomogeneous Partially Observed Markov Decision Process (POMDP) model to examine the latent health outcomes of online health community members. The transition between different health states is modeled as a probability function that incorporates different forms of social support that patients receive via online communication, and other factors that impact patients' online behaviors. We find that patients gain benefits from learning from others, and their participation in the online community helps them to improve their health condition and better engage in their disease self-management processes. Our results also reveal the effectiveness of various forms of social support on the dynamic evolution of patients' health conditions. We find measureable evidence that informational support is the most-sought support in the online healthcare community. However, emotional support plays the most significant role in helping patients move to a healthier state. The helpfulness of social supports is found to vary with patients' health conditions. Finally, we demonstrate that our proposed POMDP model can provide accurate predictions for patients' health states, and it can be used to recover missing or unavailable information on patients' health conditions.

Keywords: healthcare, social networks, social support, Partially Observed Markov Decision Process, user generated content

1. Introduction

The Internet is changing the way people learn about health and illness (Ziebland et al. 2004). According to the Pew Research Center, the number of Americans who sought health information online in 2008 stood at 61%, up from 25% in 2000 (Pew Internet & American Life Project 2009). In 2010, the 59% of American adults who used the Internet to research health problems constituted 80% of Internet users (Pew Internet & American Life Project 2011). Unique features such as its reach to a vast audience in a cost-efficient way, 24/7 accessibility, and user anonymity make the Internet a place people can turn for social support at any time. Different from traditional health services, no need to be spatially and temporally co-presented allows the Internet to provide a safer environment for disease sufferers to engage in nonthreatening and supportive communication (Coulson 2005).

The intersection of healthcare and Internet provides enormous potential for facilitating health services, perhaps more importantly, for the development of mental health programs accessible to many who do not or cannot seek professional treatment (Christensen and Griffiths 2000). Mental health is defined as an individual's ability to respond to the many, varied experiences of life with flexibility and a sense of purpose (Oluwole et al. 2011). People with serious mental problems have difficulty balancing themselves, other people, and the surrounding environment. Take depression as an example. One in five people in the United States have some form of depression. According to the National Institute of Mental Health, there were 33,000 suicides in the United States in 2006, and more than 90% of those individuals had been diagnosed with a mental disorder. Clinical depression is now the second-most-costly disorder among all medical diseases in the United States.

Extensive studies over the past few decades indicate that the level of social support in people's lives essentially predicts their physical and mental health conditions (Clark 2006). However, people with chronic illnesses, especially mental problems, often cannot develop and maintain relationships offline (Leung 2011). As a result, many report spending the majority time alone, and experiencing feelings of social isolation or loneliness (McCorkle et al. 2008). While only a minority of those with mental disorders seek professional help (Christensen and Griffiths 2000), many find online social interactions attractive and effective to obtain much needed emotional support and companionship (Leung 2011). In light of this growing prevalent social trend, online health communities and health social networking are booming (Agarwal et al. 2010).

Despite the increasingly important role played by online health communities, how helpful this emerging patient-driven healthcare model might be for patients is largely unknown (Lamberg 2003). To the best of our knowledge, there is little research that systematically studies the social influence brought about by patients' participation in online healthcare communities and the sharing of their disease information and knowledge. The objective of this work is therefore to examine the impact of patients' activities in online social networks on their health conditions. The challenge of studying this problem is the difficulty of measuring perceived utility through patients' online behaviors in online healthcare communities, especially when patients' health conditions are hidden most of the time. To overcome this obstacle, we propose a Partially Observed Markov Decision Process (POMDP) model, where a patient's health condition is partially observed and varies over time. The transition between different health conditions is determined by the extrinsic medical help and a set of covariates measuring the benefits patients receive from their activities in online healthcare community. The number of health condition states is determined by the complex characteristics of patients' online behavior and dynamics over time. To control for individual-specific characteristics, we include a set of random-effect coefficients to capture this unobserved heterogeneity. Finally, a maximum likelihood estimation procedure is conducted for this POMDP model.

The POMDP model we propose identifies dynamic changes in health condition according to patients' online activities and provides evidence for the helpfulness and benefits of online healthcare communities. Through incorporating partially observed patients' health conditions to examine (latent) dynamic changes, we find that patients receive various forms of social support from their online activities, and that communication with other similar patients has a positive impact on their health conditions. Although information is the major social support that patients seek/provide in online healthcare communities, emotional support has a higher magnitude of influence in helping patients move to a better health condition.

This work, to the best of our knowledge, is the first study to focus on online healthcare communities, where patients share their medical histories and health information to help one another. We find measurable evidence of how patients' social interactions affect their health conditions. Our work bridges the social networking and healthcare fields, offering the following contributions. First, we study patients with mental (chronic) problems and their online activities. We find quantitative evidence that online healthcare communities change patients' disease management behaviors. Users who directly participate in social support exchanges experience support in various forms, such as information about their condition or the knowledge that others are experiencing similar stressful situations. Social support helps them to stop blaming themselves for their illness and presents them with opportunities to actively engage in mutual aid and self-assistance. Our findings on the effectiveness of informational support suggest that an online healthcare community acts as a health knowledge repository. Patients who are managing their disease and understand its progression are tremendous resources for other patients suffering similar problems. The access to a massive library of health data and visualized networking tools make it a true "health university" for the public (and provides the opportunity for post-trial analysis for the pharmaceutical industry). Especially, for rarer conditions, online health networking might be the only means for patients to interact with other similar sufferers who are geographically scattered. Second, the proposed POMDP model can help to recover patients' missing or unavailable information. It takes time and effort for patients to keep track of their health condition. And sometimes, patients may not find the opportunity to get their health condition assessed. Under such conditions, our work postulates a way to reveal the unobservable information effectively and accurately. Third, this work suggests a less costly diagnostic tool. Even if patients do have the chance to measure their health condition, the complex process reduces patients' willingness to participate because it is time-consuming and there is a significant cost in effort. Since the online platform helps patients to keep their health condition history and take control of their health and healthcare, our model thus provides a method to further simplify the procedure and generate a dynamic questionnaire covering patients' health history and previous health condition. Hence, it can increase patients' participation rate by requiring less time and effort and thus encouraging them to reveal more important and valuable information. Finally, our work is not limited to mental health problems, as it can be applied to other types of illnesses.

The rest of the paper is organized as follows. In Section 2, we review the literature and develop the theoretical framework. Section 3 discusses the research context of online health social networking and online health data. The empirical model is presented in Section 4. We explain the data set and key variables in Section 5, and present the results in Section 6. In Section 7, we provide further discussion on the model and analyses. Finally, Section 8 concludes the study.

2. Theory and Hypotheses

This study is in the context of the emerging literature on health social networking and patient-driven healthcare models. Social capital – the resources embedded in social networks – is widely considered to influence health (Abbott and Freeth 2008). It is believed that social media is well suited for healthcare and represents a promising space (Fichman et al. 2011). As such, we have observed the emergence of many healthcare related social networking websites which have, over the years, evolved to virtual platforms to bring together patients with shared interests to communicate with and help each other (Swan 2009).

Following Merton's (1976) description of the role of the "good doctor," Radley and Billig (1996) advocate that the "good patient" must be more than a patient to receive the entitlement. For this reason, aside from physical conditions, internal attitude plays an important role in "defining" patients' health conditions. Many of those who join online healthcare communities are patients who are active in their self-care process. An online healthcare community provides them with the opportunities to gain support within a virtual network formed by individuals dealing with similar issues. Different from the widely used, email-based type of support group, an online healthcare community offers significant advantages, such as access to voluminous data (the aggregated knowledge generated by members) and live online discussions. These modern social media-based communities are constructed on a commons-based peer production basis (Benkler 2002, Fichman et al. 2011) and are especially attractive to individuals with rare diseases and/or chronic health problems. For many people who participate in online healthcare communities, the platform is used to supplement traditional offline methods of support; but for others, the online venues may be the only social support available.

2.1. Types of Social Support

Social support is defined as an exchange of resources between at least two individuals (Shumaker and Brownell 1984). A positive relationship between people's health and social support has long been recognized (Langford et al. 1997). Social support is one of the most important factors in predicting individuals' overall physical health (Chernomas and Clarke 2010, Clark 2006). Cobb (1976) explains that supportive interactions among people protect against the health consequences of stress. McCorkle et al. (2008) find that social support increases patients' adherence to treatments and enhances recovery. As a result, researchers across disciplines have been studying the social support of individuals in various scenarios. Today, a major body of sociology research categorizes social support in four forms: informational support, emotional support, companionship, and instrumental assistance (Berkman et al. 2000, Wortman and Conway 1985).

Informational support involves the transmission of information, including advice and referrals. The convenient access to the mass of information makes health-related topics one of the most popular searches on the Internet (McMullan 2006). The Internet is a source for mental health information for over 10% of the general population, and more than 20% of people who have a history of mental health problems (Powell and Clarke 2006). There is also a significant amount of research showing extensive use of online healthcare communities. Members of online healthcare communities create health profiles or blogs to share geographic and demographic information such as age and gender and track the effects of various medical treatments. Rather than sorting through huge piles of test results and other paperwork, online health profiles help patients keep track of their treatment progress and medications easily and conveniently. This simplifies medical interpretation, enables patients to understand their condition better, and helps patients make treatment decisions (McMullan 2006).

Bandura (2004) has constructed a theoretical framework based on social cognitive theory to examine health promotion. This theory "specifies a core set of determinants, the mechanism through which they work, and the optimal ways of translating this knowledge into effective health practices." One of the core determinants is knowledge of health risks and benefits which creates the precondition for the change of individual health behaviors and habits. The increased access to shared health information, medical experiences, and treatment history in online healthcare communities can produce more informed patients. The knowledge gained from informational support can help provide a greater understanding of problems and possible solutions. The more health information patients learn, the better they understand their condition and the better they can take steps to care for themselves (Kassirer 2000, McMullan 2006, Wanless 2002). Learning experiential information from other patients' profiles provides an individual with a window on other's second opinions, information that is "difficult" to ask directly, and assistance in making sense of the stage of the disease (Ziebland et al. 2004). Therefore, it makes the Internet and online healthcare communities an attractive resource for information about new treatments and discoveries.

Hypothesis 1: *Informational support in online healthcare communities has a positive effect on patients' health conditions*.

Emotional support comes in the form of sharing happiness or sadness, or expressing caring and concern. It sends a signal that one is not alone, that one is taken care of and valued. This kind of support is especially important for patients with chronic mental problems. First, patients with mental health issues have difficulty in developing and maintaining relationships to receive meaningful help. At different stages, the disease can inhibit the ability of patients to cope with their illness. Family relationships can become strained and support withdrawn because of the various burdens that stem from the disease (Weinberg et al. 1995, Wright 2000). Second, due to the limits of time and resources, it may be difficult for offline relations to provide support when it is needed. However, with no geography boundary, online healthcare communities make it possible for patients to talk with other patients suffering from similar illnesses at any time (Bambina 2007, Lamberg 2003). Third, most importantly, knowing that others have faced a similar problem, and even overcome it, can provide both relief from personal blame and renewed strength (Bambina 2007, Weiss 1974, Wills et al. 1985). It is found that online healthcare community members often develop intimate and trusting relationships; among other things, they provide referrals and encourage each other to stick with therapy (Lamberg 2003).

Hypothesis 2: *Emotional support in online healthcare communities has a positive effect on patients' health conditions*.

Companionship can consist of group meetings, chatting, and other social activities. It provides support by making individuals feel there are others who enjoy their presence and that they are a valuable part of something bigger than themselves (Wellman and Wortley 1990). In an online healthcare community, such support is usually exchanged by participating in a discussion forum. The various activities in online healthcare community act as "talk" therapy and can make each individual feel they are not isolated from the world and have social connections. Finally, instrumental or practical support refers to assistance in locating life-related resources. This kind of support is usually not available in online healthcare community settings, as it requires individuals to reveal their real-life identity. In this study, social support is classified to either informational, emotional, or companionship, so their respective effect cannot be empirically identified simultaneously. We have selected companionship as the base category, so the effects of informational or emotional support are relative to that of companionship.

2.2. Social Support as a Process

While it is important to differentiate the type of social support, Jacobson (1986) points out that the "timing" or sequence of social support can affect its effectiveness. A medical problem may need different type of support as it moves through its disease stages (Pearlin 1985). This calls for social support to be examined as a dynamic process rather than just a resource or outcome (King et al. 2006). This perspective is supported by prior studies which have emphasized the importance of social-exchange processes in social support (Antonucci and Jackson 1990). In the context of this study, the dynamics is characterized by a patient's fluctuating among health conditions.

Cohen and Wills (1985) have proposed two models: the direct-effect and stress-buffering model, to explain the influence of social support on stress and health. The direct-effect model asserts that social support protects health, whether stress is present or not. However, according to the stress-buffering model, social support is less effective, or relatively unimportant, for patients experiencing low levels of stress. While our Hypotheses 1 and 2 support the direct-effect model, the stress-buffering model and the process view of social support suggest that the helpfulness of social support may vary depending on patients' instant health conditions.

Hypothesis 3: *The effect of social support in online healthcare communities is moderated by patients' health conditions.*

2.3. Antecedents of Social Support

Social network is identified as the vehicle through which social support, for example, the "give and take" of helpfulness and protections, is provided (Langford et al. 1997). In the context of online healthcare community, the network is the structure of an interactive process allowing patients to share and research information, seek help, make treatment decisions, construct social connections, and find alternative therapies for advocacy, escape, and prevention (Ziebland et al. 2004, Reeves 2001).

The virtual relationships that developed in a virtual community play an important role in meeting patients' social needs (Leung 2011). The online health social networking has become the "crowdsourcing" and that it allows individuals to observe and react to information provided by others, especially learning how to interpret data. The shared medical information, practical tips and advice online help patients to develop their own quasi-professional knowledge of their health conditions (Griffiths et al. 2012). The collective learning and experience of others can be leveraged and is particularly related to health condition (Swan 2009). The opportunity of displaying familiarity with a remarkably body of medical and experiential knowledge about the illness enables a patient to gain a modern form of competence and social fitness in the face of serious health problems (Ziebland et al. 2004). Other than competent patients, the connectedness patients have to others in the healthcare community also changes their relationship with illness. Some degree of connectedness in the network indicates a patient's social embeddedness and how she derives the support from the environment. It is found that with strong social embeddedness, even someone who is experiencing difficulties does not suffer to the same extent as a more isolated individual (Berkman and Breslow 1984).

Hypothesis 4: *Social embeddedness and social competence in online healthcare communities illustrate the depth and strength of social support and thus have a positive effect on patients' health conditions.*

Figure 1 shows the conceptual framework where social support influence patients' health conditions and there is a moderating effect. It is theorized that patients' health conditions consequently affect their online activities and also moderate other control variables.

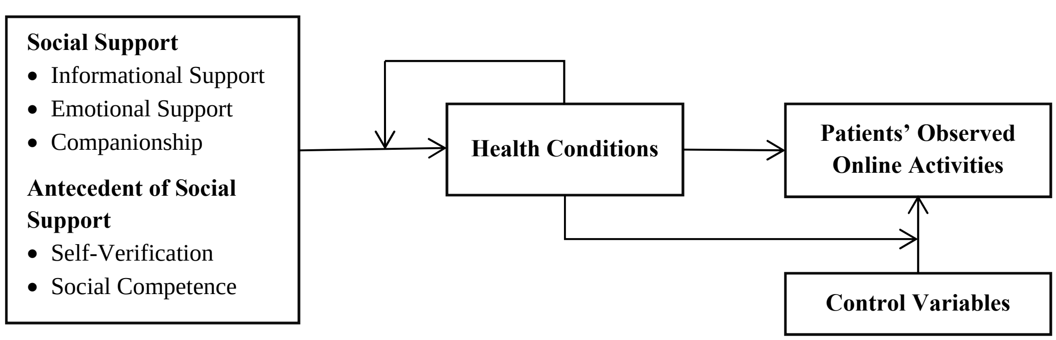


Figure 1: Conceptual Framework

2.4. Reciprocity and Altruism in Social Support

In this study, we focus on the effects of different types of social support on health conditions. In our research context, patients participate in ongoing discussions, via the public discussion forum of an online healthcare community, to exchange social support. Due to the limitation of our data, we cannot discern the support offered from that received. Therefore, we measure the support experienced, from both receiving and offering activities. While the effect of receiving support on health is more or less expected, it is less obvious for that of giving support. In the following, we argue, however, that giving support helps a patient to better health as well.

The act of seeking or giving support immediately triggers opportunities to receive support, and so can be viewed as a proxy measure of support received. This is self-evident in the case of seeking support, since often the first step in getting support is to seek it. We have observed that many posts which offer emotionally supportive things to someone else often receive a supportive message immediately in return from that person or others contributing to the thread. Before replying to a post with a supportive message, one patient "read the entire thread," got the sense of "I am not alone," and experienced implicit emotional support (Swan 2009). There are also occurrences where patients learn new things by reading other's posts before replying and hence receive implicit informational support.

All of these are anecdotes of reciprocity in social support which has been extensively examined in the literature (Antonucci and Jackson 1990). Putnam (1993) defines generalized reciprocity as "a continuing relationship of exchange that is at any given time unrequited or imbalanced, but that involves mutual expectations that a benefit granted now should be repaid in the future." Therefore, this gives the notion that "By helping others, you help yourself in the long run." or "What goes around comes around." Jung (1990) examined three aspects of social support: amount received, amount given, and reciprocity in relationship to coping with stress. Reciprocity was found to have a stronger relationship with reduced symptoms than the amount of social support that was either received or provided. Jou and Fukada (2002) developed questionnaire items to measure the support provided for, requested by, requested of, and received from others, and consequently constructed the measure for reciprocity of support. They find that the health of participants in reciprocal relationships is better than that of participants in nonreciprocal relationships.

Another mechanism that supports our argument is related to altruism and its relationship with health. Researchers have suggested that altruistic (other-regarding) emotions and behaviors are associated with greater well-being, health, happiness, and longevity (Post 2005). According to Midlarsky (1991), altruism results in deeper social integration, distraction from personal problems, enhanced meaningfulness, increased perception of self-efficacy and competence, and improved mood or more physically active lifestyle, and hence leads to better mental and physical health. Therefore, "It's good to be good." This notion is supported by empirical evidence. For example, Schwartz et al. (2003) investigated altruistic social behaviors, such as helping others, of more than 2000 members of the Presbyterian Church located throughout the United States. They find that both helping others and receiving help were associated with better mental health. However, giving help was associated with higher levels of mental health, above and beyond the benefits of receiving help.

3. Research Context

In this paper, we focus on a Health 2.0 website that primarily directs at patients and provides services for them to interact with each other. This communication platform offers the opportunity for patients to find others in similar health situations and share information about conditions, symptoms, treatments and other needs. The key benefits thus include providing a more comprehensive look at a patient's health condition and covering a deeper and broader range of conditions than traditional health expedient.

3.1. Data Description

Like other social network websites, this virtual place provides registration forms for patients to share their medical history and disease details, as well as communication platforms. To provide direct help and better service, this website is organized by health problems and patients are routed to their target communities based on the type of their disease. Members are required to disclose their health condition at the time of registration, and are thereafter directed to the targeted community at every login. Each community is a closed environment, on the belief that patients suffering from similar diseases will better understand each other and thus exchange social support more efficiently. Although members in different communities can view each other's profiles, information access is limited and patients from different communities cannot leave comments or initiate threads on any forum other than their own. Therefore, the boundaries of this online healthcare community are already defined by the website structure.

3.1.1. Individual Profile and Shared Health Information

In the community that serves people with mental problems, patients, as usual, must first create their personal profiles. Similar to other online social networking sites, users need to provide basic information to introduce themselves (for example, creating a username with geographic and demographic information, providing an email address, etc.). In the context of an online healthcare community, however, the "basic information" focuses more on their health, such as the type of their major problem (and perhaps a second or third health problem), the date of the first symptom, and the results of any diagnostic testing. Depending on how much health information a patient shared in the community, the value of profile is identified and controlled by rewarding an indicator based on the volume and quality of information.

There are four levels of data quality on patients' profiles, indicated by 0 to 3 stars. If there is only basic membership information with no health data, the profile receives no stars. One star is assigned to a patient who completes her profile with biographical and condition history information. Another star is added if the patient updates her treatments, symptoms, and mood maps for three months. Patients are also asked to provide names of prescription medications as well as significant supplements, equipment used, and other interventions. After completing four mood maps, patients receive a third star, indicating their profile is complete. Thus the profile keeps each member's shared information and online activities up to date. There is also a medical application in the profile that allows patients to update recent health data and display these in chart form. Thus visualized, it is easy for patients to track their health history. From a practical point of view, consecutive records on health conditions help a patient to better understand her disease progress and also offer an integrated overview to the patient's healthcare providers.

3.1.2. Online Communication and Social Intervention

Without the geographic boundaries, the online healthcare community provides a large pool of patients as potential contributors who have different anecdotal knowledge and motivations (O'Grady et al. 2008). They enter qualitative and quantitative health data about their own conditions, symptoms, treatments, and overall experiences. To take advantage of these shared resources, each member in the mental online healthcare community can use a search tool to easily find other patients who suffer from similar symptoms or experience similar treatments. Once members find a valuable user – and think that person might have some information they need – the members can leave comments on the profile or send private messages. Hence, the number of communications reflects the quality of a patient's profile as well as her online interactions. As a reward for sharing outcome data, other patients can set a flag to express appreciation for the profile host's hospitality and generosity.

The basic service offered by the online healthcare community is the exchange of social support. The forum, outside of the individual's profile level, is a social channel for every patient in the online healthcare community. As a broadcast-type of virtual place, members with general access can exchange general information, ask questions, seek help, provide useful information, or just chat. In addition to the functionality of the email-group based social support, the credence and the value of these conversations can be further differentiated. In particular, each post is evaluated for usefulness by other patients. A utility score is added to the post if another reader finds it helpful. Thus, although there is no hierarchical structure for social conversations occurred in forum, patients still have guidance to find the proper discussion meets their needs. Detailed data and variable descriptions are shown in Table 1.

Table 1: Data and Variable Descriptions

	Variable Name	Description
Profile Data	*gender*	the declared sexuality by a patient: female = 1, male = 0
	membership	number of days that a patient staying in this forum
	update	1 if profile is updated in the period; 0 otherwise
	info quality	the number of stars a patient receives for her profile
	# treatment	the in-period number of treatments a patient takes
	# symptom	the in-period number of symptoms a patient suffers
	posts	cumulative number of posts on forum by a patient
	usefulness	cumulative number of usefulness rated by other patients
	views	cumulative number of times a patient's profile being viewed
	thank you	cumulative votes for shared personal health information on the profile
	comments	cumulative number of comments left for a patient's profile
Social Support	*emo. support*	in-period amount of emotional support for a patient
	info. support	in-period amount of informational support for a patient
	companionship	in-period amount of companionship for a patient
Social Network	*in-degree*	the number of incoming ties in a patient's social network
	out-degree	the number of out-going ties in a patient's social network

3.2. Partially Observed Health States

The health-related information that patients upload to their profiles and share with other patients in the mood community includes their instant mood, functionality level, overall distress level, and detailed distress components, treatments, symptoms, and counseling. Different from instant mood inputs, patients need to take a weekly multi-point survey to be diagnosed so as to obtain the assessments for their health condition measurements such as functionality level defined by this online healthcare community. For example, the online survey contains detailed questions about symptoms such as sleep quality, headache severity, problems concentrating, stomach pain, nervousness, hopelessness, and treatments such as drug dosage. As functionality level is a more comprehensive measure, we discretize it and use it to operationalize the variable for health condition (state) in our model.

It can be often overwhelming for patients to fill out such a detailed survey every week. Like the difficulty of paying doctor visit offline, the time and effort needed to have this virtual diagnosis reduce patients' active engagement in the online healthcare community. The lack of effort from patients results in missing information. On average, we observe 46.09% of patients' health condition points.

3.3. Social Network Construction

In our paper, a social network is constructed based on commenting activities on patients' profiles. Since a comment is a one-way communication representing a patient's willingness to interact, the act of leaving a comment establish a network tie directed form the commenter to the recipient. The degree centrality denotes the extent to which individual patients are involved with others in the social network. This helps document how much benefit a patient receives from and contributes to the website. Figure 2 shows a snapshot of a partial social network. The size of the node represents the number of communications to or from that node; the bigger the node, the more connections. In other words, the size of a node can be viewed as a proxy for the visibility of the patient in the network. This network is constructed at every time period, to reflect the changes of patients' social status.

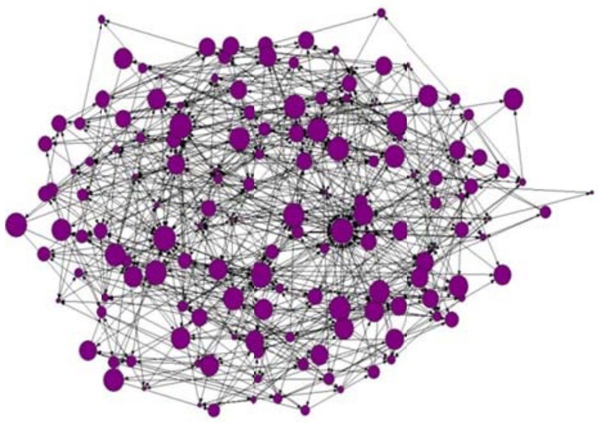


Figure 2: Patients' Social Network

4. Empirical Model

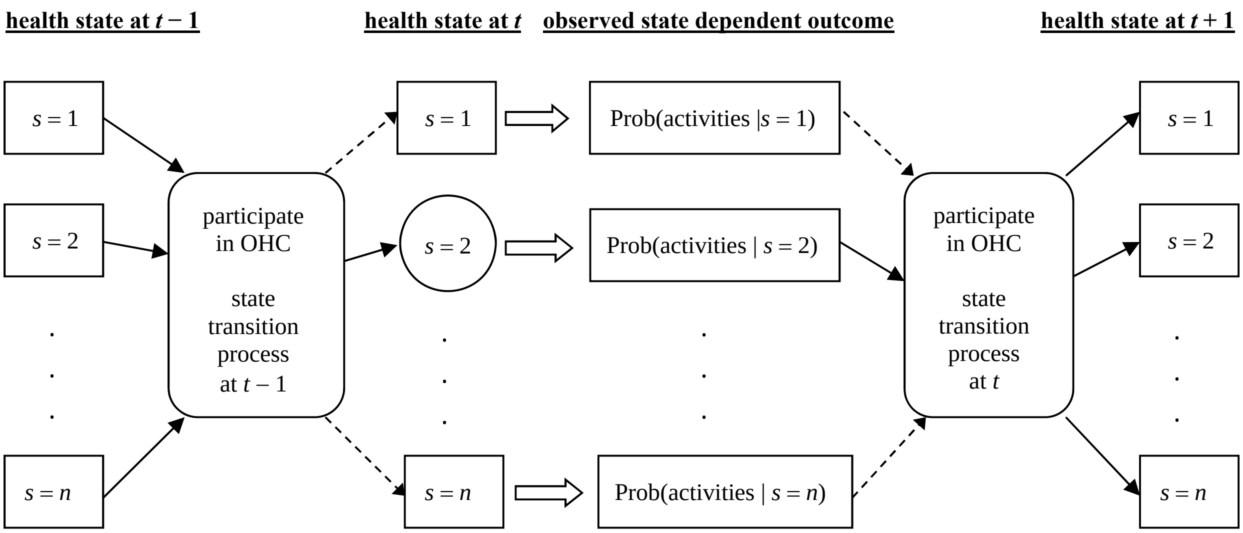
One of the challenges in this study is that the health conditions of a patient are partially observable and evolve over time. This makes it impossible to compare consecutive health conditions and draw the conclusion whether patients benefit from social support. To recover latent health conditions, we propose a model based on Partially Observed Markov Decision Process (POMDP) in which patients' health conditions can be inferred from other observables such as their online activities

4.1. Partially Observed Markov Decision Process

The POMDP starts with the Hidden Markov Model (HMM), which is modified to account for the partial observability of health conditions. HMM is a stochastic process that is not directly observable but can be inferred through another stochastic process that produces a sequence of observable outcomes (Rabiner 1989). It has been widely applied to different contexts. For example, Netzer et al. (2008) capture customers' dynamic relationships by modeling latent relationship states. Hauser et al. (2009) study customer's cognitive style in the context of website morphing. Singh et al. (2011) identify developer's learning dynamics from their past experiences and interactions with other peers in the context of open source software development.

In this study, we identify a patient's health condition as the latent state, whenever it is unobservable, and study whether a patient's online participation helps to change her latent state. Such a transition can be triggered by communication, the exchange of information and knowledge, or other activities with patients in the online healthcare community. The time variant online activities define the observed outcome sequence for a patient. The Markovian transitions account for the dependence on subsequent behaviors. Figure 3 sketches the POMDP in this study. In contrast to HMM, a patient's (latent) health condition is not completely hidden. Therefore, by considering this information we are able to reduce the randomness of the HMM and redefine the distribution of latent states.

Following the notation of Rabiner et al. (1986, 1989), the proposed POMDP model is a combination of HMM and a probability adjustment process. It consists of three main components and an additional probability recalculation process for the partially observed health condition (state): (1) the initial state distribution, π; (2) the state transition probability distribution, A; (3) the observed outcome probability distribution, B; and (4) the recalculation of transition probability distribution, A', if the state is observed at the given period. For convenience, we use a compact notation for the overall model: $\lambda = (A / A', B, \pi)$. With this model specification, we proceed to find the values of parameters in $\lambda = (A / A', B, \pi)$ to best explain the observed outcome sequence, or to maximize the probability P (outcome sequence | λ).



A rectangle indicates an unobserved state whereas a circle refers to an observed state. A solid arrow denotes a possible path whereas a dashed arrow indicates a forbidden path. OHC is the acronym for online health community.

Figure 3: POMDP diagram

4.2.State Transition Matrix

We assume that there are n states that discretize health conditions from the lowest health state 1 to the highest health state n. A patient takes medications, receives treatments, and other therapies. Even with such external controls, one still observes that very often a patient's health condition changes quite drastically. As the changes in health state can be very random, we relax the assumption of random walk in a typical POMDP model. The state transition probability is defined as $A = \{a_{it}(s,m)\}$, where $a_{it}(s,m) = P\left(S_{it+1} = m \mid S_{it} = s\right)$, $1 \le s, m \le n$; and S_{it} denotes the state of patient i at time t. For each state s, we have $\sum_{m=1}^{n} a_{it}(s,m) = 1$ and $a_{it}(s,m) \le 1$.

As discussed earlier, a patient's health condition changes according to her reception of social support through her online communications and other activities. A continuous measurement of this propensity needs to be modeled into the probability transition matrix. In other words, a patient can move to a higher health state if the benefit she gets from the online healthcare community is greater than a certain threshold, whereas she will transit to a lower state if the aggregate social impact is lower than a low threshold value. Hence, the matrix is defined as an ordered Logit model:

$$a_{it}(s,n) = 1 - \frac{\exp(\overline{\omega}_{s\to n} - \beta_s X_{it} - \xi_i)}{1+\exp(\overline{\omega}_{s\to n} - \beta_s X_{it} - \xi_i)};$$

$$a_{it}(s,n-1) = \frac{\exp(\overline{\omega}_{s\to n} - \beta_s X_{it} - \xi_i)}{1+\exp(\overline{\omega}_{s\to n} - \beta_s X_{it} - \xi_i)} - \frac{\exp(\overline{\omega}_{s\to n-1} - \beta_s X_{it} - \xi_i)}{1+\exp(\overline{\omega}_{s\to n-1} - \beta_s X_{it} - \xi_i)};$$

$$\cdots$$

$$a_{it}(s,s) = \frac{\exp(\overline{\omega}_{s\to s+1} - \beta_s X_{it} - \xi_i)}{1+\exp(\overline{\omega}_{s\to s+1} - \beta_s X_{it} - \xi_i)} - \frac{\exp(\underline{\omega}_{s\to s-1} - \beta_s X_{it} - \xi_i)}{1+\exp(\underline{\omega}_{s\to s-1} - \beta_s X_{it} - \xi_i)};$$

$$\cdots$$

$$a_{it}(s,2) = \frac{\exp(\underline{\omega}_{s\to 2} - \beta_s X_{it} - \xi_i)}{1+\exp(\underline{\omega}_{s\to 2} - \beta_s X_{it} - \xi_i)} - \frac{\exp(\underline{\omega}_{s\to 1} - \beta_s X_{it} - \xi_i)}{1+\exp(\underline{\omega}_{s\to 1} - \beta_s X_{it} - \xi_i)};$$

$$a_{it}(s,1) = \frac{\exp(\underline{\omega}_{s\to 1} - \beta_s X_{it} - \xi_i)}{1+\exp(\underline{\omega}_{s\to 1} - \beta_s X_{it} - \xi_i)}.$$

Here, s is the current state. $\underline{\omega}_{s\to k}$ is the threshold for the current state *s* to transit to a lower state *k* ($k < s$) and $\overline{\omega}_{s\to k}$ is the threshold for the current state *s* to transit to a higher state *k* ($k > s$). For a given *s*, we have $\overline{\omega}_{s\to n} \ge \overline{\omega}_{s\to n-1} \ge \cdots \ge \overline{\omega}_{s\to s+1} \ge \underline{\omega}_{s\to s-1} \ge \cdots \ge \underline{\omega}_{s\to 1}$. X_{it} is a vector containing variables that have impact for patients to switch between states. β_s is a set of state dependent parameters. ξ_i represents the patients' specific random effect and accounts for individual unobserved heterogeneity. As shown in Figure 3, patients in the lowest state can move to any one of $(n-1)$ higher states or stay idle while the highest-state patient can either stay unchanged or move down to any of $(n-1)$ lower states. Patient's health condition lying in any other states has the probability to move either up or down or stay unchanged in the same state.

4.3. State Dependent Outcome

In this paper, we choose the number of new posts a patient initiates and answers as the measurement of her observed online activity. Following Singh et al. (2011), we model that the number of new posts in a period, a count variable, follows a negative binominal (NB) distribution for a given health condition state:

$$P\left(O_{it} \mid S_{it}=s\right)= f_s\left(O_{it} \mid Y_{it};\gamma_s,\theta_s^2\right)=\frac{\Gamma(O_{it}+\theta_s^{-2})}{(O_{it}!)\Gamma(\theta_s^{-2})}\left(\frac{\theta_s^{-2}}{\theta_s^{-2}+\exp(Y_{it}\gamma_s+\eta_i)}\right)^{\theta_s^{-2}}\left(\frac{\exp(Y_{it}\gamma_s+\eta_i)}{\theta_s^{-2}+\exp(Y_{it}\gamma_s+\eta_i)}\right)^{O_{it}},$$

where O_{it} is the number of posts for patient *i* at time period *t;* and θ_s is the state-dependent parameter to capture the possible over-dispersion in O_{it}. Y_{it} is the vector composing variables that have direct impact on the outcome for patient *i* at period *t;* γ_s is the vector containing state-dependent parameters; and $\exp(\gamma_s Y_{it}+\eta_i)$ specifies the expected value of O_{it}, according to NB distribution. The symbol η_i is the patient-specific random effect that accounts for patient's unobserved heterogeneity.

4.4. Adjustment for State Transition Probability with Observed Patient Health Condition

The unobserved states are handled by HMM. If patient *i*'s health state at time period t is observed, her state transition matrix *A* must be modified. Recall that $a_{it}(s,m)=P\left(S_{it+1}=m \mid S_{it}=s\right)$ is an element of *A*. If we observe that $S_{it+1}=m'$, that is, at time period (*t* + 1), with certainty, patient *i* enters a health state of m', then the corresponding state transition probability $a_{it}(s,m)$ is replaced by

$$a_{it}'\left(s,m\right)=\begin{cases}1, & \text{if } m=m';\\ 0, & \text{if } m\neq m'.\end{cases}$$

This adjustment process is in line with that of Kaelbling et al. (1998). For the time period that has the observed state information, all states from the previous time period will enter state m' with a probability of 1, and for the next time period, the possible routes will be initiated only from this state m'.

4.5. Likelihood of an Observed Sequence of Outcomes

Consider an observed sequence of outcomes $O\left(i\right)=O_{i1}O_{i2}\cdots O_{iT}$ for patient *i,* and a sequence of states $S\left(i\right)=S_{i1}S_{i2}\cdots S_{iT}$. The conditional likelihood, on two random effect control variables η and ξ, which account for unobserved patient heterogeneity, is the sum over all possible paths, explicitly,

$$L\left(O\left(i\right) \mid \eta,\xi\right)=\sum_{s_1=1}^{n}\sum_{s_2=1}^{n}\cdots\sum_{s_T=1}^{n}P\left(S_{i1}=s_1\right)\prod_{t=2}^{T}P\left(S_{it}=s_t \mid S_{it-1}=s_{t-1}\right)\prod_{t=1}^{T}P\left(O_{it} \mid S_{it}=s_t\right),$$

where $s_t\in\left\{1,\cdots,n\right\}$ is the state in which a patient can possibly reside in time period *t*. The likelihood of patient *i* can be obtained by integrating over η and ξ:

$$L\left(O\left(i\right)\right)=\int_{\eta}\int_{\xi}L\left(O\left(i\right) \mid \eta,\xi\right)dH\left(\xi \mid \eta\right)dG\left(\eta\right),$$

where the probabilities H and G are evaluated non-parametrically, that is, their supports and corresponding probability masses are considered as model parameters to be estimated.

5. Variable Set and Description

Our data are sampled weekly for 16 weeks. The sample includes patients' online activities on the website as well as their interactions with other members. Specifically, we collect both levels of patients' activities in the online healthcare community: patients' forum activities statistics (the total number of posts is the aggregate number of conversation threads, including topic initiations and replies), helpfulness marks (other patients reward the post by marking it as helpful), as well as their profile activities. Table 2 explains the variable sets and shows detailed summary statistics. The correlation matrix is presented in the online supplement (OLS).

Table 2: Data Statistics

Variable Set		**Variable**	**Mean**	**St. Dev.**	**Min**	**Max**
Online Behavior		*new posts*	2.814	1.551	0	5
Personal Characteristics		*gender*	0.498	0.500	0	1
		info quality	1.481	1.117	0	3
		membership	45.34	33.415	0	118
		update	0.436	0.496	0	1
		# treatment	5.424	4.423	0	28
		# symptom	4.066	3.470	0	27
Antecedents of Social Support	Social Embeddedness	*views*	14.739	7.922	0	43
		thank you	4.255	2.724	0	15
		comments	4.748	2.803	0	16
		in-degree	8.483	5.151	0	24
	Social Competence	*posts*	27.028	14.506	1	70
		usefulness	8.991	5.225	0	29
		out-degree	7.998	4.892	0	23

5.1. Social Support Measurements

As discussed in Section 2, there are four forms of social support. Since we focus on online activities in this research, only three of them are considered: informational support, emotional support, and companionship. We followed the coding scheme proposed by Bambina (2007), and the details are provided in Table 3.

Table 3: Social Support Coding Scheme

Support Categories	Support Subcategories
Informational Support	Advice
	Referral
	Teaching
	Information broadcasting/seeking
	Personal experience
Emotional Support	Understanding/empathy
	Encouragement
	Affirmation/validation
	Sympathy
	Caring/concern
Companionship	Chatting
	Humor/teasing
	Groupness

Social support measures are extracted from forum discussions. Different from the comments on the user profile, the forum is a place commonly used for various kinds of social interactions, allowing richer insights into the experiences and needs of individuals affected by mental problems. Focusing on the patients who participated (and not those who just "lurked") in the forum, we used LingPipe[1] to conduct a semantic analysis on the forum threads. There are 5,192 topics initiated in the Forum and 371,562 posts made during our data collection period. For each post, a number was returned to indicate the probability that this post belonged in a certain category. Because patients tend to provide multiple pieces of information in each post, it would be improper to classify a post into one category only. Therefore, we assign three probabilities (adding up to 1) to each post, corresponding to the topics addressed. With this classification scheme, we interpret the probability as the amount of social support that a patient is offered or asks for. Consistent with the general perception, the informational support is the most exchanged social support type in the online community, followed by emotional support and then companionship. Table 4 gives the statistical details.

Table 4: Social Support Statistics

Variable	Mean	Median	St. Dev.	Min	Max
informational support (weekly)	1.843	1.680	1.063	0.19	7.92
emotional support (weekly)	0.712	0.560	0.578	0	3.68
companionship (weekly)	0.542	0.420	0.442	0	2.8

[1] Alias-i. 2008. LingPipe 3.8.2. http://alias-i.com/lingpipe

5.2. Variables Directly Impacting Patients' Health Conditions

In addition to the social support that would affect patient's health condition, we also conducted analysis of variables that may affect patients' health condition dynamics, and all of these constitute vector X_{it} in the proposed model. For example, the cumulative number of posts by patient *i* till time period *t*, called "posts," could indicate her social competence. This variable measures the familiarity with medical and experiential knowledge about the illness and hence enables a patient to gain a form of competence and social fitness in the face of serious health problems. Another variable involves recognition and appreciation from other patients, which can help a patient to feel capable and valuable. This is measured by "views," which shows the total number of times that a patient's profile has been checked by other community members. The number of times a profile is displayed also indicates the visibility of the patient in the community. If members find a profile particularly valuable, "thank you" is the simple way of showing appreciation for sharing information, while "comments" involve more detailed communications. Not all posts contribute equally; some might contain trivial information while others tend to be more useful. Therefore, "usefulness" is used to measure the value of posts, as assessed by other patients. Each patient can vote only once, for any post except her own. Finally, the willingness to communicate with other patients (out-degree) captures a patient's direct online activity initiated by herself for the collaborative learning. This behavior describes her perception of the online healthcare community and her attitude towards it. In addition, a patient's medical control, including the symptoms she is suffering and the treatments she is taking, can directly affect her health condition.

5.3. Variables Directly Impacting Patients' Online Behavior

There are factors that might directly affect her online behavior pattern. For instance, the quality of a profile indicates a patient's level of concern about her disease, and therefore could directly impact her online activities. As four or more mood maps provide an accurate history of disease progress, it is important for not only sharing with the community but also working with the clinician. For this reason, we define "information quality" to measure a patient's data quality. There are other characteristics that could have a direct impact on patients' online behaviors. It is possible that female patients are more active in the online healthcare community, and this could lead to a different behavior pattern for their observed online behavior given a certain health state. For example, McPherson et al. (2001) find that gender is significant in predicting communication patterns. Hence, we include a “gender” variable in our model to control for this possibility. The date the patient joined the online healthcare community (membership), and the frequency with which the patient updates his profile (update). Thus these variables are proxies used to capture the patient’s activity pattern.

Last, we use a patient’s instant online activity, the number of new posts she initiates or replies in the period, to describe her state-dependent outcome. The change of this measure could result from a change in the patient’s health condition. In other words, the social support she has is insufficient and she needs to communicate in the online healthcare community to fulfill her social demand. The more posts she contributes, the more likely her health condition has changed. Note, this does not imply whether a patient’s health condition has deteriorated or improved. It suggests only that a patient’s online activity relates to her current health state.

5.4. Estimation and Model Selection

We started with a latent class model to estimate the initial distribution for the latent health state, and then used the maximum likelihood method to estimate the model parameters. To control for patients’ heterogeneity, modeled by η and ξ, we followed the approach by Heckman and Singer (1984). The approximation process for the underlying unknown probability distribution was evaluated by finite sampled supporting points associated with probability mass distributions. After rescaling η and ξ by two parameters C_η and C_ξ, respectively, we set the boundary for each of the random effect variables to be between 0 and 1. The number of states n was chosen by the selection criteria of Bayesian Information Criterion (BIC):

$$BIC = \ln L - k \times \ln P / 2$$

where P is the sample size (the number of patients), L is the likelihood of the model, and k is the number of parameters to be estimated. The goal of the model selection process was to choose the model with a probability that approached one as the sample size got larger (Anderson et al. 1998). The results are shown in Table 5. Our estimation indicates that the three-state POMDP outperforms other models.

6. Findings

In this section, we report the results from the POMDP model with three health states (bad, fair, and good). The initial state distribution probabilities are (0.7756, 0.15026, 0.07414), obtained from latent class model.

The estimated parameters are presented in Table 6, where the corresponding standard errors are shown in parentheses.

Table 5: Selecting the Number of States

Number of States	Log-Likelihood	Variables	BIC
1	–14583.9	16	–14655.3
2	–14346.2	34	–14497.9
3	**–13847.3**	**54**	**–14088.3**
4	–14205.8	76	–14544.9

6.1. Hypothesis Test

As shown in Table 6, the estimated parameters for the effects of both informational and emotional supports are positive and significant across all three states. Hence, both Hypotheses 1 and 2 are supported. To test Hypothesis 3, we compare the estimated parameters across the states for informational or emotional support. We find that the effect of emotional support increases from bad to fair, and from fair to good state. Informational support is more effective for patients in a bad state. However, there is no significant difference between fair and good state. Therefore, Hypothesis 3 is partially supported. The impacts of social embeddedness and social competence are discussed in Sections 6.4.2 and 6.4.3. Overall, our estimated parameters suggest that Hypothesis 4 is partially supported.

6.2. State-Dependent Outcome

The parameters for state-dependent outcomes describe the variables that affect a patient's activities in online healthcare community at a given health state. It is interesting to note that – as indicated by the state-dependent constants that give the intrinsic propensity to contribute – the patients turn to create fewer posts as they progress to better health state. In the online supplement, we calculate the expected intrinsic number of new posts which are respectively 2.75 for bad state, 1.56 for fair state, and 1.13 for good state. These numbers are statistically different. The patients in a bad state want to learn more about their disease and hence have relatively more problems or questions to ask than those in fair or good state.

As shown in Table 6, women participated more actively than men in the online healthcare community across all health states. Female patients tended to post more when they were in either the worst (3.44 more posts than the intrinsic number) or best condition (additional 0.96 posts). This may be due to the fact that women are more sensitive to changes in their emotional and physiological states (Hunt

Table 6: Estimated Parameters for the Three-State POMDP[2]

Parameter	**State 1 (bad)**		**State 2 (fair)**		**State 3 (good)**	
θ Dispersion	0.7512***	(0.1213)	0.4758***	(0.0254)	0.5372**	(0.1982)
Variables Impacting State Transition						
β_1 [views]	0.9561***	(0.1837)	0.9287***	(0.2382)	0.7271***	(0.1642)
β_2 [thank you]	1.6298***	(0.1259)	0.6284***	(0.1983)	2.2823***	(0.3487)
β_3 [comments]	3.0172***	(0.1834)	0.8768***	(0.2674)	1.6417**	(0.7283)
β_4 [usefulness]	0.9265***	(0.3342)	0.3372***	(0.1028)	1.4419***	(0.1482)
β_5 [in-degree]	–1.0564***	(0.2166)	0.2209*	(0.1192)	–0.6198***	(0.2128)
β_6 [out-degree]	0.8293***	(0.1698)	0.2231**	(0.1012)	0.6126***	(0.1871)
β_7 [info. support]	0.7637***	(0.1095)	0.6123***	(0.1482)	0.6218***	(0.1001)
β_8 [emo. support]	0.6234***	(0.1243)	0.8198***	(0.2031)	1.0821***	(0.2237)
β_9 [posts]	2.4298***	(0.1771)	1.8728***	(0.1749)	4.4925***	(0.4548)
β_{10} [# treatment]	1.2372***	(0.1210)	0.7213**	(0.2845)	0.8832***	(0.2837)
β_{11} [# sympt]	–0.9218***	(0.0972)	–1.0023***	(0.2693)	0.2178***	(0.0415)
Thresholds						
State 1 (bad)			0.9227***	(0.1894)	3.1593***	(0.6831)
State 2 (fair)	–2.5327***	(0.3287)			2.6481***	(0.4037)
State 3 (good)	–3.0126***	(0.6044)	–1.0469***	(0.1362)		
Variables Impacting State Dependent Outcome						
γ_0 [constant]	1.1126***	(0.2436)	0.5479***	(0.1928)	0.2281***	(0.0244)
γ_1 [gender]	0.8127***	(0.1902)	0.7487*	(0.4235)	0.6127**	(0.2823)
γ_2 [info quality]	0.2841***	(0.0512)	0.4120	(0.2876)	0.1298*	(0.0685)
γ_3 [membership]	0.8236***	(0.0075)	0.5824***	(0.0046)	0.9237***	(0.0029)
γ_4 [update]	–0.7218***	(0.0951)	–0.5218	(0.3827)	–0.9218***	(0.1148)

Unobserved Heterogeneity (η , ξ)

$C_\eta = -0.204, C_\xi = -0.151$

	$\eta_1 = 0$	$\eta_2 = 0.3174$	$\eta_3 = 0.5313$	$\eta_4 = 1$
Probability $G(\eta)$	0.0639	0.3487	0.4015	0.1859
Conditional Distribution: $H(\xi \mid \eta)$				
$\xi_1 = 0$	0.9417	0.2312	0.1302	0.9271
$\xi_2 = 0.4134$	0.0295	0.5675	0.1934	0.0369
$\xi_3 = 1$	0.0288	0.2013	0.6764	0.0360

Significance levels: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

[2] The following rescaling is performed: “membership” is log transformed; “views”, “thank you”, “comments”, “usefulness”, “in-degree”, and “out-degree” are scaled down by a factor of 100; “posts” is scaled down by a factor of 1000; “info. Support”, “# treatment”, “# sympt”, “gender”, “info quality”, and “update” are scaled down by a factor of 10.

et al. 1981) and, hence, more willing to express themselves and their emotions when they are in extreme conditions (bad or good). This finding is consistent with prior studies (e.g., Hunt et al. 1981, Mechanic et al. 1978) that show a sex differentiation in admitting to certain problems.

A patient with good information quality is one who keeps close track of her disease progression. She prefers to seek out social supports, find out underlying reasons, and try to improve her condition. This is confirmed by the significant and positive coefficients in our results. The duration of membership measures the attitude commitment to online healthcare communities. Our results show that a member with a longer tenure, when in good or bad condition, tended to contribute more to an online healthcare community than a newcomer. The significant and negative coefficients for patients' profile update for all states suggest the various form of patients' usage of this health social networking platform and its functionality. For a patient who prefers to use this online tool as a means of self-reporting and documentation, she is more focusing on her own health condition and experiences and thus less likely to participate in the collaborative activities.

6.3. State Transitions and Baseline Results

The thresholds provide the intrinsic propensity to transition from one state to another. As we allow patients to "jump" among different states, these thresholds ensure that moving involves some positive boundary requirements. The intrinsic probabilities[3] to transit among states are shown in Table 7.

Table 7: Intrinsic Transition Matrix

	Bad	Fair	Good
Bad	0.7254	0.2357	0.0389
Fair	0.0771	0.8597	0.0632
Good	0.0492	0.2204	0.7304

Although a patient's health state could change dramatically (even to the point of jumping to a nonadjacent state in our model), our results showed that patients are indeed relatively stable in their health states. Different from mood changes, health status describes a patient's physical and mental ability and the variation is minimized by medication control. The stickiness in the current state could result from the effect of medical treatments that patients received for their mental disease. As for mood problems, medication is not always recommended for those with mild depression because the risks outweigh the

[3] They are evaluated with the estimated threshold values and random effects. The values of all variables were set to zero.

benefits. In our dataset, an average of 35% of patients were taking medical treatments. However, without help from external resources (e.g., various services provided in the online healthcare community), a patient has a lower probability of improving her health condition to a better level. Without participation in an online healthcare community, a patient is more likely to stay at her current health state or get worse.

6.4. Factors Influencing Patients' Health Transition

As our primary objective was to determine the helpfulness of online healthcare communities in improving patients' health condition, we provided detailed discussions about the variables that affect patients' health state and consequently influence their behavior in online healthcare communities. In what follows, we categorize these variables into three groups. The transition probabilities are evaluated with the average value of the focal variable and the values of the other variables set at zero, and compared with the intrinsic transition probabilities.

6.4.1. Impact of Social Support on Health Condition

Table 8 shows the change of transition probabilities, due to informational support, from the intrinsic ones in Table 7 (also shown in the parentheses). By communicating with other members, a patient is more likely to get useful information and better understand her health conditions. Along with information about medical terms and symptom descriptions, personal advice and referrals make the communication more valuable. The firsthand experience available on the online healthcare community helps patients muster the strength to fight their disease; it is also a place to find guidance for self-management. All of this helps to increase the probability that patients will transit to a better health state. For example, compared with intrinsic propensity transition, the probability of a patient in a bad state moving to a fair state is increased by 2.33%[4] after experiencing 1 unit of informational support. With an increase of probability of 2.2%, a patient already at a good health status is more likely to stay in good condition when she receives information support. Our results also indicate that the possibility of moving down to a worse health condition decreases because of informational support.

Emotional support has a significant influence on patients in different states. As shown in Table 9, we observe the same pattern seen with informational support: The benefits of emotional support are significant and positive in all three states. In other words, it increases the probability that patients move to

[4] The changes in probabilities are statistically significant if the corresponding parameter is.

a better health condition. Many studies have found that emotional support plays a critical role in a patient’s outcome. For example, in a study of heart failure, emotional support was found to have significant association with risk for heart disease (Krumhole et al. 1998). We also find evidence supporting the importance of emotional support. With 1 unit of emotional support, patients in a bad state had a 7.6% higher possibility of moving to a fair state and a 9.64% lower possibility of staying in a bad state. A patient who was already in a good state was shown to be more likely to stay in good condition with a 12.37% higher probability.

Table 8: Change in Transition Probability – Informational Support

	Bad	Fair	Good
Bad	–0.0289 (0.7254)	0.0233 (0.2357)	0.0056 (0.0389)
Fair	–0.0077 (0.0771)	0.0006 (0.8597)	0.0070 (0.0632)
Good	–0.0051 (0.0492)	–0.0169 (0.2204)	0.0220 (0.7304)

Table 9: Change in Transition Probability – Emotional Support

	Bad	Fair	Good
Bad	–0.0964 (0.7254)	0.0760 (0.2357)	0.0204 (0.0389)
Fair	–0.0326 (0.0771)	–0.0121 (0.8597)	0.0447 (0.0632)
Good	–0.0258 (0.0492)	–0.0987 (0.2204)	0.1237 (0.7304)

Severe disease affects patients and changes their everyday activities. Researchers in psychosocial and social science have examined social supports in various contexts. Such work includes studying patients’ need for emotional support (e.g., Slevin et al. 1996) and emotional and informational support for patients’ relatives (Eriksson and Lauri 2000). The requirements for such social support change according to the magnitude and time in need. Tables 8 and 9 show that emotional support is overall more influential in changing patients’ conditions, although patients receive more units of informational support in this community.

6.4.2. Impact of Social Embeddedness

Multiple measurements can be used to evaluate how well patients communicate with other community members, and how personal images are built in such a virtual world. By searching for similar patients with certain criteria, a patient can learn more from those members by viewing their detailed profiles. Therefore, the number of times that her profile is viewed indicates how visible a patient is in this community. The coefficient for profile views was positively significant and therefore increased the probability that a patient would move to a better health state. Take a patient in a bad state as an example.

The probability of moving to a fair state in the next time period increases by 2.33%,[5] while there is an increase of 2.89% in the probability of staying in the same state for the next period.

A large number of "thank you" votes indicated the quality of a patient's data. It not only confirmed the patient's effort in disease self-management but also made her feel appreciation for helping others. This satisfaction can influence a patient's ability to move among different health states. We find that patients in a bad state had a 1.13% increase in the probability of moving to a fair state and a 1.4% decrease in the probability of staying in a bad state. Patients in a good state also benefited from confirmation and encouragement, and therefore had a higher probability of staying well.

The measurement for "comments" was intended to signal patients' profile quality. As shown in Table 6, the number of comments on a patient's profile had a significant and positive impact in all health states. This may result from the profile owner being encouraged by recognition and care from other patients, thus increasing her probability of feeling better. Even a patient in a bad health condition who received an average number of comments had a higher probability of moving to a better condition (a 2.37% increase to a fair state). If she was already in good condition, the possibility of staying well increased 1.51%, as compared to when this recognition was absent.

Finally, the in-degree measures the incoming connections of a patient in the community where she was in greetings by others. It can be considered as a proxy for receiving social support from other members in the community. As shown in Table 6, the significant and negative coefficients suggest that patients who are in extreme conditions (bad or good) and receive more social support are less likely to seek out for help. The preference of more interested in talking with her own favored cluster of patients and less likely to participate in community-based communications barriers her to progress to a healthier state (a 1.43% decrease from bad to a fair state).

6.4.3. Impact of Displaying Social Competence on Health Condition

The number of posts and the helpfulness of those posts help to determine patients' social value in this online healthcare community. The number of "posts" that a patient created is an indicator of her attitude in facing the disease and her aggregated knowledge of the disease, which may contain valuable information and experience for others patients. The "usefulness" variable measured the effectiveness of

[5] The full results for the changes in transition probabilities can be found in the OLS.

her posts. In addition, the out-degree, constructed by counting a patient's self-motivated or initiated communication indicate her knowledge about a certain disease and helps to identify her familiar with health problems and potential value for other patients. Table 6 shows the positive impact of such activities on patients' health conditions, all of which suggest a patient's social competence reconstructed in the virtual space.

A patient in a bad health state could create posts in the online healthcare community to seek help. This could help her release unhappiness and pressure and receive advice about her next move, and hence prevent her from falling into a worse condition. There was a 1.33% decrease in her probability of staying in a bad state, and there was a 1.78% smaller probability that a patient in a good state would move to a fair state. The recognition and reward for competence ("usefulness") also helped patients not to get worse. For example, appreciation gave patients in a bad state a 1.37% greater chance of moving to a fair state. A patient already in good condition increased her probability of staying well by 2.47%. Finally, a patient's health condition changes with her outreach behavior as indicated by out-degree; there is a higher probability of moving to a better condition (a 2.37% increase to a fair state) and 3.3% probability increase for a patient to stay in good condition.

7. Further Analyses and Robustness Checks

The analysis above is based on a POMDP model estimated based on an aggregated data over 4 month period, focusing on social support in the online healthcare community. In this section, we present further analysis in contrasting the impact of different forms of social support, discuss the robustness checks of the qualitative findings and the limitation of the model.

7.1. Contrast between Informational Support and Emotional Support

The informational and emotional support delivered through online healthcare communities may help patients cope better with their mental problems. Bambina (2007) investigated a cancer forum and finds that on average members receive more informational support than emotional support. However, it remains unclear, due to the lack of prior empirical evidence, which type of support is better for meeting patients' social needs. It is possible that emotional support is more important than informational support for patients suffering from mental problems as they are more emotional and always feel loneliness because of their disability to maintain social relationships (McCorkle et al. 2008). The perceptions of insufficient social support barrier their recovery from medical illness. On the contrary, the anecdotal and experiential knowledge shared by individuals regarding various treatments and medications create the "wisdom of crowds" and may have an impact on patients' health decision making (O'Grady et al. 2008). Therefore, to help improve the effectiveness of online healthcare communities, it is important to contrast the respective effect of informational and emotional support.

To test the aforementioned argument, we calculate the difference between the parameters of emotional and informational support, for a given state, and the corresponding standard error. The results are reported in the online supplement. They are all significant at 1% level. Hence, we conclude that for patients with mental problems, the emotional support they receive online plays a more important role in helping them to progress to healthier conditions than informational support. That is, our empirical result suggests that emotional support is significantly more effective in helping mental patients to progress to a better state.

7.2. Posterior Analysis

For this analysis, we applied the filtering approach proposed by Hamilton (1989) to recover patients' unobserved health conditions across time periods. Once the model parameters are estimated, the likelihood can be obtained using the information until time *t*. Posterior probability for a patient in a given state can be calculated using Bayes rule. This allowed a patient to be classified, in any given time period, into a health state according to the posterior probability calculation. As observed in Figure 4,[6] about 45% to 50% of patients were in a bad health condition, and 35% to 40% of patients were in a fair state over the time.

In Figure 5, we plotted two examples of individual patient behavior. There was no unique pattern: One remained at the same level and was more or less stable, while the other fluctuated a lot among states. Since the observed information on health states from Week 17 was not used to calibrate the model, the results of posterior analysis for Week 17 and beyond were purely predicted. Figure 5 shows that our POMDP model was very accurate in predicting patients' health conditions. We have examined all the patients in our dataset, and calculated the prediction accuracy, defined as the percentage of correctly predicted health states which are observable from Week 17 to Week 32. The overall accuracy is 93.25%.

[6] Figures 4 and 5 plot the results of posterior analysis on functionality state for 32 weeks. The parameters in our POMDP model are, however, estimated using the data of the first 16 weeks, due to computational complexity.

Therefore, our result shows this is a very effective way for patients and healthcare providers to recover missing or unavailable information.

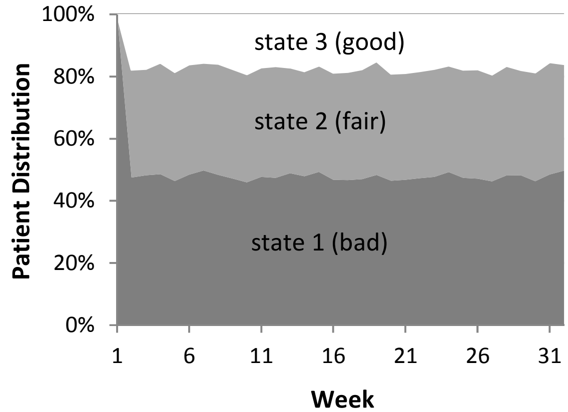


Figure 4: Posterior analysis for patient distribution

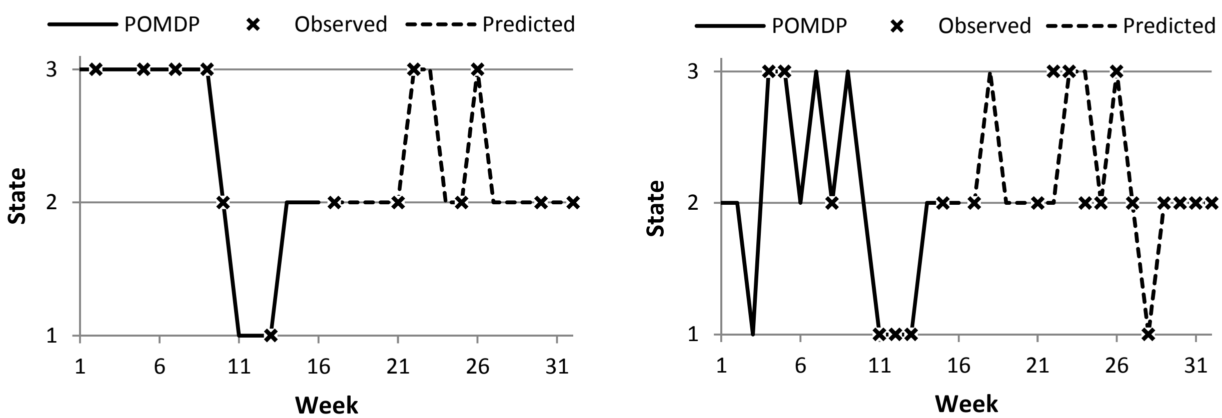


Figure 5: Posterior Analysis with Partially Observed Health State

7.3. Robustness Check

Various attempts were made to check the robustness of our results. First, we checked the robustness of observed health conditions. In our dataset, the observed functionality level was scaled from 0 to 100. We performed various classifiers to categorize (or discretize) functionality levels. The alternative trials did not produce qualitatively different results, and the likelihood does not exceed the result presented earlier. In addition, we have conducted the following analyses.[7]

Exogeneity of random effects ξ and η. To verify that the random effects are exogenous, or uncorrelated with the covariates, we follow Wooldridge (2001) to apply a variation (Mundlak 1978) of

[7] We thank an anonymous reviewer to suggest these analyses.

the Chamberlain device (Chamberlain 1980). Explicitly, we write $\xi_i = \xi^1 \overline{x}_i + \xi_i^0$ and $\eta_i = \eta^1 \overline{y}_i + \eta_i^0$ where $\overline{x}_i = T^{-1} \sum_{t=1}^{T} x_{it}$ and $\overline{y}_i = T^{-1} \sum_{t=1}^{T} y_{it}$ are vectors of means of covariates; ξ^1 and η^1 are vectors of coefficients to be estimated; and ξ_i^0 and η_i^0 are handled non-parametrically as before. The estimation results, as presented in the online supplement, show that ξ^1 and η^1 are insignificant, with the exception of membership. While the exogeneity condition is not strictly satisfied, relaxing it with the Chamberlain device does not result in significantly different findings. Hence, we retain the original model which is more parsimonious.

Stationarity of negative binomial distribution (NBD). We model the dependent variable, number of new posts, using NBD which is stationary (Morrison and Schmittlein 1988). In general, patients' online behaviors vary over time. To account for this non-stationary nature, we allow the mean of NBD to change from one period to another. Therefore, the overall process is stationary within a period but non-stationary across periods. As the length of time period reduces, this approximation becomes a more accurate presentation. We recalculated the values of covariates using half week and 2 weeks as the time period, respectively, and re-conducted the analysis. The results showed no significant difference. The prediction accuracies are 93.1% and 92.7%, respectively. Hence, we chose a week as the length of time between observations, to coincide with the practice of the online healthcare community which routinely asked patients to update their profiles weekly.

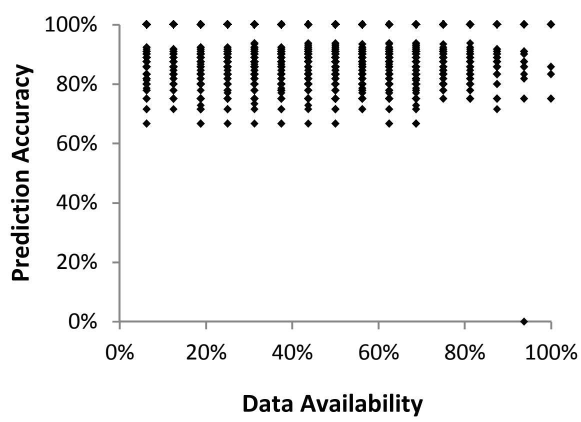


Figure 6: Effect of data availability on prediction accuracy

Effects of missing data. For each patient, we calculated the percentage of data availability, for example, if a patient reported her conditions in 8 weeks out of 16, her data availability is 50%. Figure 6 is a scatterplot of prediction accuracy on data availability for all 7512 patients in the data set. There is no apparent pattern and the correlation between the two is insignificant at –0.0069.

To further analyze the effects of missing data on prediction accuracy, we created stratified subsamples (around 1000 patients) based on the percentage of observed health states. Three subsamples have 13.37%, 21.91%, and 34.73% of available health states. Since there is a very low correlation of missing pattern between the 1^{st} 16 weeks data and the 2^{nd} (holdout), the data availability for the holdout is almost the same (around 37%) for three subsamples. The prediction accuracies are 86.59%, 87.21%, and 89.18%, which do not differ significantly from each other.

More analyses, as reported in the OLS, show the patterns of missing data is not MCAR (missing completely at random). It also appears that some patients report their health information when they are under certain health conditions; hence, it is unlikely that MAR (missing at random) is true. We should note our POMDP framework does not require either of the above conditions to hold. Similarly, we created stratified subsamples based on the variability of data availability across health states which was operationalized using coefficient of variation (COV). For example, in the case of three health states, if a patient reported (2, 2, 2) times (a total of 6 times out of 16), the COV is zero. The missing of data for her is independent of her health conditions and hence is more likely to be MAR. The other extreme is (6, 0, 0) which is not MAR. The two subsamples have the COV of 0.40 and 1.68 respectively. However, the prediction accuracies are virtually identical, 89.09% and 89.52%. Therefore, we conclude that our POMDP model is robust with respect to the missing data issues.

7.4. Limitations

There are several limitations in our study, and we do not want to overstate our findings. First, in contrast with qualitative study we conduct quantitative examine on the helpfulness of social support. In our findings, social supports are shown to have significant impact on patients' health condition changes, but our dataset does not allow us to distinguish between active and passive social supports. Instead of considering social support as a discrete time-limit act with immediate or delayed effects (King et al. 2006), we take the different perspective by taking the flow of support and evolving meaning of support over time. Although we are unable to separate "providing" from "seeking" social supports and, hence, cannot precisely measure the impact of each. The reciprocal aspect better fits the context of online healthcare community. Second, we use the number of posts as the measure for patients' online healthcare community outcomes. It is very possible, however, that patients possess different preferences in their online activities. For example, some patients may spend more time observing, rather than actively participating in others' communications. It would be helpful to incorporate more measures of patients' online behavior patterns. The third limitation is that we only considered direct communications among patients. Social supports can, however, also be transferred by word-of-mouth via common friends. Therefore, including other network measures could potentially shed more light on how online healthcare community can benefit patients.

8. Conclusion and Implications

In this paper, we developed a Partially Observed Markov Decision Process model to study patients' dynamic health condition outcomes. The POMDP model was estimated by a maximum likelihood procedure. Three health condition states were identified to best explain the data. Our results offered several insights into the driving forces behind patients' health condition changes and, hence, demonstrated the usefulness and value of online healthcare communities.

Despite the sizeable body of research on social support and various findings confirming the positive impact of social support on individuals' health conditions, there is less attention drawn on the magnitude and contrast of these impacts. Especially, patents' online behavior pattern and the impact of such online activities to their health is not fully understood yet. Although the utilization of online healthcare communities (for example, emotional support and information sharing) is thought to be positive to patients' health, there is also a need for more careful explanation of the processes by which support and lack of support might affect patients' health outcome. Thus, the main contributions of this work are (1) our proposed framework to measure how helpful an online social network can be and (2) new evidence of the efficiencies and benefits of such online services. Growing participation in online healthcare communities is well documented. While research on how these forms of social networking work and how well they serve patients' needs is underway, the important question of how social supports change patients' health outcomes remains unanswered (Lamberg 2003). By investigating the online activities of patients suffering from mental disease, our results revealed the benefits and advantages of online healthcare communities in helping patients improve their health conditions.

Our procedure to identify a patient's unseen health condition distinguishes our model from other social networking studies on healthcare. Extending the sociological research on patient behavior, we used the POMDP model to explain patients' health condition changes with respect to social supports they received online. Patients are actively involved in disease self-management. Their participation in online discussions enabled them to learn from other patients, and enjoy a partial prevention effect that reduced the possibility of their condition deteriorating. These findings can be used to encourage users who are passively participating in online healthcare communities to reduce "lurking" behaviors. That would result in online healthcare communities becoming a place where social support are contributed by an even more diverse membership. The investigation of transition distributions for various effects revealed that such communications were more effective for patients in good health conditions. We showed that a healthier patient gains more benefits from online healthcare community and has a higher probability of staying well.

We found, in our empirical analysis, that informational support was the most useful of all available online social supports. It was the main draw for patients and their family members to join an online healthcare community. However, its impact on changing patients' health condition was relatively lower than that of emotional support. Our results also indicated that recognition and positive feedback from other patients helped to improve an individual's health condition and encouraged patients to play their social roles competently. This effect was enhanced in the "sticky" dormant states.

Finally, our work is just a first step. It revealed the importance of studying the role of information systems in the context of healthcare. Fichman and his colleagues (2011) agree: The intersection of social media and healthcare is a promising direction for study. Our work combined theoretical modeling and data validation and exhibited the quantitative results. These findings signify a potential direction for healthcare reform, and suggest the effective and encouraging consequences of incorporating patients' self-assistance efforts into health management. These possibilities are tantalizing for both information systems and healthcare practices research.

References

Abbott, S., D. Freeth. 2008. Social Capital and Health: Starting to Make Sense of the Role of Generalized Trust and Reciprocity. *Journal of Health Psychology* **13**(7) 874-883.

Agarwal, R., Gao, G, DesRoches, C., Jha, A.K. 2010. The Digital Transformation of Healthcare: Current Status and the Road Ahead. *Inform. Sys. Res.* **21** (4) 796-809

Anderson, D.R., K.P. Burnham, G.C. White. 1998. Comparison of AIC and CAIC for model selection and statistical inference from capture-recapture studies. *Journal of Applied Statistics* **25** 263 – 282.

Antonucci, T.C., J.S. Jackson. 1990. The role of reciprocity in social support. In B.R. Sarason, I.G. Sarason, G.R. Pierce (Eds.), *Social support: An interactional view* (pp. 173-198). New York: Wiley.

Bambina, A.D. 2007. Online Social Support: The interplay of Social Networks and Computer-Mediated Communication.

Bandura, A. 2004. Health Promotion by Social Cognitive Means. *Health Education & Behavior* **31**(2) 143-164.

Benkler, Y. 2002. Coase's penguin, or, Linux and the nature of the firm. *Yale Law J.* **112**(3) 369–446.

Berkman, L.A., L. Breslow. 1984. Health and Ways of Living: The Alameda County Studies. *The Journal of Ambulatory Care Management* **7**(1) 80.

Berkman, L.F., T. Glass, I. Brissette, T.E. Seeman. 2000. From social integration to health: Durkheim in the new millennium. *Social Science and Medicine* **51** 843-857.

Chamberlain, G. 1980. Analysis of Covariance with Qualitative Data. *Review of Economic Studies* XLVII 225-238.

Chernomas, W.M, Clarke, D.E. Social Support and Women Living with serious Mental Illness. *Project #23 of the Prairie Women's Health Centre of Excellence* **ISBN 0-9689692-4-9**.

Christensen H, Griffiths K. 2000. The Internet and mental health literacy. *Aust N Z J Psychiatry* **34**(6) 975-9.

Clark, C.M. 2006. Relations Between Social Support and Physical Health. Working paper, Rochester Institute of Technology.

Cobb, S. 1976. Social Support as a Moderator of Life Stress. *Psychosomatic Medicine* **38**(5).

Cohen, S., T.A. Wills. 1985. Social Support, Stress and the Buffering Hypothesis. *Psychological Bulletin* **98** 310-357.

Coulson, N.S. 2005. Receiving social support online: an analysis of a computer-mediated support group for individuals living with irritable bowel syndrome. *CyberPsychology & Behaviour* **8** 580–584.

Eriksson, E., S. Lauri. 2000. Informational and emotional support for cancer patients' relatives. *European Journal of Cancer Care*. 9: 8–15. doi: 10.1046/j.1365-2354.2000.00183.x

Fichman R. G., Kohli, R., Krishnan, R. and Kane, G.C. 2011. The Role of Information Systems in Healthcare: Current Research and Future Trends, lead article by the senior editors in *Information Systems Research*, special issue on Healthcare IT.

Griffiths, F., Cave J., Boardman, F., Ren, J., Pawlikowska, T., Ball, R. Clarke, A., Cohen, A. 2012. Social networks – the future for health care delivery. *Social Science and Medicine* **75** 2233-2241.

Hamilton, J.D. 1989. A New Approach to the Economic-Analysis of Nonstationary Time-Series and the Business-Cycle. *Econometrica* **57**(2) 357-384.

Hauser, J.R., G.L. Urban, G. Liberali, M. Braun. 2009. Website Morphing. *Marketing Sci.* **28** 202 - 223.

Heckman, J., Singer, B., 1984. A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data. *Econometrica* **52**(2) 271-320

Hunt, S.M., S.P. Mckenna, J. Mcewen, J. Williams, E. Papp. 1981. The Nottingham Health Profile – Subjective Health-Status and Medical Consultations. *Social Science and Medicine Part A-Medical Sociology* **15**(3) 221-229.

Jacobson, D.E., 1986. Types and Timing of Social Support. *Journal of health and Social Behavior* **27** 250-264.

Janakiraman, R., S. Dutta, C. Sismeiro, P. Stern. 2008. Physicians' Persistence and Its Implications for Their Response to Promotion of Prescription Drugs. *Management Science* **54**(6) 1080-1093.

Jou, Y.H., H. Fukada. 2002. Stress, Health, and Reciprocity and Sufficiency of Social Support: The Case of University Students in Japan. *Journal of Social Psychology* **142**(3) 353-370.

Jung, J. 1990. The role of reciprocity in social support. Basic & Applied Social Psychology **11** 243-253.

Kaelbling, L.P., M.L. Littman, A.R. Cassandra. 1998. Planning and acting in partially observable stochastic domains. *Artificial Intelligence* **101** 99-134.

Kane, G.C., R.G. Fichman, J. Gallaugher, J. Glaser. 2009. Community Relations 2.0. *Harvard Business Review*.

Kassirer, J.P. 2000. Patients, Physicians, and the Internet. *Health Affairs* **19**(6) 115-123.

King, G., Wiloughby, C., Specht, J.A., Brown, E. 2006. Social Support Processes and the Adaptation of Individuals With Chronic Disabilities. *Qualitative Health Research* **16**(7) 902-925.

Lamberg, L. 2003. Online Empathy for Mood Disorders-Patients Turn to Internet Support Groups. *American Medical Association* **289**(233) 3073.

Langford C.P., Bowsher J., Maloney J.P., Lillis P.P. 1997. Social support: a conceptual analysis. *J Adv Nurs*. **25**(1) 95-100.

Leung, L. 2011. Loneliness, social support, and preference for online social interaction: the mediating effects of identity experimentation online among children and adolescents. *Chinese Journal of Communication* **4**(4) 381-399.

McCorkle, B.H., Rogers, E. S., Dunn, E.C., Lyass, A., Wan, Y.M. 2008. Increasing Social Support for Individuals with Serious Mental Illness: Evaluating the Compeer Model of Intentional Friendship. *Community Mental Health Journal* **44** 359-366.

Mcmullan, M. 2006. Patients using the Internet to obtain health information: How this affects the patient-health professional relationship. *Patient Education and Counseling* **63** 24-28.

McPherson, M., L. Smith-Lovin, J.M. Cook. 2000. Birds of a Feather: Homophily in Social Networks? *Annu. Rev. Sociol.* **27** 415-44.

Merton, R.K. 1976. *Sociology Ambivalence and Other Essays*. Free Press, New York.

Midlarsky, E. 1991. Helping as Coping. *Prosocial Behavior: Review of Personality and Social Psychology* **12** 238-264.

Morrison, D.G., D.C. Schmittlein. 1988. Generalizing the NBD Model for Customer Purchases: What Are the Implications and Is It Worth the Effort? *J. Business and Economic Statistics* **6** 145–159.

Mundlak, Y. 1978. On the Pooling of Time Series and Cross Section Data. *Econometrica* **46** 69-85.

Netzer, O., J.M. Lattin, V. Srinivasan. 2008. A Hidden Markov Model of Customer Relationship Dynamics. *Marketing Science* **27**(2) 185-204.

O'Grady, L.A., Witteman, H., Wathen, C.N. 2008. "The experiential health information processing model: supporting collaborative web-based patient education," *BMC Medical Informatics and Decision Making* **8**:58.

Oluwole, D.A., Hammed, A.T., Awaebe, J. 2011. Patterns of Stress, Social Spoort, and Mental Health among Nigerian Women. *Advancing Women in Leadership* AdvancingWomen.com.

Pearlin, L.I. 1985. Social Structure and Processes of Social Support. *Social Support and Health* 43-60.

Post, S.G. 2005. Altruism, Happiness, and Health: It's Good to Be Good. International Journal of Behavioral Medicine **12**(2) 66-77.

Powell, J., Clarke, A. 2006. Internet information-seeking in mental health: Population survey. *British Journal of Psychiatry* **189** 273-277

Putnam, R.D. 1993. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton, NJ: Princeton University Press.

Rabiner, L.R. 1989. A Tutorial on Hidden Markov-Models and Selected Applications in Speech Recognition. *Proceedings of the IEEE* **77**(2) 257-286.

Rabiner, L.R., B.H. Juang, S.E. Levinson, M.M. Sondhi. 1986. Recognition of Isolated Digits Using Hidden Markov-Models with Continuous Mixture Densities. *AT&T Technical Journal* **64**(6) 1211-1222.

Radley, A., M. Billig. 1996. Accounts of health and illness: Dilemmas and representations. *Sociology of Health & Illness* **18**(2) 220-240.

Schwartz, C., J.B. Meisenhelder, Y. Ma, G. Reed. 2003. Altruistic Social Interest Behaviors Are Associated with Better Mental Health. *Psychosomatic Medicine* **65** 778-785.

Shumaker, S.A., Bronwell, A. 1984. Toward a theory of social support: Closing conceptual gaps. *Journal of Social Issues* **40**(4) 11-33.

Singh, P.V., Y. Tan, N. Youn. 2011. A Hidden Markov Model of Developer Learning Dynamics in Open Source Software Projects. *Information Systems Research*. Forthcoming.

Slevin, M.L., S.E. Nichols, S.M. Downer, P. Wilson, T.A. Lister, S. Arnott, J. Maher, R.L. Souhami, J.S. Tobias, A.H. Goldstone, M. Cody. 1996. Emotional support for cancer patients: what do patients really want? *British Journal of Cancer* **74** 1275-1279.

Swan, M. 2009. Emerging Patient-driven Health Care Models: An Examination of Health Social Networks, Consumer Personalized Medicine and Quantified Self-Tracking. *Int. J. Environ. Res. Public Health* **6** 492-525.

Tessler, R., D. Mechanic. 1978. Psychological Distress and Perceived Health Status. *J. Health and Social Behavior* **19**(3) 254-262.

Wanless, D. 2002. Securing our future health: taking a long-term view. London: HM Treasury.

Weinberg, N., Schmale, J.D., Uken, J., Wessel, K. 1995. Computer-mediated support groups. *Social Work with Groups* **17** 43–54.

Weiss, R.S. 1974. The provisions of social relationships. In Z. Rubin (Ed.) Englewood Cliffs, NJ: Prentice-Hall.

Wellman, B., Wortley, S. 1990. Different Strokes from Different Folks: Community Ties and Social Support. *The American Journal of Sociology* **96**(3) 558-88.

Wills, T.A. 1985. Supportive functions of interpersonal relationships. In S. Cohen & S.L. Syme (Eds.) *Social support and health*. New York: Academic Press.

Wooldrige, J.M. 2001. *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

Wortman C. & Conway T. 1985. The role of social support in adaptation and recovery in physical illness. *Social Support and Health* Wortman C. & Conway T. eds. Academic Press, New York.

Wright, K. 2000. Computer-Mediated Social Support, Older Adults, and Coping. *Journal of Communication* **50** (3) 100-118.

Ziebland, S., A. Chapple, C. Dumelow, J. Evans, S. Prinjha, L. Rozmovits. 2004. How the internet affects patients' experience of cancer: a qualitative study. *British Medical Journal* **328** 564-569.

Online Supplement to

Feel Blue so Go Online: An Empirical Study of Social Support among Patients

A1. Statistics for Testing Hypothesis 3

To test Hypothesis 3, we check whether the estimated coefficient for informational ($\beta_{7,s}$) or emotional ($\beta_{8,s}$) support varies across states. For example, we test whether the effect of informational support in state 1 is significantly different from that in state 2, that is, $\beta_{7,1} - \beta_{7,2}$.

	Parameter Difference between State					
	1 (bad) and 2 (fair)		1 (bad) and 3 (good)		2 (fair) and 3 (good)	
info. support β_7	0.1514^{***}	(0.0576)	0.1419^{*}	(0.0428)	–0.0095	(0.0665)
emo. support β_8	-0.1964^{***}	(0.0377)	-0.4587^{***}	(0.0475)	-0.2623^{***}	(0.0615)

Note: Standard errors in parentheses. Significance level: *** 1%; * 10%.

A2. Comparison of Intrinsic Propensity to Contribute across States

Here, we test the constant coefficient $\gamma_{0,s}$ is different across state *s*. The following table shows the differences of these estimated parameters and their standard errors.

Parameter Difference between State					
1 (bad) and 2 (fair)		1 (bad) and 3 (good)		2 (fair) and 3 (good)	
0.5647^{**}	(0.2300)	0.8845^{***}	(0.1570)	0.3198^{***}	(0.1331)

Note: Standard errors in parentheses. Significance level: *** 1%; ** 5%.

A3. Comparison of Effect of Emotional Support and Informational Support

To test which type of support is better for meeting patients' social needs, we compare the effect of emotional support ($\beta_{8,s}$) and informational support ($\beta_{7,s}$) for a given state *s*. Since we have rescaled the covariant for informational support by dividing a factor of 10, $\beta_{7,s}$ measures the effect of 10 units of informational support while $\beta_{8,s}$ measures 1 unit of emotional support. For state 1, we compare $\beta_{8,s}$ with $\beta_{7,s}/2$ (or, the effect of 5 units of informational support).

Parameter Difference between Emotional and Information Support		
State 1 (bad) $\beta_{8,1}-(\beta_{7,1}/2)$	State 2 (fair) $\beta_{8,2}-\beta_{7,2}$	State 3 (good) $\beta_{8,3}-\beta_{7,3}$
0.2416*** (0.0756)	0.2075*** (0.0657)	0.4603*** (0.0427)

Note: Standard errors in parentheses. *** Significant at 1%.

A4. Calculation of State Transition Probabilities

The support points and probability masses for individual heterogeneity in the state transition are:

Support Point ξ_k	0	0.4134	1
Probability $P(\xi_k)$	0.4715	0.3446	0.1839

The transition thresholds are:

$\overline{\omega}_{1\to2}$	$\overline{\omega}_{1\to3}$	$\overline{\omega}_{2\to3}$	$\underline{\omega}_{2\to1}$	$\underline{\omega}_{3\to1}$	$\underline{\omega}_{3\to2}$
0.9227	3.1593	2.6481	−2.5327	−3.0126	−1.0469

The intrinsic propensity to transition from i to j (Table 8) is calculated as follows with the rescaling parameter $C_\xi = -0.151$. Specifically, for $i=1, j=2$:

$$P_{1,2} = \sum_{k=1}^{3}\left(\frac{\exp\left(\overline{\omega}_{1\to3} - C_\xi \xi_k\right)}{1+\exp\left(\overline{\omega}_{1\to3} - C_\xi \xi_k\right)} - \frac{\exp\left(\overline{\omega}_{1\to2} - C_\xi \xi_k\right)}{1+\exp\left(\overline{\omega}_{1\to2} - C_\xi \xi_k\right)}\right)P(\xi_k) = 0.2357;$$

for $i=1, j=3$:

$$P_{1,3} = \sum_{k=1}^{3}\left(1-\frac{\exp\left(\overline{\omega}_{1\to3} - C_\xi \xi_k\right)}{1+\exp\left(\overline{\omega}_{1\to3} - C_\xi \xi_k\right)}\right)P(\xi_k) = 0.0389;$$

for $i=2, j=1$:

$$P_{2,1} = \sum_{k=1}^{3}\left(\frac{\exp\left(\underline{\omega}_{2\to1} - C_\xi \xi_k\right)}{1+\exp\left(\underline{\omega}_{2\to1} - C_\xi \xi_k\right)}\right)P(\xi_k) = 0.0771;$$

for $i=2, j=3$:

$$P_{2,3} = \sum_{k=1}^{3}\left(1-\frac{\exp\left(\overline{\omega}_{2\to3} - C_\xi \xi_k\right)}{1+\exp\left(\overline{\omega}_{2\to3} - C_\xi \xi_k\right)}\right)P(\xi_k) = 0.0632;$$

for $i=3, j=1$:

$$P_{3,1} = \sum_{k=1}^{3} \left(\frac{\exp\left(\underline{\omega}_{3\to1} - C_{\xi}\xi_k\right)}{1+\exp\left(\underline{\omega}_{3\to1} - C_{\xi}\xi_k\right)} \right) P\left(\xi_k\right) = 0.0492;$$

and for $i = 3, j = 2$:

$$P_{3,2} = \sum_{k=1}^{3} \left(\frac{\exp\left(\underline{\omega}_{3\to2} - C_{\xi}\xi_k\right)}{1+\exp\left(\underline{\omega}_{3\to2} - C_{\xi}\xi_k\right)} - \frac{\exp\left(\underline{\omega}_{3\to1} - C_{\xi}\xi_k\right)}{1+\exp\left(\underline{\omega}_{3\to1} - C_{\xi}\xi_k\right)} \right) P\left(\xi_k\right) = 0.2204.$$

A5. Changes of State Transition Probabilities

The transition probabilities are evaluated with the average value of the focal variable and the values of the other variables set at zero, and compared with the intrinsic transition probabilities as presented in Table 7 in the paper.

Table A5.1: Probability Change – views

	Bad	Fair	Good
Bad	–0.0289	0.0233	0.0056
Fair	–0.0092	–0.0006	0.0086
Good	–0.0048	–0.0158	0.0206

Table A5.2: Probability Change – thank you

	Bad	Fair	Good
Bad	–0.0140	0.0113	0.0027
Fair	–0.0019	0.0003	0.0016
Good	–0.0043	–0.0143	0.0187

Table A5.3: Probability Change – comments

	Bad	Fair	Good
Bad	–0.0294	0.0237	0.0057
Fair	–0.0029	0.0004	0.0025
Good	–0.0035	–0.0115	0.0151

Table A5.4: Probability Change – In-Degree

	Bad	Fair	Good
Bad	0.0175	–0.0143	–0.0032
Fair	–0.0013	0.0002	0.0011
Good	0.0025	0.0080	–0.0105

Table A5.5: Probability Change – Posts

	Bad	Fair	Good
Bad	–0.0132	0.0107	0.0025
Fair	–0.0035	0.0005	0.0031
Good	–0.0054	–0.0178	0.0232

Table A5.6: Probability Change – Usefulness

	Bad	Fair	Good
Bad	–0.0169	0.0137	0.0032
Fair	–0.0021	0.0003	0.0018
Good	–0.0057	–0.0190	0.0247

Table A5.7: Probability Change – Out-Degree

	Bad	Fair	Good
Bad	–0.0134	0.0108	0.0026
Fair	–0.0013	0.0002	0.0011
Good	–0.0022	–0.0073	0.0095

A6. Expected Intrinsic Number of New Posts

The support points and probability masses for individual heterogeneity in the outcome model are:

Support Point η_k	0	0.3174	0.5313	1
Probability $G(\eta_k)$	0.0639	0.3487	0.4015	0.1859

The rescaling parameter is $C_\eta = -0.204$. The expected number of new posts for state 1:

$$\sum_{k=1}^{4} \exp\left(\gamma_{0,1} + C_\eta \eta_k\right) G(\eta_k) = 2.75;$$

for state 2:

$$\sum_{k=1}^{4} \exp(\gamma_{0,2} + C_\eta \eta_k) \Pr(\eta_k) = 1.56;$$

for state 3:

$$\sum_{k=1}^{4} \exp(\gamma_{0,3} + C_\eta \eta_k) \Pr(\eta_k) = 1.13.$$

A7. Characteristics of Missing Data on Health Conditions

As mentioned in the paper, we partition the data into two sets. We use the first 16 weeks to estimate the model parameters, and the second 16 weeks data is the holdout to assess the accuracy of prediction. Figure A1 plots the distribution for the percentage of unobserved health states among patients, for example, roughly 10% patients did not report their health conditions 50% to 60% of the time. Overall, we observe 46.09% of observed states for the first set, and 37.5% for the second. Overall, there is a significant drop over time as patients are less willing to provide their health information, and the average of the 3 years during which we observed the website is around 11% only. We find that the correlation of patients missing data between the two data sets is insignificant, at 0.0053, indicating over time the missing data behavior is random.

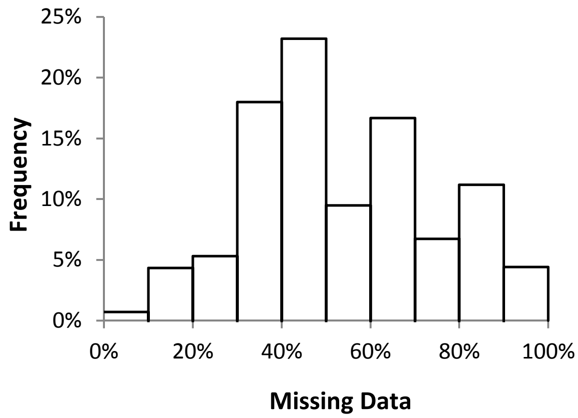


Figure A1: Distribution of missing data

We ran a logistic regression of the binary variable, $avail_{it}$ (= 1 if patient i reported her health condition at period t; 0 otherwise) on all covariates used in our study as well as the dependent variable: new posts. The results are presented in the following table, from which we can conclude that the missing data is NOT MCAR (missing completely at random).

new posts	–0.0021	(0.0050)
constant	-1.1430^{***}	(0.0415)
views	–0.0041	(0.0261)
thank you	–0.0849	(0.0666)
comments	–0.0692	(0.0634)
usefulness	–0.0077	(0.0391)
in_degree	-0.1762^{***}	(0.0325)
out_degree	-0.1792^{***}	(0.0343)
info. support	0.1493^{**}	(0.0683)
emo. support	0.2283^{*}	(0.1292)
posts	-1.1223^{***}	(0.1911)
# treatment	–0.0041	(0.0532)
# sympt	0.0091	(0.0673)
gender	–0.0913	(0.2389)
info quality	0.1015	(0.1273)
membership	0.3500^{***}	(0.0155)
update	-0.4532^{*}	(0.1433)

We have also examined the pattern of missing data with respect to the health condition of a patient. It is observed that patients are heterogeneous in that some reported their health information regardless of their health conditions, while others were more selective. Hence, MAR (missing at random) did not occur either.

We should note that our POMDP framework does not require either (MAR or MCAR) to work. However, we have conducted more analyses with sampled data to examine the effects of missing data on prediction accuracy.

A8. Robustness Check

We have conducted 3 sets of robustness check.

1. *Exogeneity of random effects*. Table A8.1 reports the results of POMDP with Chamberlain device.
2. *Stationarity of negative binomial distribution*. Tables A8.2-1 and A8.2-2 present the estimation results for the time period of half-week or two weeks.
3. *Effects of missing data*. We created stratified samples from 7512 patients based on two criteria: percentage of observed states or coefficient of variation (COV) in the 1st 16 weeks data. The following subsamples have the percentage of missing data at three levels. Due to close to zero correlation (0.0053) between the 1st and 2nd sets, all other measures remain similar.

Sample Characteristics						
Subsample	Size (Number of patients)	Percentage of Observed States		Coefficient of Variation (COV)		Prediction Accuracy
		1st Set	2nd Set	1st Set	2nd Set	
8.3-1	1171	**13.37%**	36.83%	1.28	0.73	**86.59%**
8.3-2	1000	**21.91%**	36.89%	1.03	0.73	**87.21%**
8.3-3	1252	**34.73%**	38.02%	1.05	0.73	**89.18%**

The following two subsamples are used to examine the effects of MAR. The COV is calculated across health conditions. A low COV shows that a patient reports her health information independently of her health condition at the time, hence indicates a higher degree of MAR.

Sample Characteristics						
Subsample	Size (Number of patients)	Percentage of Observed States		Coefficient of Variation (COV)		Prediction Accuracy
		1st Set	2nd Set	1st Set	2nd Set	
8.3-4	1209	52.26%	38.08%	**0.40**	0.73	**89.09%**
8.3-5	1285	28.45%	37.17%	**1.68**	0.72	**89.52%**

Table A8.1: Results for the Three-State POMDP with Chamberlain Device

Parameter	State 1 (bad)	State 2 (fair)	State 3 (good)	Mean of Covariates
θ Dispersion	0.6783***(0.1732)	0.4234***(0.0747)	0.5350***(0.1744)	
Variables Impacting State Transition				
β_1 [views]	0.8761***(0.1626)	0.9545** (0.4743)	0.7533***(0.1342)	**0.7916 (0.6518)**
β_2 [thank you]	1.1501***(0.1183)	0.6753***(0.2074)	2.6453***(0.3344)	**1.2115 (0.7414)**
β_3 [comments]	3.2374***(0.1927)	0.7433***(0.2146)	1.4523** (0.6986)	**1.1599 (0.9482)**
β_4 [usefulness]	0.8512***(0.2734)	0.5423***(0.1708)	1.3464***(0.1324)	**0.8476 (0.7458)**
β_5 [in-degree]	–1.1843***(0.2743)	0.2745** (0.1357)	–0.7543***(0.2235)	**–0.9031 (0.7884)**
β_6 [out-degree]	0.9023***(0.1574)	0.1975* (0.1173)	0.5975***(0.1874)	**0.2960 (0.2927)**
β_7 [info. support]	0.8935***(0.0432)	0.1467***(0.0745)	0.5674* (0.3231)	**0.7684 (0.5437)**
β_8 [emo. support]	0.6823***(0.0927)	0.5474***(0.1025)	0.7634***(0.0954)	**0.6059 (0.5352)**
β_9 [posts]	1.6422***(0.1965)	1.9644***(0.2087)	4.4325***(0.2346)	**1.0981 (0.6997)**
β_{10} [# treatment]	0.9515***(0.1321)	0.6543 (0.4667)	0.3456 (0.5465)	**0.7161 (0.5827)**
β_{11} [# sympt]	–0.8404***(0.0865)	–0.9746***(0.3345)	0.2987***(0.0753)	**–0.8551 (0.6593)**
Thresholds				
State 1 (bad)		0.6823***(0.1642)	3.3941***(0.5894)	
State 2 (fair)	–2.2745***(0.6680)		2.8483***(0.4026)	
State 3 (good)	–3.2936***(0.7275)	–1.1898***(0.1903)		
Variables Impacting State Dependent Outcome				
γ_0 [constant]	0.8747***(0.2877)	0.5654***(0.1678)	0.2185***(0.0324)	
γ_1 [gender]	0.7254***(0.2014)	0.8454 (0.5343)	0.5672** (0.2872)	
γ_2 [info quality]	0.1212 (0.0765)	0.3556 (0.2745)	0.1653***(0.0238)	**0.2190 (0.2035)**
γ_3 [membership]	0.6542***(0.0065)	0.5444***(0.0025)	0.9012***(0.0022)	**0.5166***(0.0065)**
γ_4 [update]	–0.6354***(0.1245)	0.6854* (0.4014)	–0.8533***(0.1649)	**–0.8578 (0.7962)**

Unobserved Heterogeneity (η , ξ)

$C_\eta = -0.204, C_\xi = -0.152$

	$\eta_1 = 0$	$\eta_2 = 0.3208$	$\eta_3 = 0.5215$	$\eta_4 = 1$
Probability $G(\eta)$	0.0627	0.3488	0.4037	0.1848
Conditional Distribution: $H(\xi \mid \eta)$				
$\xi_1 = 0$	0.9324	0.2132	0.1469	0.9147
$\xi_2 = 0.4135$	0.0371	0.5826	0.1775	0.0431
$\xi_3 = 1$	0.0305	0.2042	0.6756	0.0422

Table A8.2-1: Results for the Three-State POMDP with Half-Weekly Data

Parameter	State 1 (bad)		State 2 (fair)		State 3 (good)	
θ Dispersion	0.6917***	(0.1645)	0.4183***	(0.0336)	0.5817***	(0.1901)
Variables Impacting State Transition						
β_1 [views]	0.9871***	(0.1863)	1.0283***	(0.2384)	0.8182***	(0.1717)
β_2 [thank you]	1.5890***	(0.2055)	0.7213***	(0.1718)	2.1875***	(0.3974)
β_3 [comments]	2.7265***	(0.1358)	0.9074***	(0.3038)	1.7357**	(0.7183)
β_4 [usefulness]	1.2918***	(0.3658)	0.3465**	(0.1436)	1.3864***	(0.1092)
β_5 [in-degree]	–1.3726***	(0.2982)	0.2837*	(0.1642)	–0.7018***	(0.2701)
β_6 [out-degree]	0.8843***	(0.1837)	0.2816	(0.1843)	0.6074***	(0.1548)
β_7 [info. support]	0.8032***	(0.1656)	0.5921***	(0.1652)	0.6592***	(0.0922)
β_8 [emo. support]	0.7129***	(0.1417)	0.7929***	(0.2077)	0.9813***	(0.2734)
β_9 [posts]	2.6265***	(0.2018)	1.9563***	(0.1123)	4.3071***	(0.4045)
β_{10} [# treatment]	1.3821***	(0.1754)	0.5837***	(0.2098)	0.9368***	(0.2989)
β_{11} [# sympt]	–0.9474***	(0.0937)	–0.9185***	(0.2853)	0.2654***	(0.0471)
Thresholds						
State 1 (bad)			0.9852***	(0.1184)	3.0736***	(0.6836)
State 2 (fair)	–2.4707***	(0.4193)			2.7163***	(0.4365)
State 3 (good)	–3.0046***	(0.5932)	–1.1947***	(0.1483)		
Variables Impacting State Dependent Outcome						
γ_0 [constant]	1.2465***	(0.3010)	0.6026***	(0.2038)	0.2645***	(0.0314)
γ_1 [gender]	0.7927***	(0.2186)	0.7876	(0.4875)	0.6923**	(0.2928)
γ_2 [info quality]	0.2548***	(0.0487)	0.4875	(0.2962)	0.1893***	(0.0703)
γ_3 [membership]	0.9273***	(0.0981)	0.6542	(0.4468)	1.2931**	(0.5837)
γ_4 [update]	–0.8267***	(0.1071)	–0.5098	(0.3991)	–0.9587***	(0.1293)

Unobserved Heterogeneity (η , ξ)

$C_\eta = -0.206, C_\xi = -0.154$

	$\eta_1 = 0$	$\eta_2 = 0.3182$	$\eta_3 = 0.5496$	$\eta_4 = 1$
Probability $G(\eta)$	0.0682	0.3513	0.3853	0.1952
Conditional Distribution: $H(\xi \mid \eta)$				
$\xi_1 = 0$	0.9398	0.2572	0.1279	0.9182
$\xi_2 = 0.4146$	0.0317	0.6061	0.2154	0.0415
$\xi_3 = 1$	0.0285	0.1367	0.6567	0.0403

Posterior Analysis

Prediction Accuracy **93.1%**

Table A8.2-2: Results for the Three-State POMDP with Bi-Weekly Data

Parameter	State 1 (bad)		State 2 (fair)		State 3 (good)	
θ Dispersion	0.5637***	(0.1953)	0.2827***	(0.1028)	0.3923**	(0.1954)
Variables Impacting State Transition						
β_1 [views]	0.7263***	(0.1372)	0.6732*	(0.3746)	0.5837***	(0.2027)
β_2 [thank you]	1.0621***	(0.1826)	0.4283**	(0.2138)	1.6352***	(0.5632)
β_3 [comments]	2.1834***	(0.1583)	0.4736**	(0.2316)	1.3923**	(0.7283)
β_4 [usefulness]	0.7533**	(0.3632)	0.3028**	(0.1473)	1.0182***	(0.2137)
β_5 [in-degree]	–0.7827***	(0.2718)	0.2394	(0.1472)	–0.5833**	(0.2748)
β_6 [out-degree]	0.6234***	(0.1933)	0.2723**	(0.1582)	0.4739***	(0.1578)
β_7 [info. support]	0.6258***	(0.2315)	0.4867**	(0.2052)	0.4962***	(0.1472)
β_8 [emo. support]	0.5247**	(0.2275)	0.7321***	(0.2698)	0.8431***	(0.3017)
β_9 [posts]	2.0281***	(0.2583)	1.3809***	(0.1992)	3.6490***	(0.7874)
β_{10} [# treatment]	1.0187***	(0.2381)	0.6213**	(0.3034)	0.7769**	(0.3145)
β_{11} [# sympt]	–0.7948***	(0.1592)	–0.8937**	(0.3412)	0.2091***	(0.0518)
Thresholds						
State 1 (bad)			0.7629***	(0.1823)	2.7188***	(0.7390)
State 2 (fair)	–2.1382***	(0.4608)			2.2523***	(0.4927)
State 3 (good)	–2.3744***	(0.7725)	–0.9175***	(0.1802)		
Variables Impacting State Dependent Outcome						
γ_0 [constant]	1.0044***	(0.3036)	0.4481**	(0.2148)	0.1825***	(0.0364)
γ_1 [gender]	0.7152***	(0.2021)	0.5837	(0.4792)	0.5825**	(0.2901)
γ_2 [info quality]	0.2197***	(0.0674)	0.3827	(0.2583)	0.1204*	(0.0728)
γ_3 [membership]	0.9725***	(0.0986)	0.5638	(0.4911)	1.0183	(0.7194)
γ_4 [update]	–0.7113***	(0.1295)	–0.4923	(0.4012)	–0.8214***	(0.1593)

Unobserved Heterogeneity (η , ξ)				
C_η = –0.206, C_ξ = –0.148				
	$\eta_1 = 0$	$\eta_2 = 0.3283$	$\eta_3 = 0.5071$	$\eta_4 = 1$
Probability $G(\eta)$	0.0915	0.4178	0.3923	0.0984
Conditional Distribution: $H(\xi \mid \eta)$				
$\xi_1 = 0$	0.9027	0.2126	0.1574	0.9031
$\xi_2 = 0.4362$	0.0625	0.6592	0.2133	0.0589
$\xi_3 = 1$	0.0348	0.1282	0.6293	0.0380
Posterior Analysis				
Prediction Accuracy	**92.7%**			

Table A8.3-1: Results for the Three-State POMDP with Sampled Data

Sample Characteristics					
Size (Number of patients)	Percentage of Observed States		Coefficient of Variation (COV)		Prediction Accuracy
	1st 16 Weeks	2nd 16 Weeks	1st 16 Weeks	2nd 16 Weeks	
1171	**13.37%**	36.83%	1.28	0.73	**86.59%**

Model Estimation						
Parameter	**State 1 (bad)**		**State 2 (fair)**		**State 3 (good)**	
θ Dispersion	0.3762**	(0.1535)	0.3056***	(0.0195)	0.3041**	(0.1212)
Variables Impacting State Transition						
β_1 [views]	0.4107***	(0.1565)	0.8319***	(0.2748)	0.9381***	(0.1635)
β_2 [thank you]	1.9873***	(0.1864)	0.3120***	(0.0706)	2.2781***	(0.2922)
β_3 [comments]	1.1965***	(0.0548)	0.9920***	(0.2126)	1.6075*	(0.9094)
β_4 [usefulness]	0.4044	(0.2816)	0.3827***	(0.0787)	0.5369***	(0.0927)
β_5 [in-degree]	–0.6067**	(0.2598)	0.1286**	(0.0510)	–0.6627***	(0.2474)
β_6 [out-degree]	0.3626**	(0.1632)	0.1100	(0.1692)	0.3322*	(0.1921)
β_7 [info. support]	0.4201**	(0.1639)	0.5935***	(0.0743)	0.4706***	(0.0863)
β_8 [emo. support]	0.2574***	(0.0440)	0.8138***	(0.1909)	0.7475***	(0.2707)
β_9 [posts]	2.5988***	(0.1746)	1.6231***	(0.1473)	2.5664***	(0.4303)
β_{10} [# treatment]	1.2087***	(0.0774)	0.5741	(0.3599)	0.8417***	(0.0877)
β_{11} [# sympt]	–0.3886***	(0.0581)	–0.7014***	(0.1594)	0.2976***	(0.0202)
Thresholds						
State 1 (bad)			0.9333***	(0.092)	2.9865***	(0.7966)
State 2 (fair)	–1.9421***	(0.3338)			1.5534***	(0.1358)
State 3 (good)	–2.2113***	(0.4453)	–0.3787***	(0.1301)		
Variables Impacting State Dependent Outcome						
γ_0 [constant]	1.2582***	(0.2978)	0.2463*	(0.1274)	0.2802***	(0.0192)
γ_1 [gender]	0.6649***	(0.2213)	0.5757**	(0.2657)	0.4780	(0.3468)
γ_2 [info quality]	0.1029	(0.0722)	0.5144*	(0.2991)	0.1776***	(0.0637)
γ_3 [membership]	1.0940***	(0.0654)	0.6353	(0.6104)	0.4422	(0.6189)
γ_4 [update]	–0.4155***	(0.1131)	–0.2304	(0.1419)	–0.9089***	(0.1357)

Unobserved Heterogeneity (η , ξ)				
$C_\eta = -0.317, C_\xi = -0.184$				
	$\eta_1 = 0$	$\eta_2 = 0.3046$	$\eta_3 = 0.5118$	$\eta_4 = 1$
Probability $G(\eta)$	0.0688	0.3764	0.4339	0.1209
Conditional Distribution: $H(\xi \mid \eta)$				
$\xi_1 = 0$	0.9371	0.2228	0.1991	0.9039
$\xi_2 = 0.4326$	0.0362	0.5827	0.2496	0.0498
$\xi_3 = 1$	0.0267	0.1945	0.5513	0.0463

Table A8.3-2: Results for the Three-State POMDP with Sampled Data

Sample Characteristics					
Size (Number of patients)	Percentage of Observed States		Coefficient of Variation (COV)		Prediction Accuracy
	1st 16 Weeks	2nd 16 Weeks	1st 16 Weeks	2nd 16 Weeks	
1000	**21.91%**	36.89%	1.03	0.73	**87.21%**

Model Estimation						
Parameter	**State 1 (bad)**		**State 2 (fair)**		**State 3 (good)**	
θ Dispersion	0.3310^{***}	(0.0867)	0.2444^{***}	(0.0145)	0.6139^{**}	(0.1947)
Variables Impacting State Transition						
β_1 [views]	0.8762^{***}	(0.1039)	0.4159^{**}	(0.2011)	0.5614^{***}	(0.0498)
β_2 [thank you]	1.7755^{***}	(0.1178)	0.6307^{***}	(0.0706)	2.1188^{***}	(0.1559)
β_3 [comments]	3.0349^{***}	(0.0916)	0.8022^{***}	(0.1007)	0.6495	(0.9624)
β_4 [usefulness]	1.1628^{***}	(0.2910)	0.1297^{***}	(0.0425)	1.0032^{***}	(0.0610)
β_5 [in-degree]	-0.9255^{***}	(0.1894)	0.3137^{***}	(0.1167)	-0.3074	(0.3237)
β_6 [out-degree]	0.5144^{***}	(0.1224)	0.1154	(0.1576)	0.1890^{***}	(0.1281)
β_7 [info. support]	0.5650^{***}	(0.0476)	0.7649^{***}	(0.1380)	0.4458^{***}	(0.0712)
β_8 [emo. support]	0.4453^{***}	(0.1399)	0.9521^{***}	(0.2524)	1.3001^{***}	(0.2594)
β_9 [posts]	1.1768^{***}	(0.0736)	1.2766^{***}	(0.1125)	2.2124^{***}	(0.4972)
β_{10} [# treatment]	0.4758^{***}	(0.0798)	0.5741^{**}	(0.2806)	0.6089^{***}	(0.2221)
β_{11} [# sympt]	-0.4338^{***}	(0.0340)	-1.3113^{***}	(0.1459)	0.2152^{***}	(0.0165)
Thresholds						
State 1 (bad)			0.3752^{***}	(0.0646)	2.4835^{***}	(0.6360)
State 2 (fair)	-2.6133^{***}	(0.3302)			2.2642^{***}	(0.2634)
State 3 (good)	-3.6376^{***}	(0.6986)	-0.5936^{***}	(0.0775)		
Variables Impacting State Dependent Outcome						
γ_0 [constant]	1.5698^{***}	(0.2197)	0.7165^{***}	(0.2043)	0.1890^{***}	(0.0238)
γ_1 [gender]	0.4347^{*}	(0.2249)	0.4226	(0.4915)	0.4443^{*}	(0.2532)
γ_2 [info quality]	0.1326^{**}	(0.0613)	0.1444	(0.2303)	0.2077^{**}	(0.0825)
γ_3 [membership]	0.4335^{***}	(0.0395)	0.5136	(0.3360)	0.9476^{**}	(0.4387)
γ_4 [update]	-0.7006^{***}	(0.1112)	-0.4550^{**}	(0.1829)	-0.8254^{***}	(0.0701)

Unobserved Heterogeneity (η , ξ)				
$C_\eta = -0.315, C_\xi = -0.186$				
	$\eta_1 = 0$	$\eta_2 = 0.3219$	$\eta_3 = 0.5475$	$\eta_4 = 1$
Probability $G(\eta)$	0.0624	0.3276	0.4540	0.1560
Conditional Distribution: $H(\xi \mid \eta)$				
$\xi_1 = 0$	0.9371	0.2161	0.2011	0.9220
$\xi_2 = 0.4327$	0.0351	0.5944	0.2421	0.0513
$\xi_3 = 1$	0.0278	0.1895	0.5568	0.0267

Table A8.3-3: Results for the Three-State POMDP with Sampled Data

Sample Characteristics

Size (Number of patients)	Percentage of Observed States		Coefficient of Variation (COV)		Prediction Accuracy
	1st 16 Weeks	2nd 16 Weeks	1st 16 Weeks	2nd 16 Weeks	
1252	**34.73%**	38.02%	1.05	0.73	**89.18%**

Model Estimation

Parameter	State 1 (bad)		State 2 (fair)		State 3 (good)	
θ Dispersion	0.6395***	(0.1364)	0.1702***	(0.0214)	0.2872*	(0.1543)
Variables Impacting State Transition						
β_1 [views]	0.6480***	(0.1127)	0.7949***	(0.1966)	0.7535***	(0.1635)
β_2 [thank you]	0.7819***	(0.1088)	0.6714***	(0.0772)	0.7291**	(0.3604)
β_3 [comments]	2.1011***	(0.0548)	0.3019	(0.1987)	1.0392**	(0.4926)
β_4 [usefulness]	0.3236	(0.3004)	0.1557	(0.1149)	1.0314***	(0.1103)
β_5 [in-degree]	–0.4113**	(0.2015)	0.1028	(0.0724)	–0.8882***	(0.0816)
β_6 [out-degree]	0.7167***	(0.1670)	0.1852***	(0.0549)	0.4926***	(0.1137)
β_7 [info. support]	0.3332***	(0.1243)	0.3304***	(0.0576)	0.6130***	(0.0592)
β_8 [emo. support]	0.2783***	(0.0451)	0.5452***	(0.0954)	1.1591***	(0.1843)
β_9 [posts]	1.2259***	(0.0907)	1.9331***	(0.0729)	4.3363***	(0.3468)
β_{10} [# treatment]	0.4372***	(0.1560)	0.6487***	(0.1304)	0.3403	(0.3428)
β_{11} [# sympt]	–0.6778***	(0.0350)	–1.1182***	(0.3268)	0.1259***	(0.0149)
Thresholds						
State 1 (bad)			0.7503***	(0.1367)	1.2260***	(0.4076)
State 2 (fair)	–1.8092***	(0.2035)			0.8688**	(0.3622)
State 3 (good)	–2.6203***	(0.3401)	–1.2178***	(0.0852)		
Variables Impacting State Dependent Outcome						
γ_0 [constant]	1.3181***	(0.2515)	0.3751***	(0.0881)	0.1325***	(0.0233)
γ_1 [gender]	0.8524***	(0.1161)	0.6048	(0.4605)	0.8213***	(0.2009)
γ_2 [info quality]	0.0869	(0.0653)	0.5775***	(0.1662)	0.1072**	(0.0510)
γ_3 [membership]	0.3406***	(0.0783)	0.4325	(0.3691)	1.6045**	(0.6430)
γ_4 [update]	–0.5866***	(0.1218)	–0.5010	(0.3397)	–0.9923***	(0.0445)

Unobserved Heterogeneity (η , ξ)

$C_\eta = -0.312, C_\xi = -0.181$

	$\eta_1 = 0$	$\eta_2 = 0.3028$	$\eta_3 = 0.5183$	$\eta_4 = 1$
Probability $G(\eta)$	0.0726	0.3799	0.3455	0.2020
Conditional Distribution: $H(\xi \mid \eta)$				
$\xi_1 = 0$	0.9465	0.2250	0.1931	0.8858
$\xi_2 = 0.4325$	0.0355	0.6002	0.2521	0.0493
$\xi_3 = 1$	0.0180	0.1748	0.5548	0.0649

Table A8.3-4: Results for the Three-State POMDP with Sampled Data

Sample Characteristics					
Size (Number of patients)	Percentage of Observed States		Coefficient of Variation (COV)		Prediction Accuracy
	1st 16 Weeks	2nd 16 Weeks	1st 16 Weeks	2nd 16 Weeks	
1209	52.26%	38.08%	**0.40**	0.73	**89.09%**

Model Estimation

Parameter	**State 1 (bad)**		**State 2 (fair)**		**State 3 (good)**	
θ Dispersion	0.8501***	(0.0654)	0.2095***	(0.0145)	0.4449**	(0.2186)
Variables Impacting State Transition						
β_1 [views]	0.4746***	(0.0488)	0.5823**	(0.2301)	0.9529***	(0.1308)
β_2 [thank you]	1.6452***	(0.1267)	0.7392***	(0.161)	0.8658***	(0.2208)
β_3 [comments]	1.2840***	(0.0438)	1.0437***	(0.1707)	1.3639***	(0.3259)
β_4 [usefulness]	0.7078***	(0.1940)	0.1103	(0.1636)	0.5934***	(0.0375)
β_5 [in-degree]	–0.7918***	(0.2064)	0.2442***	(0.0898)	–0.5807**	(0.2527)
β_6 [out-degree]	0.6577***	(0.0638)	0.1100	(0.1808)	0.5785***	(0.2049)
β_7 [info. support]	0.4201***	(0.1203)	0.2937*	(0.1578)	0.8050***	(0.0481)
β_8 [emo. support]	0.4871***	(0.0812)	0.6917***	(0.2397)	0.8558***	(0.0728)
β_9 [posts]	1.2504***	(0.1216)	0.7112***	(0.1331)	2.0354***	(0.2381)
β_{10} [# treatment]	1.4145***	(0.1402)	0.2833	(0.1955)	0.5283**	(0.2248)
β_{11} [# sympt]	–1.0121***	(0.0390)	–0.9962***	(0.1323)	0.2106***	(0.0500)
Thresholds						
State 1 (bad)			0.6771***	(0.0758)	2.6407***	(0.3335)
State 2 (fair)	–1.1056***	(0.3378)			1.9219***	(0.3622)
State 3 (good)	–3.0519***	(0.6677)	–0.4196**	(0.1874)		
Variables Impacting State Dependent Outcome						
γ_0 [constant]	0.9586***	(0.3411)	0.6662***	(0.1199)	0.1477***	(0.0277)
γ_1 [gender]	0.4773**	(0.2050)	0.7578***	(0.1993)	0.3905**	(0.1954)
γ_2 [info quality]	0.0983	(0.0602)	0.2301**	(0.0973)	0.1189***	(0.0416)
γ_3 [membership]	1.2592***	(0.0502)	0.8853**	(0.3786)	0.6822	(0.6370)
γ_4 [update]	–0.6029***	(0.0996)	–0.2476	(0.4405)	–0.9181***	(0.0612)

Unobserved Heterogeneity (η , ξ)

$C_\eta = -0.318, C_\xi = -0.183$

	$\eta_1 = 0$	$\eta_2 = 0.3230$	$\eta_3 = 0.5417$	$\eta_4 = 1$
Probability $G(\eta)$	0.0605	0.2962	0.4098	0.2335
Conditional Distribution: $H(\xi \mid \eta)$				
$\xi_1 = 0$	0.9558	0.2161	0.1991	0.9310
$\xi_2 = 0.4329$	0.0366	0.5769	0.2446	0.0498
$\xi_3 = 1$	0.0076	0.2070	0.5563	0.0192

Table A8.3-5: Results for the Three-State POMDP with Sampled Data

Sample Characteristics					
Size (Number of patients)	Percentage of Observed States		Coefficient of Variation (COV)		Prediction Accuracy
	1st 16 Weeks	2nd 16 Weeks	1st 16 Weeks	2nd 16 Weeks	
1285	28.45%	37.17%	**1.68**	0.72	**89.52%**

Model Estimation

Parameter	**State 1 (bad)**		**State 2 (fair)**		**State 3 (good)**	
θ Dispersion	0.4514^{***}	(0.1691)	0.2750^{***}	(0.0067)	0.2422^{*}	(0.1249)
Variables Impacting State Transition						
β_1 [views]	0.7210^{***}	(0.0576)	0.6563^{***}	(0.0693)	0.3620^{***}	(0.1394)
β_2 [thank you]	1.9710^{***}	(0.1715)	0.5629^{***}	(0.0608)	2.2783^{***}	(0.3702)
β_3 [comments]	3.1225^{***}	(0.0369)	0.5262	(0.3609)	0.9255^{***}	(0.2349)
β_4 [usefulness]	1.2335^{***}	(0.1502)	0.1849	(0.1536)	0.7488^{***}	(0.0622)
β_5 [in-degree]	-0.5142^{*}	(0.2792)	0.3008^{***}	(0.1046)	-0.5807^{***}	(0.1026)
β_6 [out-degree]	1.0203^{***}	(0.0956)	0.3194^{***}	(0.0463)	0.4181^{***}	(0.1297)
β_7 [info. support]	0.2318^{**}	(0.0992)	0.3671^{**}	(0.1578)	0.3220^{***}	(0.0993)
β_8 [emo. support]	0.5427^{***}	(0.1241)	0.4720^{**}	(0.1909)	0.6175^{***}	(0.1684)
β_9 [posts]	3.1382^{***}	(0.1472)	0.9119^{***}	(0.1616)	1.5487^{***}	(0.5431)
β_{10} [# treatment]	0.6172^{***}	(0.1415)	0.3132^{*}	(0.1842)	0.5910^{***}	(0.0932)
β_{11} [# sympt]	-1.0935^{***}	(0.0651)	-1.2910^{***}	(0.0972)	0.1282^{**}	(0.0508)
Thresholds						
State 1 (bad)			0.6954^{***}	(0.0808)	3.2910^{***}	(0.2140)
State 2 (fair)	-2.9286^{***}	(0.5502)			1.2889^{***}	(0.2594)
State 3 (good)	-3.0154^{***}	(0.2802)	-0.6550^{***}	(0.1983)		
Variables Impacting State Dependent Outcome						
γ_0 [constant]	0.8268^{***}	(0.3122)	0.5094^{***}	(0.1181)	0.2476^{***}	(0.0326)
γ_1 [gender]	1.0655^{***}	(0.2304)	0.5757^{***}	(0.2170)	0.6395^{***}	(0.2312)
γ_2 [info quality]	0.1898^{***}	(0.0516)	0.4828^{**}	(0.1970)	0.1943^{***}	(0.0624)
γ_3 [membership]	1.3314^{***}	(0.0996)	0.4190	(0.5489)	1.3266^{***}	(0.2223)
γ_4 [update]	-1.0428^{***}	(0.0977)	-0.6508^{*}	(0.3882)	-0.8810^{***}	(0.1068)

Unobserved Heterogeneity (η , ξ)

$C_\eta = -0.307, C_\xi = -0.189$

	$\eta_1 = 0$	$\eta_2 = 0.3294$	$\eta_3 = 0.5011$	$\eta_4 = 1$
Probability $G(\eta)$	0.0696	0.3415	0.3737	0.2152
Conditional Distribution: $H(\xi \mid \eta)$				
$\xi_1 = 0$	0.9558	0.2206	0.2051	0.8949
$\xi_2 = 0.4326$	0.0362	0.5769	0.2446	0.0493
$\xi_3 = 1$	0.0080	0.2025	0.5503	0.0558

A9. Correlation Matrix

Table A9: Correlation Matrix for Key Variables

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
views [1]	1											
thank you [2]	0.7397	1										
comments [3]	0.7152	0.7007	1									
usefulness [4]	0.7612	0.7392	0.7201	1								
in_degree [5]	0.5652	0.5440	0.5262	0.5610	1							
out_degree [6]	-0.0291	-0.0274	-0.0280	-0.0294	0.0580	1						
info. support [7]	-0.0733	-0.0715	-0.0697	-0.0649	-0.0562	-0.0236	1					
emo. Support [8]	-0.0528	-0.0499	-0.0478	-0.0462	-0.0430	-0.0179	0.1396	1				
companionship [9]	-0.0554	-0.0529	-0.0513	-0.0517	-0.0426	-0.0158	0.2220	0.3501	1			
posts [10]	0.8318	0.8074	0.7862	0.8542	0.6164	-0.0322	0.0003	0.0007	-0.0052	1		
# treatment [11]	0.0315	0.0308	0.0305	0.0353	0.0263	0.0024	-0.0025	-0.0032	-0.0046	0.0374	1	
# sympt [12]	0.0407	0.0343	0.0413	0.0324	0.0423	-0.0046	-0.0079	-0.0033	-0.0045	0.0373	-0.0084	1
gender [13]	0.0027	-0.0020	-0.0113	0.0014	0.0039	-0.0034	0.0046	0.0008	0.0091	0.0058	0.0136	-0.0143
info quality [14]	0.0000	0.0003	-0.0021	0.0013	-0.0004	-0.0013	0.0020	-0.0008	0.0015	0.0035	-0.0079	-0.0138
membership [15]	0.7730	0.7483	0.7301	0.7783	0.6104	-0.1106	-0.0638	-0.0481	-0.0500	0.8467	0.0378	0.0533
update [16]	0.1380	0.1343	0.1286	0.1389	0.2106	0.0633	-0.0151	-0.0158	-0.0157	0.1506	0.0082	0.0107

	[13]	[14]	[15]	[16]
gender [13]	1			
info quality [14]	0.0052	1		
membership [15]	0.0094	0.0039	1	
update [16]	0.0021	0.0049	0.1676	1