Prioritizing Hepatitis C Treatment in U.S. Prisons

Turgay Ayer, Can Zhang, Anthony Bonifonte
H. Milton Stewart School of Industrial & Systems Engineering, Georgia Institute of Technology, Atlanta, GA

Anne Spaulding
Rollins School of Public Health, Emory University, Atlanta, GA

Jagpreet Chhatwal
Massachusetts General Hospital Institute for Technology Assessment, Harvard Medical School, Boston, MA

About one out of six inmates in the United States (U.S.) is infected with hepatitis C virus (HCV). HCV prevalence in prison systems is ten times higher than the general population, and hence prison systems offer a unique opportunity to control the HCV epidemic. New HCV treatment drugs are very effective, but providing treatment to all inmates is prohibitively expensive, which precludes universal HCV treatment in prison systems. As such, current practice recommends prioritizing treatment based on clinical and incarceration-related factors, including disease staging, remaining sentence length, and injection drug use (IDU) status. However, there is controversy about how these factors should be incorporated because of the complicated tradeoffs. In this study, we propose a restless bandit modeling framework to support hepatitis C treatment prioritization decisions in U.S. prisons. We first prove indexability for our problem and derive several structural properties of the well-known Whittle's index, based on which, we derive a closed-form expression of the Whittle's index for patients with advanced liver disease. From the interpretation of this closed-form expression, we anticipate that the performance of the Whittle's index would degrade as the treatment capacity increases; and to address this limitation, we propose a capacity-adjusted closed-form index policy. We parameterize and validate our model using real-world data from Georgia state prison system and published studies. We test the performance of our proposed policy using a detailed, clinically-realistic simulation model and show that our proposed policy can significantly improve the overall effectiveness of the hepatitis C treatment programs in prisons compared with the current practice and other benchmark policies, including the commonly used Whittle's index policy. Our results also shed light on several controversial health policy issues in hepatitis C treatment prioritization in the prison setting and have important policy implications including: 1) prioritization based on only liver health status, a commonly practiced policy, is suboptimal compared with many other policies we consider. Further, considering remaining sentence length of inmates and IDU status in addition to liver health status in prioritization decisions can lead to a significant performance improvement; 2) the decision of whether to prioritize patients with shorter or longer remaining sentence lengths depends on the treatment capacities inside and outside the prison system, and prioritizing patients with shorter remaining sentence lengths may be preferable in some cases, especially if the treatment capacity inside the prison system is not very tight and linkage-to-care level outside prison system is low; and 3) among patients with advanced liver disease, IDUs should not be prioritized unless their reinfection is very-well controlled.

Key words: public health; hepatitis C; resource allocation; treatment prioritization; prisons; multi-armed bandits; index-based policies; agent-based simulation.

1. Introduction

More than 3 million persons in the United States (U.S.) have hepatitis C virus (HCV) infection (Denniston et al. 2014). Chronic HCV infection is the leading cause of liver cirrhosis, liver failure, and liver cancer (Ghany et al. 2009), the fastest-growing cause of cancer-related deaths in the U.S. (El-Serag 2011).

Recently, there have been major advancements in hepatitis C treatment, which could change the overall strategy in HCV management. In December 2013, the Food and Drug Administration (FDA) approved the first non-interferon-based, all-oral regimen HCV treatment drugs (Gane et al. 2013). This was followed by the development of several other all-oral new direct-acting agents (DAAs), which changed the landscape in HCV treatment. On June 28, 2016 the FDA approved the first pan-genotypic (for all genotypes) combination drug, Sofosbuvir with Velpatasvir, for hepatitis C. Compared with the old drugs, the new treatment regimens have much higher cure rates (>95% vs. <50%), shorter duration of treatment (8-12 weeks vs. 24-48 weeks), and significantly fewer side effects (Lawitz et al. 2013). These incredible advancements on the treatment side have made many experts and health organizations, including the World Health Organization, believe that the time has come for global HCV elimination (Mitruka et al. 2015, Hellard et al. 2016).

The criminal justice system is one of the areas with the highest concentration of HCV-infected people (Rich et al. 2014). This is because most HCV-infected people are current or past injection drug users (IDUs) (Page et al. 2013), and most Americans who inject drugs have been incarcerated at some point in their lives (CDC 2015, Rich et al. 2014). Indeed, with about one out of every six inmates being infected, HCV prevalence in correctional institutions is about 10 times higher than the in general population (17% vs. 1-2%).

Given that more than 10 million Americans cycle in and out of prisons and jails every year, including about one of every three HCV-infected patients (Varan et al. 2014), many experts have argued that the criminal justice system has the potential to be one of the best places to focus for effective HCV management (Spaulding and Thomas 2012, Liu et al. 2014, Rich et al. 2014, He et al. 2015). In addition, the criminal justice system is an ideal place to identify, test and treat high-risk patients that are hard to reach when in the community (Larney et al. 2014). The facts that linkage-to-care, the process of engaging HCV-infected patients into primary care, is higher in prisons than the community, and that a sizable portion of the released prisoners do not seek cure upon release provide further reasons for effectively treating inmates while in prisons (Liu et al. 2014).

However, one of the biggest barriers against the plan of effectively treating infected inmates while in prison is the prohibitive cost of the new generation HCV drugs. With a list price of up to $1000 per pill, a single line of treatment with the new drugs can cost up to $84,000 per patient (Chhatwal et al. 2015). The U.S. prisons operate under tight resource constraints, and healthcare budgets of state prisons are allocated by the state legislature a year in advance. Neither private insurance nor federal funds supplement healthcare costs of inmates. As such, given that treating one HCV-infected inmate costs as much as the overall healthcare costs of about 17 prisoners (Baker et al. 2015), most correctional facilities do not have resources to treat all HCV infected inmates. Indeed, a recent study has shown that if all HCV-infected inmates in a state prison system were to get treatment, the cost would be about twice as high as the entire system's annual total healthcare budget (Baker et al. 2015). Although prison systems may negotiate for better prices for HCV drugs, these new drugs are still significantly more expensive than the old regimens (cost about six times higher), which precludes universal treatment (Chhatwal et al. 2015).

Given that universal treatment is not affordable, the Federal Bureau of Prisons encourages correctional providers to prioritize inmates for HCV treatment (Federal Bureau of Prisons 2016). Current correctional system treatment prioritization guidelines and the published literature have suggested considering disease staging, remaining sentence length, and IDU status in prioritization decisions (Jayasekera et al. 2014, Baker et al. 2015, Federal Bureau of Prisons 2016). However, there is controversy about how these factors should be incorporated because of the complicated tradeoffs, and to our knowledge such tradeoffs have not been evaluated while formulating prioritization guidelines. For instance, prioritizing IDUs can reduce the risk of spreading the disease to a larger population; however IDUs are more likely to be reinfected after treatment, hence prioritizing such patients may simply lead to a waste of scarce resources. Similarly, prioritizing patients with longer sentence lengths can maximize the life expectancy of the prison population, but prioritizing patients with shorter sentence lengths may be more effective in controlling the disease because of the reduced risk of spread in the general population. To summarize, in practice, there is no consensus, treatment initiation decisions are ad hoc and vary across prisons, and a standardized HCV treatment prioritization protocol is lacking (Liu et al. 2014, Baker et al. 2015). Because of limited budget in prisons and lack of clarity on who is prioritized for treatment, many inmates have filed lawsuits against the prisons (Loftus 2015, Snowbeck 2015, Boucher 2015).

In this study, we propose a mathematical modeling-based approach for prioritizing HCV treatment in state prison systems. In particular, we propose a restless multi-armed bandit (RMAB) modeling approach to systematically compare a large number of treatment prioritization rules and identify effective prioritization strategies. To our knowledge, ours is the first to study HCV

treatment prioritization decisions in prisons, and while we primarily focus on HCV treatment prioritization in this study, our proposed RMAB model is flexible and can be applied in more general resource allocation decisions for chronic diseases. For example, another application area for our approach could be resource allocation in middle and low income countries, or in state Medicaid programs and private insurance companies. Indeed, current HCV drug prices are prohibitive even for the general population and hence treatment capacity is also limited in the community. As such, many insurance companies are also prioritizing treatment decisions and limiting treatment to selected patients in the community (Silverman 2014, Horberg 2014), where only about 5-10% of the more than 3 million HCV infected Americans receive treatment (Yehia et al. 2014, Bidell et al. 2016).

The general RMAB problem is notoriously difficult to solve since the state space explodes exponentially as the number of bandits increases. Papadimitriou and Tsitsiklis (1999) showed that this problem is PSPACE-hard, even in the special case of deterministic transitions. To solve this problem, Whittle (1988) proposed an index-type policy, well-known as the Whittle's index, which provides a complete ordering for all bandits at any possible states, and bandits with higher indices are prioritized at each period. Such a policy is easy to interpret and implement. Further, although the optimality of the Whittle's index is not guaranteed, it has been shown to be very competitive and perform near optimally, if not optimally, in a wide range of application areas (Glazebrook et al. 2005, Archibald et al. 2009, Liu and Zhao 2010).

In this study, we consider Whittle's index and other index-type policies. However, the implementation of Whittle's index is nontrivial because i) Whittle's index is only applicable to a restricted class of problems that satisfy the so-called indexability property, which is difficult to check in general (Whittle 1988, Glazebrook et al. 2006); ii) the computation of Whittle's index involves repeated solution of single-bandit dynamic programs, which is difficult for our case since each single bandit has a multi-dimensional state space.

To address the above challenges, we first establish indexability of our problem through a novel approach. Unlike most machine maintenance literature where indexability is only shown under the assumption of perfect repair (e.g., Glazebrook et al. (2006)), our analysis allows the possibility of an unsuccessful treatment, which is common in disease treatment. Then, we derive several structural properties of the Whittle's index, based on which, we develop a closed-form expression of Whittle's index for patients with advanced liver disease. This closed-form expression of Whittle's index significantly expedites the computation, and more importantly, provides an intuitive interpretation allowing us to predict under what circumstances Whittle's index would perform well or poorly for our problem. In particular, based on this closed-form expression, we anticipate that the performance of the Whittle's index would degrade as the treatment capacity increases (note that Whittle's

index is blind to the treatment capacity). Therefore, building upon this closed-form expression of Whittle's index, we propose a capacity-adjusted closed-form index policy that explicitly captures the treatment capacity. We parameterize and validate our proposed RMAB model using real data from state prison systems and published studies. We further develop a detailed realistic agent-based simulation model to test the performance of our proposed capacity-adjusted index policy against Whittle's index and other competitive benchmark policies such as myopic policy, and show that our proposed policy has a significant performance improvement.

Our research has important policy implications. First, we find that prioritization based on only liver health status, a commonly practiced policy, is suboptimal compared with many other policies we consider. Further, we show that simultaneously considering health state, remaining sentence length, IDU status, and age in prioritization decisions can lead to a significant performance improvement. Second, we find that the decision of whether to prioritize patients with shorter or longer remaining sentence lengths depends on the treatment capacities inside and outside the prison system, and prioritizing inmates with shorter length of sentences may be preferable in some cases, especially if the treatment capacity inside the prison system is not very tight and linkage-to-care level outside prison system is low. In that regard, we reiterate that treatment capacity is also limited in the community, and a possibly higher treatment capacity in the community may not necessarily translate to a higher treatment rate for released inmates due to lower linkage-to-care. Third, we shed light on the controversy surrounding the prioritization of IDUs and findthat for patients with advanced liver disease, the decision is very sensitive to reinfection rate, and IDUs should not be prioritized unless reinfection is well controlled.

The remainder of this paper is organized as follows. In §2, we present a review of relevant literature. In §3, we provide more background on the disease epidemiology and present a formal model formulation. In §4, we establish indexability for our problem and derive several structural properties for Whittle's index. In §5, we develop a closed-form expression of Whittle's index for patient with advanced liver disease and propose a capacity-adjusted closed-form index policy. In §6, we describe the key components of our simulation model and also provide an overview of parameter estimation and calibration. In §7, we test the performance of our proposed policy, compare with other benchmark policies, and analyze how the priorities may change with respect to several key parameters. Finally, in §8, we summarize our findings and conclude.

2. Literature Review

Our work is related to the literature on resource allocation problems for chronic disease management. Kaplan and Pollack (1998), Brandeau et al. (2005), Kaplan and Merson (2002), Alistar et al. (2014) and Deo et al. (2015) studied the problem of allocating limited resources for HIV prevention

and control; Deo et al. (2013) studied a capacity allocation problem in community-based childhood asthma treatment programs; Khademi et al. (2015) examined prioritizing initiation and termination decisions for HIV treatment; Deo and Sohoni (2015) and Deo et al. (2016) respectively studied the optimal allocation of point of care devices and the optimal location of labs and diagnostic equipment to improve the efficiency of early infant diagnosis of HIV. There is also a growing stream of research focusing on organ allocation among patients with end-stage diseases, such as liver cancer or kidney failure (e.g., Zenios et al. (2000), Su and Zenios (2004, 2005), Akan et al. (2012), Dai (2012), Bertsimas et al. (2013), Anderson et al. (2015).) Among the above mentioned studies, Deo et al. (2013) and Khademi et al. (2015) are most relevant to ours. In particular, these studies have a similar motivation to ours in the sense that both of them considered allocation of a fixed capacity/budget at each period for the optimal control of a chronic disease. More specifically, Deo et al. (2013) studied capacity allocation for childhood asthma using a restless bandit framework, and Khademi et al. (2015) studied resource allocation for HIV control using an approximate dynamic programming approach. On the other hand, our work is distinct because in addition to studying a very different problem with unalike dynamics, the state space of our problem is multi-dimensional and not completely ordered, which requires a different analysis. Further, unlike these and many other studies in the literature, we consider an open population by allowing movements into and out of the prison system via arrest and release events; as such, the heuristic policies developed in the above studies perform poorly in our case.

Our work also relates to the stream of research that focuses on optimal chronic disease treatment strategies. In this context, optimal radiation therapy planning for cancer treatment has received a wide attention (e.g., Romeijn et al. (2006), Bortfeld et al. (2008), Taskin et al. (2010), Lavieri et al. (2012), Chan et al. (2014)). Further, Lee et al. (2008) studied the optimal initiation of dialysis for patients afflicted with chronic kidney failure, Shechter et al. (2008) examined the optimal time to initiate HIV therapy under ordered health states, Mason et al. (2014) investigated blood pressure and cholesterol control strategies in diabetic patients, Helm et al. (2015) considered the timing of periodic examinations for patients with glaucoma, Liu et al. (2015) studied the optimal patient treatment decisions for chronic diseases considering that treatment technologies improve over time, and Bertsimas et al. (2016) developed models for the analysis and design of clinical trials to discover drug combinations with the highest potential for improved clinical outcomes. A key distinction between these and our settings is that, this stream of work mainly focuses on the optimal treatment strategies for individual patients, while our study focuses on the optimal allocation of resources and treatment prioritization strategies among multiple patients.

Finally, our work relates to the stream of research that studies operations management problems using a restless bandit framework. For example, Katehakis and Derman (1986) studied the optimal

sequential allocation of treatments in clinical trials, Glazebrook et al. (2005, 2006) studied the dynamic scheduling of maintenance for a collection of machines, Glazebrook et al. (2009) examined the admission control and routing of customers to heterogeneous service stations, Archibald et al. (2009) studied a class of inventory routing problems with direct deliveries, and Argon et al. (2009) and Ding and Glazebrook (2013) considered the dynamic routing of customers/products in a set of parallel queues. While also using a restless bandit modeling framework, our work differs from the above studies mainly in the following two aspects: i) While most of the above studies deal with a single-dimensional and completely-ordered state space, the state space for our problem is multi-dimensional and partially ordered, which imposes several challenges in proving indexability and deriving closed-form expressions for index policies; and ii) while both the Whittle's index and the myopic policy are shown to perform near optimally in most of the existing studies, we show that such capacity-blind policies may fail to perform well in our problem, and we further present a novel capacity-adjusted policy. We believe this observation and our proposed policy could also be useful in other similar applications.

3. Background and Model Formulation

HCV infection causes inflammation and ultimately scarring in liver, referred to as fibrosis. The extent of liver fibrosis is measured by a scoring system, called the METAVIR scoring system. In the METAVIR scoring system, the degree of fibrosis ranges from F0 to F4 (Bedossa and Poynard 1996), where F0-F3 stages represent none, mild, moderate, and advanced fibrosis respectively, and F4 stage represents late-stage fibrosis (often referred to as compensated cirrhosis), where scarring is widespread in entire liver (Fattovich et al. 1997). Patients with hepatitis C may further develop decompensated cirrhosis (DC) and hepatocellular carcinoma (HCC), a common type of liver cancer. The target population for hepatitis C treatment includes inmates with F0-F4 stage fibrosis, and the objective is to select and prioritize patients for treatment at each period to maximize the overall health outcomes measured by total expected quality-adjusted life years (QALYs), which simultaneously consider quantity and quality of life (Gold et al. 1996, Drummond et al. 2005). The response to treatment is measured by sustained viral response (SVR), defined as the absence of hepatitis C virus in blood serum for at least 12 weeks after discontinuing the treatment. Patients achieving SVR are considered effectively treated. However, even after viral clearance, reinfection is possible following new exposure to HCV via risky activities, such as injection drug use. When reinfected, the degradation in liver resumes from the stage at which treatment was applied (e.g., a treated F4 patient, denoted as F4SVR, directly transitions back to F4 upon reinfection; see also Figure 1 for a schematic illustration of disease progression).

We formulate the HCV treatment prioritization problem as a restless multi-armed bandit (RMAB) problem, which is first introduced by Whittle (1988) to address the sequential allocation

of limited resources among multiple projects, or bandits. The RMAB falls into a special case of the so-called weakly-coupled Markov Decision Process (MDP), where different bandits are only coupled through a capacity constraint, and is formally defined by $\{(S_n, P_n^a, P_n^b, r_n^a, r_n^b), 1 \leq n \leq N, M, \beta\}$, where N denotes the total number of bandits (i.e., patients in our setting), M denotes the maximum number of bandits that can be activated (i.e., treated in our setting) at each period, and β denotes the discount factor. For each bandit $n = 1, ..., N$, S_n denotes the state space, P_n^a and r_n^a denote the active (i.e., treated) transition probability matrix and the corresponding reward vector respectively, and P_n^b and r_n^b denote the passive (i.e., not treated) transition probability matrix and the corresponding reward vector respectively. The problem is to select which bandits to activate at each period such that the total discounted reward is maximized.

Inline with the common practice, we assume that the prison system has an annual budget, and at most M prisoners can be treated every year. We also assume that throughout the planning horizon, the number of prisoners is fixed and equal to N. Note that this is not a restrictive assumption because most prisons operate at full capacity (West et al. 2010). Also, this assumption can be easily relaxed by considering virtual patients with perfect health state and one year of sentence to fill in the extra capacity. Lastly, for tractability, we assume no transmission of disease inside the prison system in the analytical model. While transmission of HCV inside prisons does occur, its rate is typically much smaller compared with the community setting. This is because in a prison, with persons living alone or two in a cell, there is no homogeneous mixing of populations, and the supply of heroin and other drugs is limited due to heightened enforcement of prohibition. We further relax this assumption in our simulation model described in §6 and show that our proposed policies are robust against transmission rate inside the prison system.

We next define state space, active and passive transition probabilities and rewards for each individual patient $n = 1, ..., N$ as follows. For notational simplicity, we drop the dependence on n.

- $S = S_d \times S_f$: The state space of each patient, where S_d denotes the set of patient characteristics that may dynamically evolve over time, and S_f denotes the set of characteristics that typically remain static, which we describe in detail in the followings.
- $S_d = \{(i, j, k) \in H \times L \times A\}$: The set of patient characteristics that may dynamically evolve over time, where H denotes the set of liver health states, L denotes the set of remaining sentence lengths, and A denotes the set of ages. In particular, let $i \in H = \{1,...,14\}$, where $i = 1$ denotes uninfected; $i = 2, ..., 6$ denote SVR after treatment for patients in F0-F4 (i.e., F0SVR-F4SVR), respectively; $i = 7, ..., 11$ denote five increasingly severe stages of fibrosis (i.e., F0-F4), respectively, which are *candidate states* for HCV treatment; $i = 12$ denotes decompensated cirrhosis (DC), $i = 13$ denotes hepatocellular carcinoma (HCC), and $i = 14$ denotes death (see also Figure 1). As discussed earlier, we differentiate F0SVR-F4SVR patients because when

reinfected, F0SVR-F4SVR patients directly transition back to their original state prior to treatment. We assume fibrosis staging information is available, which is reasonable because staging information can be easily obtained through low-cost non-invasive screening tests, such as Fibrosure (He et al. 2015), and regular screening using these low-cost screening tests is increasingly common in prison systems (He et al. 2015, Federal Bureau of Prisons 2016). As for the remaining length of sentence, we consider up to 15 years, as 98% of the patients have less than 15 years of sentence (Georgia Department of Corrections 2014). Finally, we consider age up to 100 years, which is inline with the relevant published literature (He et al. 2015).

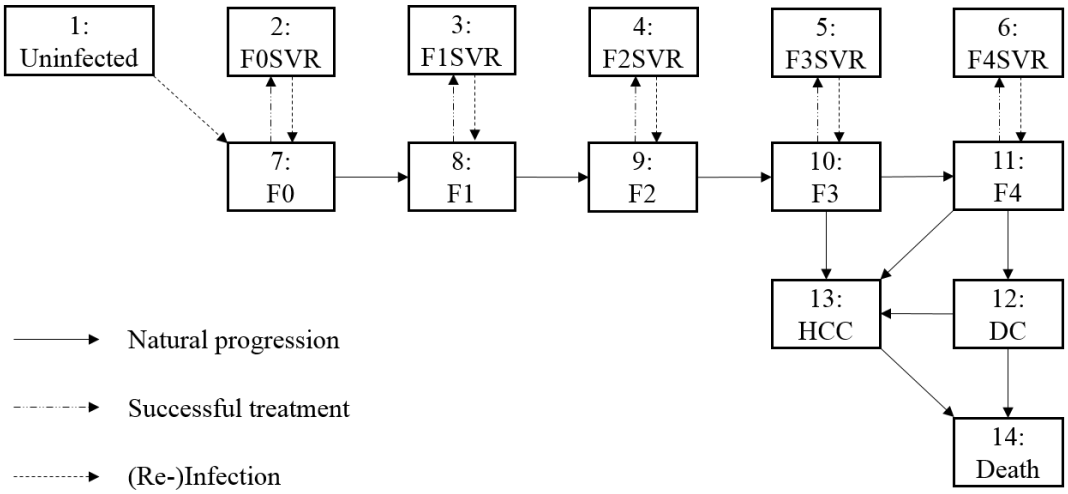


Figure 1 **A schematic illustration of hepatitis C natural history and possible state transitions.**

Notes. Background mortality rates are not captured in this figure.

- S_f: The set of patient characteristics that typically remain static, including gender (male/female), HCV genotype (G1/G2/G3/G4), and IDU status (IDU/non-IDU). We remark that, although rare, IDU status may change over time. While we consider IDU status to be a static factor in the analytical model for tractability, we explicitly model the dynamic change in the drug use behaviors in the simulation model described in §6.
- R: The 14×14 natural history progression matrix of hepatitis C, which is upper triangular, representing that there is no natural recovery from the disease (Thein et al. 2008). Possible transitions among disease states are shown in Figure 1. We remark that in our computational analysis, we allow R to be age dependent, as the background mortality changes with age. In our theoretical analysis though, we drop this dependency on age for notational simplicity; however, all of our results can be easily extended to consider an age-dependent matrix R.
- Q: The 14×14 treatment matrix of hepatitis C. As described earlier, F0-F4 (i.e., health states $i = 7, ..., 11$) are the candidate states for HCV treatment, and upon treatment, such patients

either achieve SVR and move to F0SVR-F4SVR (i.e., $i = 2, ..., 6$), or else continue to stay in their current states. Hence, for a candidate health state i, we have $Q_{i,i-5} + Q_{i,i} = 1$, where $Q_{i,i-5}$ captures the SVR rate. For non-candidate health states i, we simply set $Q_{i,i} = 1$. Then, for those patients undergoing treatment, state transitions are governed by QR.

- P^a and P^b: The active and passive transition probability matrices, respectively. As discussed earlier, we assume that the prison operates at full capacity, implying that released prisoners are replaced by new ones immediately. Then, the active and passive transition matrices for patients of state $(i, j, k) \in S_d$ are defined as follows:

$$P^a_{(i,j,k),(i',j',k')} = \begin{cases} (QR)_{i,i'} & \text{if } j > 1, j' = j-1, k' = k+1 \\ \pi_{i',j',k'}, & \text{if } j = 1, \\ 0, & \text{otherwise,} \end{cases}$$

$$P^b_{(i,j,k),(i',j',k')} = \begin{cases} R_{i,i'} & \text{if } j > 1, j' = j-1, k' = k+1 \\ \pi_{i',j',k'}, & \text{if } j = 1, \\ 0, & \text{otherwise,} \end{cases}$$

  where $\pi_{i,j,k}$ represents the distribution of the incoming (i.e., arrested) inmate population with respect to health state i, sentence length j, and age k.

- r^a and r^b: The active and passive reward vectors, capturing QALYs accumulated by treated and non-treated patients, respectively. We let $r_{i,k}$ represent the annual pre-release QALYs for a patient with health state i and age k. After release, the released patients may contribute to the overall QALYs accumulated in two ways: i) by QALYs gained by themselves over time, and ii) by possibly infecting other people in the society, leading to a reduction in overall QALYs in the population. Therefore, for a given state $(i, j, k) \in S_d$, the active and passive rewards are defined as follows:

$$r^a_{i,j,k} = \begin{cases} r_{i,k} + \beta \sum_{i'} (QR)_{i,i'} z_{i',k+1}, & \text{if } j = 1, \\ r_{i,k}, & \text{otherwise,} \end{cases}$$

$$r^b_{i,j,k} = \begin{cases} r_{i,k} + \beta \sum_{i'} R_{i,i'} z_{i',k+1}, & \text{if } j = 1, \\ r_{i,k}, & \text{otherwise,} \end{cases}$$

  where β represents the discount factor and $z_{i,k}$ represents the post-release lump sum reward accumulated from a released patient with health state i and age k. Note that while (re)infection is negligible in prisons due to strict regulations, it is common especially among IDUs in the general population, which is captured in the estimation of $z_{i,k}$. In particular, we estimate the potential reduction in overall QALYs in the population due to possible spread of the disease to others in the society as the total expected number of infections caused by the released patient times the expected lifelong loss of QALYs for each HCV infection (more detailed descriptions about the estimations of this and all other parameters can be found in Appendix B).

4. Indexability Analysis and Structural Properties

As discussed in §1, since the restless bandit problem is intractable due to the curse of dimensionality, we focus on index-type policies. An index policy significantly reduces the computational burden by decoupling the original problem into subproblems of individual bandits, where each bandit is assigned with an index value at each period, and bandits with higher indices are prioritized. Index-type policies are also appealing from an implementation perspective due to their simplicity and ease of interpretation. However, the well-known and widely used Whittle's index is not universally applicable to all restless bandit problems; instead, it is only applicable to a restricted class of problems that satisfy the so-called indexability property, which is difficult to check in general (Glazebrook et al. 2006, Whittle 1988).

In §4.1, we first establish indexability for our problem under mild conditions which are perfectly satisfied by real data. Then, in §4.2, we derive several structural properties for the Whittle's index, which helps us better understand the behavior and performance of this index-type policy. This analysis leads to the development of an exact closed-form expression of Whittle's index for patients with late-stage fibrosis (F4), and a capacity-adjusted closed-form index policy for all patients, which we present in the following section. Unless presented in the main text, the proofs for all results presented in this and the following sections can be found in Appendix A.

4.1. Indexability Analysis

To analyze an individual bandit, it is sufficient to consider the set of characteristics that dynamically evolve over time (i.e., S_d), and we omit the dependence on the set of static characteristics (i.e., S_f) for notational simplicity. For a given individual bandit and any state $(i,j,k) \in S_d$, consider a passive *subsidy* W, i.e., assume that when passive action is taken, the reward is $r^b_{i,j,k} + W$ instead of $r^b_{i,j,k}$. Let $V_{i,j,k}(W)$ be the optimal value function of state (i,j,k) when the passive subsidy is W. Then, $V_{i,1,k}(W) = \max\{r^a_{i,1,k}, r^b_{i,1,k} + W\}$, and for $j \geq 2$:

$$V_{i,j,k}(W) = \max\left\{ r^a_{i,j,k} + \beta \sum_{i'} (QR)_{i,i'} V_{i',j-1,k+1}(W), r^b_{i,j,k} + W + \beta \sum_{i'} R_{i,i'} V_{i',j-1,k+1}(W) \right\}.$$

Next, we provide the definitions of indexability and Whittle's index as follows.

Definition 1 (Indexability). Given subsidy level W, let $\Pi(W)$ be the set of states for which the passive action is optimal. A bandit is *indexable* if $\Pi(W)$ is increasing in W, i.e., $W_1 \leq W_2 \implies \Pi(W_1) \subseteq \Pi(W_2)$.

Definition 2 (Whittle's index). For an indexable bandit, its *Whittle's index* for each state (i,j,k) is defined as: $W_{i,j,k} = \inf\{W : (i,j,k) \in \Pi(W)\}$.

Intuitively, these definitions say that a bandit is indexable if the set of states at which passive action is optimal is increasing with the subsidy level, and the Whittle's index of a given state is the

smallest subsidy level at which passive action becomes optimal. Although indexability may appear to be a natural property, it may not hold in general. Before proving indexability for our problem, we first present some natural conditions below, based on which we establish the indexability.

CONDITION 1. *Recall that $i = 2, ..., 6$ respectively represent F0SVR-F4SVR, $i = 7, ..., 11$ respectively represent F0-F4, and $i = 14$ represents death. Then,*

- *Mortality rates are smaller and rewards (i.e. QALYs gained) are larger in successfully treated patients achieving SVR, compared with non-SVR patients. That is, for $i = 7, ..., 11$ and for all k , $R_{i-5,14} \leq R_{i,14}, r_{i-5,k} \geq r_{i,k}$, and $z_{i-5,k} \geq z_{i,k}$.*
- *Mortality rates are non-decreasing and the rewards are non-increasing in health states. That is, $R_{i,14}$ is non-decreasing in i, and for all k, $r_{i,k}$ and $z_{i,k}$ are non-increasing in i.*
- *SVR rates are non-decreasing in liver health state. That is, for $i = 7, ..., 11$, $Q_{i,i-5}$ is non-decreasing in i.*

The first two conditions are natural conditions and respectively imply that the total benefits from treatment is higher than no treatment, and relatively healthier states are preferable over sicker states with respect to mortality and rewards. The third condition implies that the response rate to HCV treatment does not worsen as the liver condition deteriorates, which is also reasonable because newest HCV drugs typically show similar effectiveness across liver fibrosis stages (Afdhal et al. 2014).

In proving indexability, the main challenge is that the optimal value function $V_{i,j,k}(W)$ is unknown. To overcome this challenge, we first construct an easy-to-compute alternative value function, and then show that indexability can be established by replacing the true optimal value function with the constructed one, as we describe in detail below.

For a given passive subsidy $W$and state (i, j, k) , let $\tilde{V}_{i,j,k}(W)$ denote the value function when active action is taken at candidate states (i.e., F0-F4) and passive action is taken at non-candidate states at all periods. Also, let $\mathbf{e}_0$ denote the 14-dimensional column vector with the entries of non-candidate health states equal to 1 and entries of candidate health states equal to 0, and let $\mathbf{e}_i$ denote the 14-dimensional column vector with the ith entry equal to 1 and other entries equal to zero. Then, we have the following results, which we use in proving indexability in Theorem 1.

LEMMA 1. $V_{i,j,k}(W_2) - V_{i,j,k}(W_1) \geq \tilde{V}_{i,j,k}(W_2) - \tilde{V}_{i,j,k}(W_1)$, *for* $0 \leq W_1 \leq W_2, \forall i, j, k$

LEMMA 2. $\tilde{V}_{i,j,k}(W_2) - \tilde{V}_{i,j,k}(W_1) = (W_2 - W_1)\boldsymbol{e}_i^T \left(\sum_{j'=1}^{j} \beta^{j'-1}(QR)^{j'-1}\boldsymbol{e}_0 \right)$, *for* $0 \leq W_1 \leq W_2, \forall i, j, k$.

LEMMA 3. $\beta \boldsymbol{e}_i^T (QR - R) \left(\sum_{j'=1}^{j-1} \beta^{j'-1}(QR)^{j'-1}\boldsymbol{e}_0 \right) \leq 1, \forall i, j, k$.

Lemma 1 says that the difference between the optimal value functions under two different passive subsidy levels (i.e., $V_{i,j,k}(W_2) - V_{i,j,k}(W_1)$) is at least as high as the difference between the constructed value functions under the same subsidy levels (i.e., $\tilde{V}_{i,j,k}(W_2) - \tilde{V}_{i,j,k}(W_1)$). This is a key result which enables us to replace the unknown optimal value function (i.e., $V_{i,j,k}(W)$) with the value function we constructed (i.e., $\tilde{V}_{i,j,k}(W)$) in the analysis. Lemma 2 provides a closed-form expression for $\tilde{V}_{i,j,k}(W_2) - \tilde{V}_{i,j,k}(W_1)$. Lastly, in Lemma 3, $\beta \mathbf{e}_i^T QR\left(\sum_{j'=1}^{j-1} \beta^{j'-1}(QR)^{j'-1}\mathbf{e}_0\right)$ represents the total units of subsidies collected by a patient if he is always treated whenever at a candidate state, while $1 + \beta \mathbf{e}_i^T R\left(\sum_{j'=1}^{j-1} \beta^{j'-1}(QR)^{j'-1}\mathbf{e}_0\right)$ represents the total units of subsidies collected by a patient if he is not treated at the current period but is always treated in the following periods whenever at a candidate state. Then, Lemma 3 says that the subsidy collected in the latter case, i.e., when treatment is delayed by one period, is larger. Given these lemmas, we are now ready to present the main result of this section as follows.

THEOREM 1. *Under Condition 1, the hepatitis C treatment prioritization problem is indexable.*

Proof. By definition, to show indexability, it is sufficient to show that the set of states at which the passive action is optimal is non-decreasing in the passive subsidy level W. That is, for any given state, as W increases, the optimal action can only switch from active to passive, and never switch back.

For a non-candidate state, since the state transitions and rewards are independent of the actions, clearly active action is optimal when $W \leq 0$ and passive is optimal when $W \geq 0$. Therefore, it is sufficient to consider candidate states $i = 7, ..., 11$. Further, given Condition 1, it is straightforward to show that active action is optimal for all candidate states when $W \leq 0$. Therefore, we focus on the cases where $W \geq 0$.

Recall that for any state (i, j, k), $V_{i,j,k}(W)$ denotes the optimal value function when the passive subsidy is W. Define $V^a_{i,j,k}(W) = r^a_{i,j,k} + \beta \sum_{i'} (QR)_{i,i'} V_{i',j-1,k+1}(W)$, and $V^b_{i,j,k}(W) = r^b_{i,j,k} + W + \beta \sum_{i'} R_{i,i'} V_{i',j-1,k+1}(W)$. Then, obviously when $V^a_{i,j,k}(W) - V^b_{i,j,k}(W) > 0$, active action is optimal for state (i, j, k); and passive is optimal otherwise. Therefore, to show indexability, it is sufficient to show that $V^a_{i,j,k}(W) - V^b_{i,j,k}(W)$ is non-increasing in W for $i = 7, ..., 11$ and $W \geq 0$.

For $j = 1$, we have $V^a_{i,1,k}(W) - V^b_{i,1,k}(W) = r^a_{i,1,k} - r^b_{i,1,k} - W$, which is clearly non-increasing in W. For $j \geq 2$, we have:

$$
\begin{aligned}
& V^a_{i,j,k}(W) - V^b_{i,j,k}(W) \\
=& r^a_{i,j,k} + \beta \sum_{i'} (QR)_{i,i'} V_{i',j-1,k+1}(W) - r^b_{i,j,k} - W - \beta \sum_{i'} R_{i,i'} V_{i',j-1,k+1}(W) \\
=& r_{i,k} + \beta \sum_{i'} (QR)_{i,i'} V_{i',j-1,k+1}(W) - r_{i,k} - W - \beta \sum_{i'} R_{i,i'} V_{i',j-1,k+1}(W)
\end{aligned}
$$

$$= -W + \beta \sum_{i'} (QR - R)_{i,i'} V_{i',j-1,k+1}(W).$$

Recall that $Q_{i,i-5} + Q_{i,i} = 1$. Then, for $0 \le W_1 \le W_2$, we have:

$$\begin{aligned}
&V^a_{i,j,k}(W_2) - V^b_{i,j,k}(W_2) - (V^a_{i,j,k}(W_1) - V^b_{i,j,k}(W_1)) \\
=& -W_2 + W_1 + \beta \sum_{i'} (QR - R)_{i,i'} (V_{i',j-1,k+1}(W_2) - V_{i',j-1,k+1}(W_1)) \\
=& -W_2 + W_1 + \beta \sum_{i'} (Q_{i,i-5} R_{i-5,i'} + Q_{i,i} R_{i,i'} - R_{i,i'}) (V_{i',j-1,k+1}(W_2) - V_{i',j-1,k+1}(W_1)) \\
=& -W_2 + W_1 + \beta \sum_{i'} (Q_{i,i-5} R_{i-5,i'} - Q_{i,i-5} R_{i,i'}) (V_{i',j-1,k+1}(W_2) - V_{i',j-1,k+1}(W_1)) \\
=& -W_2 + W_1 + \beta \sum_{i'} Q_{i,i-5} R_{i-5,i'} (V_{i',j-1,k+1}(W_2) - V_{i',j-1,k+1}(W_1)) \\
& - \beta \sum_{i'} Q_{i,i-5} R_{i,i'} (V_{i',j-1,k+1}(W_2) - V_{i',j-1,k+1}(W_1)) \\
=& -W_2 + W_1 + \beta \sum_{i'} Q_{i,i-5} R_{i-5,i'} (\tilde{V}_{i',j-1,k+1}(W_2) - \tilde{V}_{i',j-1,k+1}(W_1)) \\
& - \beta \sum_{i'} Q_{i,i-5} R_{i,i'} (V_{i',j-1,k+1}(W_2) - V_{i',j-1,k+1}(W_1)) && (1) \\
\le& -W_2 + W_1 + \beta \sum_{i'} (Q_{i,i-5} R_{i-5,i'} - Q_{i,i-5} R_{i,i'}) (\tilde{V}_{i',j-1,k+1}(W_2) - \tilde{V}_{i',j-1,k+1}(W_1)) && (2) \\
=& -W_2 + W_1 + \beta \sum_{i'} (QR - R)_{i,i'} (\tilde{V}_{i',j-1,k+1}(W_2) - \tilde{V}_{i',j-1,k+1}(W_1)) \\
=& (W_2 - W_1) \left(\beta \mathbf{e}_i^T (QR - R) \left(\sum_{j'=1}^{j-1} \beta^{j'-1} (QR)^{j'-1} \mathbf{e}_0 \right) - 1 \right) && (3) \\
\le& 0, && (4)
\end{aligned}$$

where Equation 1 holds because once a patient achieves SVR, he can only transition to non-candidate states, and for any non-candidate state i' we have $V_{i',j-1,k+1}(W) = \tilde{V}_{i',j-1,k+1}(W)$ for any $W \ge 0$. Also, statements 2-4 follow from Lemmas 1-3, respectively, which completes the proof. □

4.2. Structural Properties of the Whittle's Index

In this section, we derive several structural properties of the Whittle's index for our problem, which are instrumental for the derivation of the closed-form expression of Whittle's index, as we show in the next section. This derived closed-form expression of the Whittle's index helps us understand the behavior and performance of this index-type policy under different settings, based on which we construct our capacity-adjusted policy in §5. Recall that Whittle's index is well-defined only when the problem is indexable; therefore all of the properties in this section are derived under Condition 1. We start with a definition.

Definition 3. For two probability distributions $\mathbf{p}$ and $\mathbf{q}$, $\mathbf{q}$ stochastically dominates $\mathbf{p}$ if $\sum_{i'=1}^{i} p_{i'} \ge \sum_{i'=1}^{i} q_{i'}, \forall i$, and this relationship is denoted by $\mathbf{p} \preceq \mathbf{q}$.

We show in the following proposition that under mild conditions, the Whittle's index is non-decreasing in liver health state.

PROPOSITION 1. *Suppose for* $i = 7, ..., 10$, $R_i \preceq R_{i+1}$, *and the following condition holds:*

$$(\boldsymbol{e}_{i-5}^T - \boldsymbol{e}_{i-4}^T)\left(\sum_{j'=1}^{j-1} \beta^{j'} R^{j'} r_{:,k+j'} + \beta^j R^j z_{:,k+j}\right) \leq (\boldsymbol{e}_i^T - \boldsymbol{e}_{i+1}^T)\left(\beta R\big(z_{:,k+1}\mathbb{1}_{\{j=1\}} + r_{:,k+1}\mathbb{1}_{\{j\geq 2\}}\big)\right), \forall j, k. \tag{5}$$

Then, $W_{i+1,j,k} \geq W_{i,j,k}$, *for* $i = 7, ..., 10, \forall j, k$.

The left hand side in Inequality 5 represents the cumulative reward difference in adjacent fibrosis stages after treatment (e.g., F3SVR and F4SVR), and the right hand side represents one period reward difference in untreated adjacent states (e.g., F3 and F4). Then, this condition says that the cumulative reward difference in adjacent states after treatment would be smaller than one period reward difference in these states before treatment. This condition is natural and is satisfied by real data, because treatment is equally effective in all candidate states (Afdhal et al. 2014), and hence post-treatment life expectancies in different candidate states are very similar.

Unlike in the case for health states, it is less clear how the Whittle's index would behave when the sentence length changes, which we explore next. As discussed in §1, there is much controversy on whether patients with shorter or longer sentence lengths should be prioritized, and we anticipate that the Whittle's index may not be monotone in remaining sentence length for the general case. However, as we show in the following proposition, monotonicity with respect to sentence length is indeed preserved for F4 patients. Although this result may appear to be somewhat surprising, there is an intuitive reasoning behind: while F0-F3 patients can transition to more advanced fibrosis stages, F4 patients will never naturally transition back to earlier fibrosis stages; hence the structure of Whittle's index for F4 patients is much simpler. Before we formally state Proposition 2 which establishes the monotonicity of the Whittle's index with respect to sentence length for F4 patients, we first present an intuitive condition that we need, which is also used in proving Proposition 3.

CONDITION 2. *For* $i = 7, ..., 11$, $r_{i-5,k} - r_{i,k}$ *and* $z_{i-5,k} - z_{i,k}$ *are both non-increasing in* k, *and for* $i = 7, ..., 13$, $r_{i,k} - r_{i+1,k}$ *and* $z_{i,k} - z_{i+1,k}$ *are both non-increasing in* k.

The above condition simply says that the differences between the rewards of an SVR state and the corresponding disease state (i.e., $i-5$ vs. i, i=7,..,11; e.g., F3SVR vs. F3) and adjacent disease states (i.e., i vs. $i+1$, i=7,...,13; e.g., F3 vs. F4) are non-increasing in age. We remark that this is a natural condition satisfied by real data.

PROPOSITION 2. *Suppose Condition 2 and the following condition hold:*

$$(\boldsymbol{e}_6^T - \boldsymbol{e}_{11}^T)\left(\sum_{j'=1}^{j} \beta^{j'} R^{j'} r_{:,k+j'} + \beta^{j+1} R^{j+1} z_{:,k+j+1}\right) \geq (\boldsymbol{e}_6^T - \boldsymbol{e}_{11}^T)\left(\sum_{j'=1}^{j-1} \beta^{j'} R^{j'} r_{:,k+j'} + \beta^{j} R^{j} z_{:,k+j}\right), \forall j, k. \tag{6}$$

Then, $W_{i,j+1,k} \geq W_{i,j,k}$, *for* $i = 11$ *(i.e., F4),* $\forall j, k$.

Intuitively, Inequality 6 says that the marginal benefit of treatment is larger for patients with longer sentence lengths. A formal proof of Proposition 2 can be found in Appendix A, and below we provide a proof sketch, which we believe is helpful in understanding why Whittle's index of F4 patients is non-decreasing in remaining sentence length. Consider two F4 patients with the same age k and one and two years of remaining sentence lengths, respectively, i.e., two patients with states $(11, 1, k)$ and $(11, 2, k)$, respectively. To show that the Whittle's index of state $(11, 2, k)$ is higher, it is sufficient to show that when the passive subsidy is equal to $W_{11,1,k}$, the optimal action for state $(11, 2, k)$ is active. First, state $(11, 2, k)$ will transition to either state $(11, 1, k+1)$ or to a non-candidate state (e.g., DC, HCC, or death) at the next period, where, by Condition 2, the optimal action must be passive when the passive subsidy is equal to $W_{11,1,k}$. Then, by Inequality 6, we know that the marginal benefit of treating a patient at state $(11, 2, k)$ is higher than that of state $(11, 1, k)$ (where the latter is exactly $W_{11,1,k}$ by definition). Therefore the optimal action for state $(11, 2, k)$ is active, which leads to our conclusion. Note also that this argument only holds for F4 patients, because F0-F3 patients may transition to a more advanced disease state at the next period, where the Whittle's index could be higher and the optimal action may be active instead.

Finally, as the last structural property, we show that the Whittle's index is non-increasing in age as follows.

PROPOSITION 3. *Under Condition 2, we have* $W_{i,j,k+1} \leq W_{i,j,k}, \forall i, j, k$.

5. A Closed-Form Expression for Whittle's Index and A Capacity-Adjusted Index Policy

In this section, based on the structural properties we derived in §4.2, we first present a closed-form expression of Whittle's indices for F4 patients (§5.1). This closed-form expression is easy to interpret and provides an overall insight about how Whittle's index policy would behave in our problem. Then, motivated by the interpretation of this closed-form expression, we present a capacity-adjusted closed-form index policy (§5.2). Later in our numerical analysis section (§7), we show that our proposed policy overcomes the shortcomings of the Whittle's index and achieves a significant performance improvement.

5.1. A Closed-Form Whittle's Index for F4 Patients

As discussed in §1, although index-type policies decouple the original problem into sub-problems of individual bandits, the computation of the Whittle's index is still nontrivial since it involves repeated solution of dynamic programs for each individual bandit, which has a multi-dimensional state space in our case. Further, since there is no closed-form expression for Whittle's index for a generic RMAB problem, it is difficult to predict how the Whittle's index behaves and under what circumstances it is expected to perform well.

In this section, we derive a closed-form expression of Whittle's index for F4 patients based on the structural properties developed in §4.2. A closed-form expression of Whittle's index significantly expedites the computation. More importantly, it provides an intuitive interpretation, which helps us to predict when the Whittle's index may perform well and to also identify its shortcomings.

We focus on F4 patients mainly due to the following two reasons. First, we know from Proposition 1 that Whittle's indices are non-decreasing in health state. As such, it is most critical to prioritize among F4 patients and thus the indices for F4 patients (with different sentence lengths, ages and IDU status) would matter the most. Second, the structure of Whittle's indices for F4 patients is relatively simple compared with that of other patients. This is because for F0-F3 patients, they can further progress into more advanced disease states; thus to compute their indices, we also need to know the optimal actions of the more advanced disease states. On the other hand, given that the disease states are irreversible, the Whittle's index of F4 patients can be computed independent of the actions of other (less severe) disease states.

Before we derive the closed-form expression, we first make the following observations. First, we know from §4.2 that the Whittle's index of an F4 patient is non-decreasing in remaining sentence length and non-increasing in age. Second, the state transitions for remaining sentence length and age are deterministic: remaining sentence length decreases by one and age increases by one at each period. Therefore, the Whittle's index of an F4 patient is always decreasing over time. Note that this may not be true for F0-F3 patients since they may transition to more advanced disease states. Based on these observations, we derive a closed-form expression of Whittle's index for F4 patients as follows.

For any given state, if the passive subsidy level is equal to the Whittle's index of this state and the Whittle's index is decreasing over time, the optimal actions for all future states will be passive. Therefore, for any state (i,j,k) and any subsidy level W, we define $\hat{V}_{i,j,k}(W)$ as the value function when passive action is taken for all states; i.e., $\hat{V}_{i,j-1,k+1}(W) = \mathbf{e}_i^T \left(\sum_{j'=1}^{j-1} \beta^{j'-1} R^{j'-1} (r_{:,k+j'} + W) + \beta^{j-1} R^{j-1} z_{:,k+j} \right)$. Define $\hat{W}_{i,j,k}$ as the subsidy level that solves the following equation:

$$r_{i,k} + \beta \sum_{i'} (QR)_{i,i'} \hat{V}_{i',j-1,k+1}(W) = r_{i,k} + W + \beta \sum_{i'} R_{i,i'} \hat{V}_{i',j-1,k+1}(W).$$

After some algebra, we have:

$$\hat{W}_{i,j,k} = \beta \mathbf{e}_i^T (QR - R) \left(\sum_{j'=1}^{j-1} \beta^{j'-1} R^{j'-1} r_{:,k+j'} + \beta^{j-1} R^{j-1} z_{:,k+j} \right). \tag{7}$$

Note that $\beta \mathbf{e}_i^T QR \left(\sum_{j'=1}^{j-1} \beta^{j'-1} R^{j'-1} r_{:,k+j'} + \beta^{j-1} R^{j-1} z_{:,k+j} \right)$ represents the expected total QALYs of a patient if he is treated at the current period and there is no treatment at future periods inside the prison, while $\beta \mathbf{e}_i^T R \left(\sum_{j'=1}^{j-1} \beta^{j'-1} R^{j'-1} r_{:,k+j'} + \beta^{j-1} R^{j-1} z_{:,k+j} \right)$ represents the expected total QALYs of a patient if he is never treated before release. Therefore, the closed-form expression in Equation 7 has a very intuitive interpretation: the marginal benefit of treating a patient assuming no treatment at future periods before release. We next show that this expression coincides with the Whittle's index for F4 patients under conditions that are perfectly satisfied by real data.

THEOREM 2. *The closed-form expression in Equation 7 coincides with the Whittle's index for F4 patients (i.e.,* $\hat{W}_{i,j,k} = W_{i,j,k}, i = 11, \forall j, k$*) if* $\hat{W}_{i,j,k} \geq \hat{W}_{i,j-1,k+1}, i = 11, j \geq 2, \forall k$.

Proof. We prove this theorem by induction on sentence length j. First, by definition, it is straightforward to check that Equation 7 is exactly Whittle's index for $j = 1$. Assume that this is true for $1, ..., j-1$, we next show that it is also true for $j \geq 2$. Consider state (i, j, k) where $i = 11$. When the subsidy level is equal to $\hat{W}_{i,j,k}$, by induction assumption and the condition that $\hat{W}_{i,j,k} \geq \hat{W}_{i,j-1,k+1}$, the Whittle's indices for all future states are no greater than $\hat{W}_{i,j,k}$. Therefore, the optimal actions for all future states are passive. Then, by definition of $\hat{W}_{i,j,k}$, state (i, j, k) will be indifferent with the action. Therefore, we have $\hat{W}_{i,j,k} = W_{i,j,k}$ by definition of Whittle's index, which completes the proof. □

Theorem 2 says that the marginal benefit of treating a patient assuming no treatment at future periods before release is exactly equal to the Whittle's index for F4 patients, if this marginal benefit is non-increasing as remaining sentence length decreases and age increases. We remark that this condition is intuitive and is satisfied by real data.

The closed-form expression of Whittle's index significantly expedites the computation. More importantly, it provides us an intuitive interpretation that helps better understand when the Whittle's index may perform well and when not. In particular, while interpreting the Whittle's index as the marginal benefit of treatment is appealing from a practical perspective, the assumption that there is no treatment at future periods is only reasonable when the treatment capacity is very tight. This is because otherwise (i.e., when the treatment capacity is not very tight), patients with longer sentence lengths may also have the opportunity to be treated at future periods even if they miss the opportunity for treatment at the current period; in this case, it may make sense to prioritize patients with shorter sentence lengths, especially if access and/or linkage-to-care are low

in the community. However, the Whittle's index is independent of the treatment capacity, and does not capture the trade-off of chances for future treatments. Given this limitation, we anticipate the performance of the Whittle's index policy to be impaired as the treatment capacity increases. To overcome this shortcoming of the Whittle's index policy, we derive a capacity-adjusted closed-form index policy in the following section.

5.2. A Capacity-Adjusted Closed-Form Index Policy

Inspired by the closed-form expression of Whittle's index for F4 patients, in this subsection, we present a capacity-adjusted closed-form policy for all candidate patients (F0-F4). In particular, as in the closed-form expression in the previous section, we estimate the index of each patient based on the marginal benefit of treatment; however, instead of assuming no treatment at future periods, we adjust the probability of future treatment by explicitly considering the treatment capacity inside prison as follows. Let N' denote the number of total patients within prison that are eligible for treatment at the current period, and recall that M represents the number of patients that can be treated at each period. Define $\alpha = M/N'$ to be the proportion of patients that can be treated within prison. Then, α is a parameter that captures the treatment opportunity within prison. Therefore, instead of assuming no treatment within prison at future periods, we assume that each patient has a probability of α to be treated. In this way, our capacity-adjusted closed-form index for any given state (i, j, k) is defined as:

$$\bar{W}_{i,j,k} = \beta \mathbf{e}_i^T (QR - R) \left(\sum_{j'=1}^{j-1} \beta^{j'-1} (\alpha QR + (1-\alpha)R)^{j'-1} r_{:,k+j'} + \beta^{j-1} (\alpha QR + (1-\alpha)R)^{j-1} z_{:,k+j} \right). \tag{8}$$

From Equation 8, we observe that when the treatment rate $\alpha = 0$, we have $\bar{W}_{i,j,k} = W_{i,j,k}$ for $i = 11$, i.e., our capacity-adjusted index converges to the Whittle's index for F4 patients when the treatment capacity goes to zero. In this case, patients with longer sentence lengths should be prioritized (Proposition 2). However, as α increases, the chance of accessing to treatment at future periods for a patient with a long sentence length increases, and hence the marginal benefit of treatment at the current period for such a patient decreases. This implies, as α increases, the prioritization decision would tend to shift and patients with shorter sentence lenghts may start to be prioritized.

6. Simulation Model

In this section, we describe a realistic agent-based simulation (ABS) model that we use for testing the performance of our index-based policies. The simulation model allows capturing more details than the analytical model. For example, unlike our analytical modeling framework which only

considers the prison population explicitly and assumes that the arrival process is exogenous, the simulation model explicitly captures the dynamics of both prison and community population by tracking individual patients in the entire population. Hence, this simulation modeling framework is a good testbed for testing the policies generated based on our analytical model results.

This ABS model was built on an existing validated model by our group which was used to assess the potential health and economic impact of HCV screening and treatment in U.S. prisons and published in a top-tier clinical journal (He et al. 2015). For this study, we needed to modify and enhance the original simulation model to account for different treatment prioritization decisions inside prison. We further calibrated the simulation model to represent the prison system in Georgia, based on real data from the Georgia state prison system. In the following subsections, we describe the key components and modifications in the simulation model and refer the reader to He et al. (2015) for further details.

6.1. Baseline Population

We used a scale of 1:50 and generated 200,000 agents to represent the overall population in the state of GA, of whom 0.5% constituted the prison population (Guerino et al. 2011). We probabilistically assigned population characteristics to each of the agents, including age, gender, health state, incarceration status, and drug use behavior (see Tables 6-8 in Appendix C). In particular, we estimated the prevalence of HCV by age and gender from the National Health and Nutrition Examination Survey (NHANES) data, and defined the baseline distribution of the chronic HCV stages (fibrosis stages F0-F4, DC and HCC) using published studies (Denniston et al. 2014, Davis et al. 2010). We considered population growth by adding newborns to the population each year based on the annual birth rates in the US (World Bank 2014).

6.2. Disease Progression

As in the analytical model, the disease can progress through different stages of fibrosis, governed by a Markovian process. In the community, in line with current practice, patients with DC or HCC are assumed to be eligible for receiving a liver transplant. As in the analytical model, we continue to assume that fibrosis progression stops upon a successful treatment in non-cirrhotic patients (i.e., F0-F3), which is inline with the clinical evidence (Westbrook and Dusheiko 2014). On the other hand, unlike the analytical model, in the simulation model, we allow the disease to continue progress among cirrhotic patients (i.e., F4), though at a slower rate, even after a successful treatment. The estimated disease progression rates are presented in Table 9 in Appendix C.

6.3. Disease Transmission

Our agent-based model considered two types of HCV transmission: IDU-related and non-IDU-related, where IDU-related transmission contributes to 60% of all HCV infections. To simulate

HCV transmission, we constructed links among agents and probabilistically formed pairs between individuals, where IDUs had a higher probability of pairing with other IDUs, and vice versa. Once such a pair is formed, an infected individual could probabilistically transmit HCV to a susceptible individual, where the transmission probability depended on the disease state awareness of the infector, the IDU status of both infector and infectee, and whether the infectee has been infected before. The pairing and transmission related parameters are presented in Table 10 in Appendix C, where the unknown parameters are calibrated by matching the model output with the following target data: 1) the 10-year cumulative incidence of HCV, 2) the proportion of IDU-related transmission among all HCV incidences, and 3) the probability that a treated IDU can be reinfected (i.e., IDUs' reinfection rate).

6.4. Diagnosis and Treatment

As in our analytical model, we assume that diagnosed F0-F4 patients with a remaining sentence length of more than 12 months are eligible for treatment, and that the total number of patients that can be treated within the prison system is restricted by the allocated budget. In the community, patients can be diagnosed based on the current standard-of-care of HCV screening, including birth-cohort screening, risk-based screening, and usual care. We implemented the standard-of-care screening by probabilistically making unaware patients aware of infection status, and only patients aware of their infection are allowed to undergo treatment. The probabilities for diagnosis and treatment are presented in Table 11 in Appendix C.

6.5. Arrest and Release of Prisoners

We simulate movements of people in and out of the prison system by considering arrest and release of inmates. In particular, each year, agents in the community can get arrested with a certain probability, and those agents who have completed their assigned duration of sentence are released back to the community. We estimated the baseline crime probabilities for people with and without incarceration history using the Bureau of Justice Statistics (BJS) data and published studies (Langan and Levin 2002), and the probabilities are further adjusted based on the IDU status, to capture a larger probability of incarceration for IDUs. We estimate the length of sentence based on Georgia prison system data (Georgia Department of Corrections 2014). The probabilities of incarceration and length of sentence are presented in Tables 12-13 in Appendix C.

6.6. Injection Drug Users

Recall that in our analytical model, we assume that IDU status is a static characteristic of a patient that does not change over time. In the simulation model, we relax this assumption and explicitly model the dynamic change in the drug use behaviors. In particular, we allow the possibility that active IDUs could stop injecting drugs, while inactive IDUs or persons with no injection drug use

history could start injecting drugs by assigning probabilities of initiating and quitting injection drugs. Since there is no robust data for such probabilities, we use a calibration process to estimate their values. In particular, we ran our model with several combinations of these unknown parameters and selected the combination that matched model's projected annual incidence of drug use with the reported value of 0.115% per person-year.

6.7. Quality-Adjusted Life Years (QALYs)

QALYs are estimated based on liver health state and age from the published literature. The sources and the corresponding quality-of-life adjustment parameters are presented in Table 14 in Appendix C. Also, in the simulation model, we explicitly keep track of the new infections, and directly output the total QALYs gained in and out of the prison system and the QALY reductions due to new infections.

7. Computational Results and Policy Implications

In this section, we first analyze and compare the prioritization strategies proposed by our capacity-adjusted closed-form index policy and Whittle's index policy (§7.1). Then, using the simulation model described in §6, we test the performances of the capacity-adjusted index policy and compare it with other benchmark policies (§7.2). Lastly, we analyze how prioritization decisions change with respect to remaining sentence length and IDU status (§7.3 and 7.4). The estimation and values of parameters for the analytical model are included in Appendix B.

7.1. Prioritization Strategies under Whittle's vs. Capacity-Adjusted Index Policies

In this section, we analyze the prioritization strategies suggested by the Whittle's index policy (recall that the Whittle's index policy is independent of the treatment capacity), and compare it against our proposed policy, as the treatment capacity in prison vary. In particular, we consider the prioritization strategies given the health states and remaining length of sentences for an average-age (37-year-old), non-IDU patient (the patterns are similar for other age groups or IDUs). While the Whittle's indices for F4 patients and our proposed capacity-adjusted indices have closed-form expressions and can be computed very efficiently, the Whittle's indices for earlier stage patients are computed using a binary search approach, which involves repeated solution of single-bandit dynamic programs.

Figure 2 presents the prioritization strategies suggested by the Whittle's index policy, where darker areas indicate higher priority for treatment. From this figure, in line with our analytical results, we observe that Whittle's index policy i) prioritizes patients with more advanced fibrosis stages (Proposition 1), and ii) among F4 patients, prioritizes those with longer sentence lengths (Proposition 2). Note that for F0-F3 patients, the change of Whittle's indices with respect to

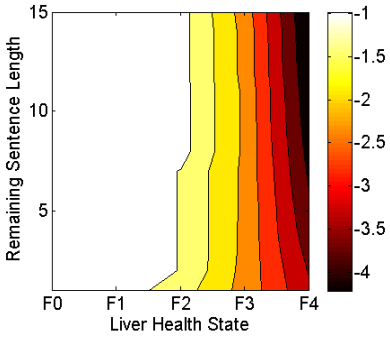
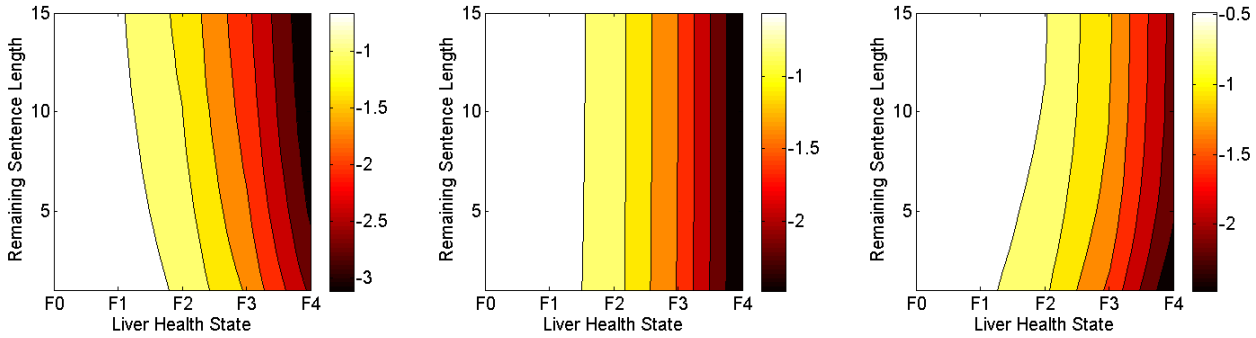


Figure 2 **Whittle's indices for average-age non-IDU patients with different health states and remaining sentence lengths.**

Note. The smoothed contour plot is presented for visualization purposes and the indices are well-defined for only discrete fibrosis state values.



(a) $\alpha = 0.05$ (b) $\alpha = 0.10$ (c) $\alpha = 0.15$

Figure 3 **Capacity-adjusted indices for average-age non-IDU patients with different health states and remaining sentence lengths under different treatment rates inside prison** $\alpha = 0.05, 0.10, 0.15$**.**

Note. Recall that $\alpha = M/N'$ represents the proportion of patients that can be treated within prison, where M is the total treatment capacity and N' is total number of patients eligible for treatment.

remaining sentence length is less structured and the monotonicity is indeed violated, as we anticipated.

Figure 3 presents the prioritization strategies suggested by our capacity-adjusted index policy as the treatment capacity varies. From these figures, we make the following observations: i) similar to the Whittle's index policy, our capacity-adjusted policy also suggests prioritizing patients with more advanced fibrosis stages, but ii) when the treatment capacity inside prison is very tight (Figure 3.a), among patients with the same health states, patients with longer remaining sentence lengths should be prioritized, and iii) as the treatment capacity increases, this strategy should shift and patients with shorter lengths of sentences should be given higher priority for treatment. Although this may appear counter-intuitive at first, it is reasonable and has an intuitive explanation: when the treatment capacity is very tight, patients with longer length of sentences may die in correctional

facilities before accessing to treatment, therefore they should be given priority. On the other hand, when the capacity in the prison system is not very tight, patients with longer sentence lengths are likely to have access to treatment before release at future periods; however, patients with shorter sentence lengths may not have access to treatment upon release due to coverage problems or lower linkage-to-care, therefore they should be given priority.

7.2. QALY Comparison of Various Policies

In this section, we test the performance of our proposed policy and compare with other benchmark policies using the simulation model described in §6. As described in §6, we generate 200,000 agents to represent the overall population, among which, about 0.5% (i.e., 1,000 agents) are inside prison, and similar to He et al. (2015), we consider a simulation horizon of 30 years. HCV prevalence within the prison system is estimated as 17.6%, among which about 75.5% have remaining sentence lengths of one year or longer, and are considered eligible for treatment. Therefore, there are initially about 133 patients within prison that are eligible for treatment. In practice, treatment capacity vary across prison systems. For example, a recent Wall Street Journal survey of all fifty state prison systems revealed that of the infected inmates, as low as less than 1% (Oklahoma) or as high as 13% (Indiana) are offered treatment in different prison systems (Loftus and Fields 2016). To capture this variation in practice, we consider five different annual treatment capacities $M = 1, 5, 10, 15, 20$.

First, we benchmark the performance of our policy with a *health state* policy, which prioritizes patients with more advanced fibrosis stages, and when patients are in the same fibrosis stage, prioritization is randomized. In current practice, while there are discussions around the role of sentence length, age, and IDU status in prioritization, there is no consensus and hence there exists variation. Therefore, in some sense, health state policy mimics the current practice. We also compare the performance of our proposed policy against myopic policy (defined as the one-period benefit of treatment, i.e., $\sum_{i'}(QR-R)_{i,i'}r_{i',k}$) and Whittle's index policy, both of which typically perform very well in bandit problems (Glazebrook et al. 2005, Archibald et al. 2009, Liu and Zhao 2010, Deo et al. 2013).

In order to compare the performances of the policies considered, we compute the *QALY gain* of each policy π over no treatment, defined as:

$$\text{QALY gain}(\pi) = \text{QALYs}(\pi) - \text{QALYs}(\text{no treatment}).$$

QALY gains of different policies are reported in Table 1 (and sensitivity analysis results against key model parameters are presented in Appendix D). From these results, we make several important observations. First, we observe that among all policies compared, health state policy performs

Table 1 Total QALY gains of different policies under different treatment capacities (numbers in parentheses are percentage improvements over health state policy).

Policies	$M=1$	$M=5$	$M=10$	$M=15$	$M=20$
Health state policy	37.6	167.0	282.1	359.3	415.0
Myopic policy	42.2 (12.1%)	172.4 (3.3%)	280.3 (-0.6%)	355.6 (-1.0%)	408.0 (-1.7%)
Whittle's index	43.8 (16.5%)	176.3 (5.6%)	285.3 (1.1%)	361.3 (0.5%)	412.7 (-0.6%)
Capacity-adjusted index	44.0 (17.1%)	177.7 (6.4%)	290.4 (2.9%)	368.3 (2.5%)	423.2 (2.0%)

Notes. Recall that M represents the annual treatment capacity, i.e., maximum number of infected inmates that can be treated every year. There are initially 133 patients eligible for treatment.

poorly, especially when the treatment capacity is very tight. This is important, because as we noted earlier, health state policy to some extent mimics the current practice.

Second, we observe that the myopic policy, which has been shown to perform very well in many other bandit problems (Liu and Zhao 2010, Deo et al. 2013), does not perform well in our case. In particular, it performs strictly worse than the Whittle's index and our capacity-adjusted index policies at all capacity levels. This finding implies that patients' remaining length of sentence, which is ignored by the myopic policy, should be an important factor in treatment prioritization decisions.

Third, we observe that the Whittle's index performs very well when the treatment capacity is very tight. However, as we anticipated based on the analytical results, the performance of the Whittle's index is significantly undermined as the treatment capacity increases. This is mainly because, as we have shown analytically, the Whittle's index for an F4 patient coincides with the marginal benefit of treating this patient assuming no future treatment before release. As a result, the Whittle's index tends to prioritize patients with longer remaining sentence lengths. However, the assumption of no future treatment opportunity is reasonable only when treatment capacity is extremely tight. Practically, this finding implies that patients with longest remaining length of sentences should be prioritized only when the likelihood of future treatment before release is very low.

Lastly, we observe that our proposed capacity-adjusted index policy performs very well at all treatment levels. In particular, when the treatment capacity is very tight, our proposed policy has a similar behavior and performance to that of the Whittle's index policy. This was expected because, as we showed analytically (Theorem 2), when the treatment capacity goes to zero, our capacity-adjusted index converges to Whittle's index for F4 patients. On the other hand, as the treatment capacity increases, our proposed policy tends to perform better than the Whittle's index, and the gap increases with the treatment capacity. This finding demonstrates that while most

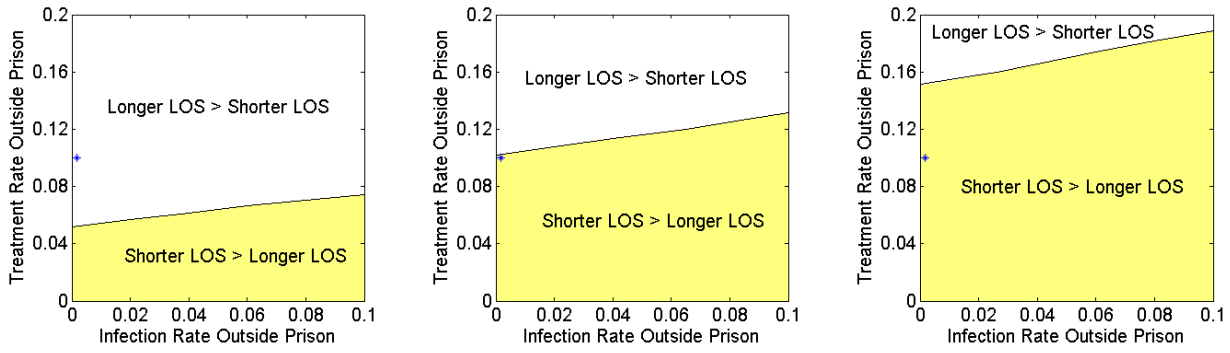
existing index-type policies are blind to the treatment capacity, in our setting treatment capacity is an important parameter in determining the prioritization rules.

7.3. Prioritization Between Patients with Shorter vs. Longer Sentence Lengths

As discussed in §1, there exists controversy about the roles of the remaining sentence length and IDU status in treatment prioritization decisions. In §7.2, we demonstrated that our proposed closed-form policy, which considers both of these factors, has a significant performance improvement over the health state policy. In this and the following subsections, we analyze the circumstances under which these factors would be most critical, and how prioritization decisions change with respect to these factors. Further, we quantify the marginal contributions of each of these factors.

Note that whether patients with shorter or longer remaining sentence lengths should be prioritized mainly depends on the differences of settings inside and outside the prison system. In that regard, given that the disease progression and treatment effectiveness are the same irrespective of the setting, the factors that could potentially drive this decision would include treatment rates inside and outside the prison, and the infection rate outside the prison. Intuitively, with everything else being held fixed and equal, one would expect that the smaller the treatment rate inside the prison system is, the higher the priority for patients with longer sentence lengths should be. Also, the smaller the treatment rate outside the prison system is, the higher the priority for patients with shorter sentence lengths should be. We also remark that while in general treatment capacity in the community may be higher than in the prison system, this may not necessarily translate to a higher treatment rate for released inmates. Indeed, the linkage-to-care level for released inmates is typically significantly lower than that in the prison setting (Liu et al. 2014). Lastly, in order to reduce the overall incidence of new infections outside, the larger the infection rate outside the prison system, the higher the priority for patients with shorter sentence lengths should be. However, it is not obvious whether and when patients with shorter or longer sentence lengths should be prioritized as all three of these factors change simultaneously. In order to address this question, we compared the prioritization decisions based on our index policy for patients with one- and five-year remaining sentence lengths, respectively, as the treatment and infection rates vary (we present the findings for average-age, non-IDU F4 patients, but the results are qualitatively similar for also other patients).

The results are presented in Figure 4. From this figure, we make several important observations. First, we observe that prioritization decisions have a simple threshold structure, where the threshold values can be easily characterized from the definition of our closed-form index policy. Second, we see that our intuition holds and that patients with shorter sentence lengths should have higher priority when the treatment capacity inside the prison system is not very tight, while in the community,



(a) $\alpha = 0.05$ (b) $\alpha = 0.10$ (c) $\alpha = 0.15$

Figure 4 Prioritization among patients with different remaining sentence lengths under different treatment rates inside prison $\alpha = 0.05, 0.10, 0.15$**.**

Notes. Presented results are for average-age, non-IDU F4 patients.
'*' represents the baseline parameter values, and LOS represents length of sentence.

treatment rate is relatively small. Third, we observe that, compared with the treatment rates (inside and outside), prioritization decisions are less sensitive to varying infection rates.

To quantify the marginal contribution of considering remaining sentence length in the prioritization decisions, we compare a policy that considers both health states and remaining sentence lengths, called *health state & sentence policy*, against the health state policy, which only considers health states in prioritization decisions. In particular, the health state & sentence policy is defined as our capacity-adjusted closed-form index for average-age non-IDU patients. The results are presented in Table 2.

Table 2 Total QALY gains under different treatment capacities/rates inside and outside prison (numbers in parentheses are percentage improvements over health state policy).

Treatment rate outside	Policies	$M = 5$	$M = 10$	$M = 15$
0.05	Health state policy	188.1	323.5	417.3
	Health state & sentence	193.5 (2.9%)	329.7 (1.9%)	428.0 (2.6%)
0.10	Health state policy	167.0	282.1	359.3
	Health state & sentence	174.5 (4.5%)	286.6 (1.6%)	365.4 (1.7%)
0.15	Health state policy	151.1	252.5	319.7
	Health state & sentence	160.3 (6.1%)	255.6 (1.2%)	322.6 (0.9%)

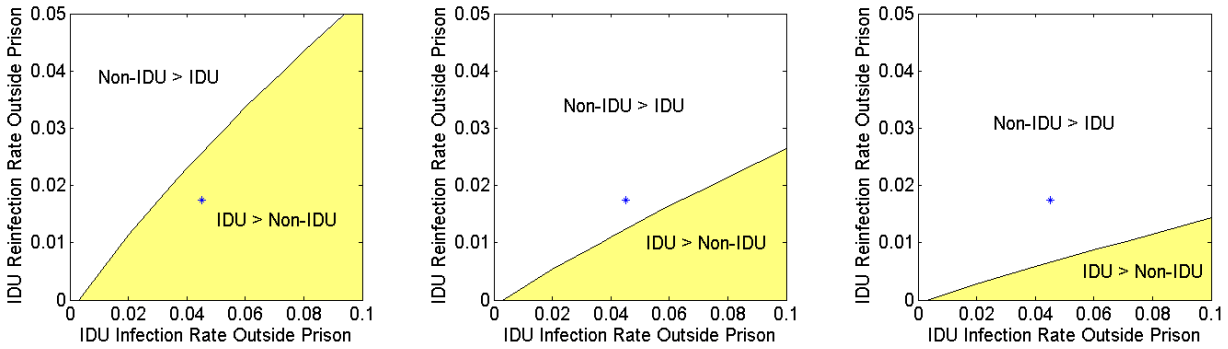
From the results, we observe that i) the performance improvement from considering remaining sentence length is significant, and ii) the performance improvement is large especially when the difference between the treatment rates inside and outside prison is large. These results suggest that considering remaining sentence length in the prioritization decisions is important, especially

when there is a substantial gap between the treatment opportunities inside and outside the prison system.

7.4. Prioritization Between IDU vs. Non-IDU Patients

As discussed earlier, if left untreated, released IDU patients may infect others after release. On the other hand, even if treated before release, the IDUs may get re-infected because of high-risk behavior, which may result in wastage of scarce resources. Therefore, whether and when IDUs should have priority over non-IDUs remains unclear.

We analyze the decision of whether and when IDUs should be prioritized by considering variable infection and reinfection rates for IDUs. The corresponding results are presented in Figure 5. First, as expected, we observe that the treatment prioritization decisions have a threshold structure: when the IDUs' infection rate is high and reinfection rate is low, IDUs should be prioritized; otherwise, non-IDUs should be prioritized. Second, we observe that among patients with less advanced fibrosis stages such as F2 (see Figure 5.a), treatment decisions are more sensitive to infection rate, and prioritizing IDUs is preferred at many combinations of parameter values. On the other hand, among patients with advanced fibrosis stages such as F4 (see Figure 5.c), prioritization decisions are more sensitive to reinfection rate, and prioritizing non-IDUs is preferred. We remark that these are important findings, because in current practice, there exists significant controversy regarding whether to prioritize IDUs over non-IDUs, and the main concern is the potential risk of reinfection. In that regard, our analysis allows quantification of the threshold risks of reinfection, above which non-IDUs should be prioritized. Further, our analysis shows that the role of reinfection differs for patients with different levels of fibrosis stages.



(a) F2 patients (b) F3 patients (c) F4 patients

Figure 5 Prioritization among patients who are IDUs and non-IDUs.

Note. '*' represents the baseline parameter values.
Presented results are for one-year-sentence, average-age patients.

To quantify the marginal contribution of considering IDU status in the prioritization decisions, similar to the previous subsection, we now compare a policy that considers both health states and IDU status, which we call *health state & IDU* policy, against the health state policy. The results are reported in Table 3.

Table 3 QALY gains under different IDUs' reinfection rates and treatment capacities (numbers in parentheses are percentage improvements over health state policy).

IDU reinfection rate	Policies	$M = 5$	$M = 10$	$M = 15$
0.05	Health state policy	162.7	275.8	351.4
	Health state & IDU	164.9 (1.3%)	276.0 (0.1%)	352.3 (0.2%)
0.10	Health state policy	157.0	267.1	340.3
	Health state & IDU	161.9 (3.1%)	268.8 (0.6%)	341.9 (0.5%)
0.15	Health state policy	152.1	258.3	330.0
	Health state & IDU	159.6 (4.9%)	261.6 (1.3%)	332.4 (0.7%)

From the results, we observe that the consideration of IDU status can lead to a significant performance improvement especially when the treatment capacity inside the prison system is tight and IDUs' reinfection rate is large. In particular, as IDUs' reinfection rate increases, the performance of the health state policy decreases very fast. On the other hand, the performance of the health state + IDU policy also decreases, but at a much slower rate. This is because when IDUs' reinfection rate becomes large, our policy prioritizes non-IDUs, so that the risk of reinfection is minimized. These results also highlight the importance of controlling reinfection of the disease for treated patients.

7.5. Prioritization Between Younger vs. Older Patients

Unlike remaining length of sentence and IDU status, the qualitative effect of considering age in prioritization decisions is more obvious: the younger the patients are prioritized, the higher the overall QALY gains would be. Indeed, we have already theoretically showed this result in Proposition 3. On the other hand, a prioritization rule based on age may be ethically controversial, and it is not obvious if consideration of age would lead to a significant QALY improvement to possibly rationalize such a prioritization policy. Therefore, in this section, we quantify the consideration of age in prioritization decisions.

In particular, similar to the previous subsections, we now compare a policy that considers both health states and age, which we call *health state & age* policy against the health state policy. The results are reported in Table 4 below. From the results, we observe that considering age in prioritization decisions can lead to a substantial QALY improvement, especially when the treatment capacity within prison is small. We believe this analysis may help policy-makers in deciding whether

such an improvement is large enough to rationalize an age-based policy and prioritize younger patients for treatment.

Table 4 QALY gains under different treatment capacities (numbers in parentheses are percentage improvements over health state policy).

Policies	$M = 5$	$M = 10$	$M = 15$
Health state policy	167.0	282.1	359.3
Health state & age	175.4 (5.1%)	288.0 (2.1%)	364.1 (1.3%)

8. Discussion and Conclusion

The concentration of HCV infection in the criminal justice system offers a unique opportunity to control the HCV epidemic. However, due to the prohibitive cost of the new generation HCV drugs and the tight resource constraints in U.S. prisons, universal treatment is not feasible. In current practice, standard treatment prioritization protocol is lacking, and hence there exists significant variation in treatment practices. Our study aims to fill this gap by providing a systematic algorithmic framework to support and standardize HCV treatment prioritization decisions in prisons. To solve this complex and computationally intractable problem, we develop a RMAB model and focus on index-type policies, which are appealing from both computation and implementation perspectives. We prove indexability for our problem and derive several structural properties for the well-known Whittle's index. We then present a closed-form expression of the Whittle's index for patients with late-stage fibrosis (F4), based on which we anticipate that the performance of the Whittle's index would degrade in our setting as the treatment capacity increases. Therefore, to address this problem, we propose a capacity-adjusted closed-form index policy. We test the performance of our proposed policy using a detailed realistic simulation model and show that our policy has a significant performance improvement over the current practice and other benchmark policies.

Our findings provide insights into the controversies in prioritization decisions for HCV treatment. First, we find that systematically considering other factors in addition to liver health state including remaining sentence length, IDU status and age can improve the current practice. Second, we show that patients with shorter sentence lengths should be prioritized when the linkage-to-care level outside prison is low while the treatment capacity in the prison system is not very tight. We remark that this is an important finding because while many other index-type policies are blind to the treatment capacity, we show that under different levels of treatment capacities, the prioritization decisions with respect to remaining sentence length can be significantly different. Third, we quantify the trade-off between treating/prioritizing IDUs versus non-IDUs and show that the decision has

a threshold structure. We further show that the decision is very sensitive to reinfection rate for patients with advanced liver disease, among whom IDUs should not be prioritized unless their reinfection is very-well controlled.

Due to simplicity in implementing prioritization policy, our work is appealing to several stakeholders including medical directors and policy makers at prisons. The authors presented this work at the 9th Academic & Health Policy Conference on Correctional Health, where the audience mostly comprised of practitioners, including policy makers and prison physicians. Based on their positive feedback, the authors believe that such index policies have high potential for implementation due to their simplicity and practicality. Such a scoring-based structured decision-making approach could also reduce prison-level variation in HCV treatment access and provide protection to prisons against litigation cases, in addition to potentially improving overall health outcomes of the society. Indeed, scoring based allocation policies are common in organ transplant (e.g., MELD scores in liver transplantation), and the authors hope to make a practical contribution in that regard by providing scientific basis for the use of scoring-based index policies for HCV treatment prioritization. Finally, our proposed modeling framework is flexible and can be extended to prioritization decisions in other resource-limited settings, such as HCV treatment in budget-constrained State Medicaid programs. In addition, this framework can also be extended to other infectious diseases that could be expensive to treat from a population level, such as a future treatment for chronic hepatitis B that can effectively cure infection .

Our study is not without limitations. First, while we show that prioritizing younger patients can lead to greater total health outcomes in the society, we do not directly capture fairness among patient with different ages. A simple approach to address this concern is to implement the index policy of an average-age patient for patients with different ages. In this way, age is excluded in the prioritization decisions. However, we remark that even in this case, it is important to include age into the state space. This is because if age is not considered and the lumpsum reward is estimated the same for patients with different ages, then patients with longer remaining sentence lengths will be implicitly prioritized since they can contribute more to the QALYs in prison while collecting the same amount of lumpsum reward after release. Second, we assume a constant treatment rate outside prison; however, in reality, the treatment probability may vary from person to person, depending on many factors such as insurance coverage, and ongoing substance abuse. While our modeling framework is flexible to capture this variation in practice, estimation of patient-specific treatment rates is difficult. To implicitly capture this variation and operationalize policies accordingly, our estimated treatment rate may be thought of representative for an average patient, and higher or lower lump-sum rewards can be used for patients with higher or lower likelihood of treatment. Third, we assume a negligible rate of in-prison HCV transmission. In a scenario where in-prison

IDU increases, such as relaxed policing for illicit drugs, this assumption may not hold and the model may no longer be appropriate. Nonetheless, in the current situation where industry has priced their products high, and state legislatures have not allocated adequate funds to prisons for universal treatment, we believe our model represents a reasonable strategy for prioritizing whom to treat.

Acknowledgments.

This study was supported in part by the National Science Foundation under award number 1452999. The authors thank Kan Li for his help on the agent-based simulation model; and Lauren Cipriano, Shan Liu, Qiushi Chen, Jan Vlachy, and the audience at the 9th Academic & Health Policy Conference on Correctional Health for their constructive comments, which were instrumental in improving our work.

References

Afdhal, N., S. Zeuzem, P. Kwo, M. Chojkier, N. Gitlin, M. Puoti, M. Romero-Gomez, J.P. Zarski, K. Agarwal, P. Buggisch, et al. 2014. Ledipasvir and sofosbuvir for untreated HCV genotype 1 infection. *New England Journal of Medicine* **370**(20) 1889–1898.

Akan, M., O. Alagoz, B. Ata, Fatih S. E., A. Said. 2012. A broader view of designing the liver allocation system. *Operations research* **60**(4) 757–770.

Alistar, S.S., E.F. Long, M.L. Brandeau, E.J. Beck. 2014. HIV epidemic control-a model for optimal allocation of prevention and treatment resources. *Health care management science* **17**(2) 162–181.

Anderson, R., I. Ashlagi, D. Gamarnik, A.E. Roth. 2015. Finding long chains in kidney exchange using the traveling salesman problem. *Proceedings of the National Academy of Sciences* **112**(3) 663–668.

Archibald, T.W., D.P. Black, K.D. Glazebrook. 2009. Indexability and index heuristics for a simple class of inventory routing problems. *Operations research* **57**(2) 314–326.

Argon, N.T., L. Ding, K.D. Glazebrook, S. Ziya. 2009. Dynamic routing of customers with general delay costs in a multiserver queuing system. *Probability in the Engineering and Informational Sciences* **23**(02) 175–203.

Arias, E. 2011. United States life tables, 2007. *National vital statistics reports: from the Centers for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System* **59**(9) 1–60.

Baker, S., J. Barnett, J.L. Gomez, et al. 2015. Strategies for addressing hepatitis C prevalence and treatment in American state prisons. URL https://wws.princeton.edu/sites/default/files/content/Domestic%20Workshop%20Report%202015.pdf.

Bedossa, P, T Poynard. 1996. An algorithm for the grading of activity in chronic hepatitis C. *Hepatology* **24**(2) 289–293.

Bertsimas, D., V.F. Farias, N. Trichakis. 2013. Fairness, efficiency, and flexibility in organ allocation for kidney transplantation. *Operations Research* **61**(1) 73–87.

Bertsimas, D., A. OHair, S. Relyea, J. Silberholz. 2016. An analytics approach to designing combination chemotherapy regimens for cancer. *Management Science* **62**(5) 1511–1531.

Bidell, Monique R, Milena McLaughlin, John Faragon, Caroline Morse, Nimish Patel. 2016. Desirable characteristics of hepatitis c treatment regimens: A review of what we have and what we need. *Infectious Diseases and Therapy* **5**(3) 299–312.

Bortfeld, T., Timothy C.Y. C., A. Trofimov, J.N. Tsitsiklis. 2008. Robust management of motion uncertainty in intensity-modulated radiation therapy. *Operations Research* **56**(6) 1461–1473.

Boucher, D. 2015. Inmates with hepatitis C sue Tennessee prison officials for treatment. URL http://www.tennessean.com/story/news/politics/2016/07/25/hepatitis-c-infected-inmates-sue-prison-officials-treatment/87527992/.

Brady, J.E., S.R. Friedman, H.L. Cooper, P.L. Flom, B. Tempalski, K. Gostnell. 2008. Estimating the prevalence of injection drug users in the us and in large us metropolitan areas from 1992 to 2002. *Journal of urban health* **85**(3) 323–351.

Brandeau, Margaret L, Gregory S Zaric, Vanda De Angelis. 2005. Improved allocation of HIV prevention resources: using information about prevention program production functions. *Health care management science* **8**(1) 19–28.

Cardoso, A.C., R. Moucari, C. Figueiredo-Mendes, M.P. Ripault, N. Giuily, C. Castelnau, N. Boyer, T. Asselah, M. Martinot-Peignoux, S. Maylin, et al. 2010. Impact of peginterferon and ribavirin therapy on hepatocellular carcinoma: incidence and survival in hepatitis C patients with advanced fibrosis. *Journal of hepatology* **52**(5) 652–657.

CDC. 2015. HIV and injection drug use in the United States. URL http://www.cdc.gov/hiv/risk/idu.html.

Chan, T.C.Y., T. Craig, T. Lee, M.B. Sharpe. 2014. Generalized inverse multiobjective optimization with application to cancer therapy. *Operations Research* **62**(3) 680–695.

Chhatwal, J., F. Kanwal, M.S. Roberts, M.A. Dunn. 2015. Cost-effectiveness and budget impact of hepatitis C virus treatment with sofosbuvir and ledipasvir in the United States. *Annals of internal medicine* **162**(6) 397–406.

Chong, C.A., A. Gulamhussein, E.J. Heathcote, L. Lilly, M. Sherman, G. Naglie, M. Krahn. 2003. Health-state utilities and quality of life in hepatitis C patients. *The American journal of gastroenterology* **98**(3) 630–638.

Cottrell, E.B., R. Chou, N. Wasson, B. Rahman, J.M. Guise. 2013. Reducing risk for mother-to-infant transmission of hepatitis C virus: a systematic review for the US Preventive Services Task Force. *Annals of internal medicine* **158**(2) 109–113.

Currie, S.L., J.C. Ryan, D. Tracy, T.L. Wright, S. George, R. McQuaid, M. Kim, H. Shen, A. Monto. 2008. A prospective study to examine persistent HCV reinfection in injection drug users who have previously cleared the virus. *Drug and alcohol dependence* **93**(1) 148–154.

Dai, Tinglong. 2012. The welfare consequences of the donor priority rule. *The 34th Annual Meeting of the Society for Medical Decision Making*. Smdm.

Davis, G.L., M.J. Alter, H.B. El-Serag, T. Poynard, L.W. Jennings. 2010. Aging of hepatitis C virus (HCV)-infected persons in the United States: a multiple cohort model of HCV prevalence and disease progression. *Gastroenterology* **138**(2) 513–521.

Denniston, M.M., R.B. Jiles, J. Drobeniuc, R.M. Klevens, J.W. Ward, G.M. McQuillan, S.D. Holmberg. 2014. Chronic hepatitis C virus infection in the United States, National Health and Nutrition Examination Survey 2003 to 2010. *Annals of internal medicine* **160**(5) 293–300.

Denniston, M.M., R.M. Klevens, G.M. McQuillan, R.B. Jiles. 2012. Awareness of infection, knowledge of hepatitis C, and medical follow-up among individuals testing positive for hepatitis C: National health and nutrition examination survey 2001-2008. *Hepatology* **55**(6) 1652–1661.

Deo, S., S. Iravani, T. Jiang, K. Smilowitz, S. Samuelson. 2013. Improving health outcomes through better capacity allocation in a community-based chronic care model. *Operations Research* **61**(6) 1277–1294.

Deo, S., K. Rajaram, S. Rath, U.S. Karmarkar, M.B. Goetz. 2015. Planning for HIV screening, testing, and care at the veterans health administration. *Operations Research* **63**(2) 287–304.

Deo, Sarang, Jeremie Gallien, Jonas Jonasson. 2016. Improving hiv early infant diagnosis supply chains in sub-saharan africa: Models and application to mozambique. *Working paper* .

Deo, Sarang, Milind Sohoni. 2015. Optimal decentralization of early infant diagnosis of HIV in resource-limited settings. *Manufacturing & Service Operations Management* **17**(2) 191–207.

Ding, L., K.D. Glazebrook. 2013. Dynamic routing in distinguishable parallel queues: an application of product returns for remanufacturing. *OR spectrum* **35**(3) 585–608.

Drummond, M.F., M.J. Sculpher, G.W. Torrance, B.J. O'Brien, G.L. Stoddart. 2005. *Methods for the economic evaluation of health care programmes*. Oxford University Press.

El-Serag, H.B. 2011. Hepatocellular carcinoma. *New England Journal of Medicine* **365**(12) 1118–1127.

Fattovich, G., G. Giustina, F. Degos, F. Tremolada, G. Diodati, P. Almasio, F. Nevens, A. Solinas, D. Mura, J.T. Brouwer, et al. 1997. Morbidity and mortality in compensated cirrhosis type c: a retrospective follow-up study of 384 patients. *Gastroenterology* **112**(2) 463–472.

Federal Bureau of Prisons. 2016. Evaluation and management of chronic hepatitis C virus (HCV) infection: Clinical practice guidelines.

Gane, E.J., C.A. Stedman, R.H. Hyland, X. Ding, E. Svarovskaia, W.T. Symonds, R.G. Hindes, M.M Berrey. 2013. Nucleotide polymerase inhibitor sofosbuvir plus ribavirin for hepatitis C. *New England Journal of Medicine* **368**(1) 34–44.

Georgia Department of Corrections. 2014. Inmate statistical profile. URL http://www.dcor.state.ga.us/sites/all/files/pdf/Research/Annual/Profile_inmate_admissions_FY2014.pdf.

Ghany, M.G., D.B. Strader, D.L. Thomas, L.B. Seeff. 2009. Diagnosis, management, and treatment of hepatitis C: an update. *Hepatology* **49**(4) 1335–1374.

Glazebrook, K.D., C. Kirkbride, J. Ouenniche. 2009. Index policies for the admission control and routing of impatient customers to heterogeneous service stations. *Operations Research* **57**(4) 975–989.

Glazebrook, K.D., H.M. Mitchell, P.S. Ansell. 2005. Index policies for the maintenance of a collection of machines by a set of repairmen. *European Journal of Operational Research* **165**(1) 267–284.

Glazebrook, K.D., D. Ruiz-Hernandez, C. Kirkbride. 2006. Some indexable families of restless bandit problems. *Advances in Applied Probability* 643–672.

Gold, M.R., J.E. Siegel, L.B. Russell, M.C. Weinstein. 1996. *Cost-effectiveness in health and medicine*.

Guerino, P., P.M. Harrison, W.J. Sabol. 2011. Prisoners in 2010. *Bureau of Justice Statistics, Washington, DC* .

Hanmer, J., W.F. Lawrence, J.P. Anderson, R.M. Kaplan, D.G. Fryback. 2006. Report of nationally representative values for the noninstitutionalized us adult population for 7 health-related quality-of-life scores. *Medical Decision Making* **26**(4) 391–400.

He, T., K. Li, M.S. Roberts, A.C. Spaulding, T. Ayer, J.J. Grefenstette, J. Chhatwal. 2015. Prevention of hepatitis C by screening and treatment in US prisons. *Annals of internal medicine* .

Hellard, M., R. Sacks-Davis, J. Doyle. 2016. Hepatitis C elimination by 2030 through treatment and prevention: think global, act in local networks. *Journal of Epidemiology and Community Health* jech–2015.

Helm, J.E., M.S. Lavieri, M.P. Van Oyen, J.D. Stein, D.C. Musch. 2015. Dynamic forecasting and control algorithms of glaucoma progression for clinician decision support. *Operations Research* **63**(5) 979–999.

Horberg, M. 2014. Kaiser permanente mid-atlantic states hepatitis C program. *Presented at National Summit on HCV Prevention and Cure 2014, Atlanta, GA*. URL http://www.viralhepatitisaction.org/national-summit-hcv-prevention-and-cure-presentations.

Howden, L.M., J.A. Meyer. 2010. Age and sex composition: 2010. *2010 Census Briefs, US Department of Commerce, Economics and Statistics Administration. US Census Bureau* .

Jayasekera, C.R., M. Barry, L.R. Roberts, M.H. Nguyen. 2014. Treating hepatitis C in lower-income countries. *N Engl J Med* **370**(20) 1869–71.

Kabiri, M., A.B. Jazwinski, M.S. Roberts, A.J. Schaefer, J. Chhatwal. 2014. The changing burden of hepatitis C virus infection in the United States: model-based predictions. *Annals of internal medicine* **161**(3) 170–180.

Kaplan, Edward H, Harold Pollack. 1998. Allocating HIV prevention resources. *Socio-Economic Planning Sciences* **32**(4) 257–263.

Kaplan, E.H., M.H. Merson. 2002. Allocating HIV-prevention resources: balancing efficiency and equity. *American Journal of Public Health* **92**(12) 1905–1907.

Katehakis, Michael N, Cyrus Derman. 1986. Computing optimal sequential allocation rules in clinical trials. *Lecture notes-monograph series* 29–39.

Khademi, Amin, Denis R Saure, Andrew J Schaefer, Ronald S Braithwaite, Mark S Roberts. 2015. The price of nonabandonment: HIV in resource-limited settings. *Manufacturing & Service Operations Management* **17**(4) 554–570.

Lang, K., N. Danchenko, K. Gondek, S. Shah, D. Thompson. 2009. The burden of illness associated with hepatocellular carcinoma in the United States. *Journal of hepatology* **50**(1) 89–99.

Langan, P.A., D.J. Levin. 2002. Recidivism of prisoners released in 1994. *Federal Sentencing Reporter* **15**(1) 58–65.

Larney, Sarah, Curt G. Beckwith, Nickolas D. Zaller, Brian T. Montague, Josiah Rich. 2014. seek, test, treat and retain for hepatitis c in the united states criminal justice system. *International journal of prisoner health* **10**(3) 164–171.

Lavieri, MS, ML Puterman, S Tyldesley, WJ Morris. 2012. When to treat prostate cancer patients based on their psa dynamics. *IIE Transactions on Healthcare Systems Engineering* **2**(1) 62–77.

Lawitz, E., A. Mangia, D. Wyles, M. Rodriguez-Torres, T. Hassanein, S.C. Gordon, M Schultz, M.N. Davis, Z. Kayali, K.R Reddy, et al. 2013. Sofosbuvir for previously untreated chronic hepatitis C infection. *New England Journal of Medicine* **368**(20) 1878–1887.

Lee, Chris P, Glenn M Chertow, Stefanos A Zenios. 2008. Optimal initiation and management of dialysis therapy. *Operations Research* **56**(6) 1428–1449.

Liu, K., Q. Zhao. 2010. Indexability of restless bandit problems and optimality of whittle index for dynamic multichannel access. *Information Theory, IEEE Transactions on* **56**(11) 5547–5567.

Liu, S., M.L. Brandeau, J.D. Goldhaber-Fiebert. 2015. Optimizing patient treatment decisions in an era of rapid technological advances: the case of hepatitis C treatment. *Health care management science* 1–17.

Liu, S., D. Watcha, M. Holodniy, J.D. Goldhaber-Fiebert. 2014. Sofosbuvir-based treatment regimens for chronic, genotype 1 hepatitis C virus infection in us incarcerated populations: a cost-effectiveness analysis. *Annals of internal medicine* **161**(8) 546–553.

Loftus, P. 2015. Prisoners sue Massachusetts for withholding hepatitis C drugs. *The Wall Street Journal* June 11, 2015.URL http://blogs.wsj.com/pharmalot/2015/06/11/prisoners-sue-massachusetts-for-withholding-hepatitis-c-drugs/.

Loftus, P., G. Fields. 2016. High cost of new hepatitis C drugs strains prison budgets, locks many out of cure. *The Wall Street Journal* Sep 12, 2016.URL http://www.wsj.com/articles/high-cost-of-new-hepatitis-c-drugs-strains-prison-budgets-locks-many-out-of-cure-1473701644.

Lok, A.S., L.B. Seeff, T.R. Morgan, Adrian M. D.B., R.K. Sterling, T.M. Curto, G.T. Everson, K.L. Lindsay, W.M. Lee, H.L. Bonkovsky, et al. 2009. Incidence of hepatocellular carcinoma and associated risk factors in hepatitis C-related advanced liver disease. *Gastroenterology* **136**(1) 138–148.

Mason, J.E., B.T. Denton, N.D. Shah, S.A. Smith. 2014. Optimizing the simultaneous management of blood pressure and cholesterol for type 2 diabetes patients. *European Journal of Operational Research* **233**(3) 727–738.

Mitruka, K., T. Tsertsvadze, M. Butsashvili, A. Gamkrelidze, P. Sabelashvili, E. Adamia, M. Chokheli, J. Drobeniuc, L. Hagan, A.M. Harris, et al. 2015. Launch of a nationwide hepatitis C elimination programgeorgia, april 2015. *MMWR Morb Mortal Wkly Rep* **64** 753–7.

Mumola, C.J., J.C. Karberg, et al. 2006. *Drug use and dependence, state and federal prisoners, 2004*. US Department of Justice, Office of Justice Programs, Bureau of Justice Statistics Washington, DC.

Nainan, O.V., M.J. Alter, D. Kruszon-Moran, F. Gao, G. Xia, G. McQuillan, H.S. Margolis. 2006. Hepatitis C virus genotypes and viral concentrations in participants of a general population survey in the United States. *Gastroenterology* **131**(2) 478–484.

Page, K., M.D. Morris, J.A. Hahn, L. Maher, M. Prins. 2013. Injection drug use and hepatitis C virus infection in young adult injectors: using evidence to inform comprehensive prevention. *Clinical infectious diseases* **57**(suppl 2) S32–S38.

Papadimitriou, C.H., J.N. Tsitsiklis. 1999. The complexity of optimal queuing network control. *Mathematics of Operations Research* **24**(2) 293–305.

Planas, R., B. Ballesté, M.A. Álvarez, M. Rivera, S. Montoliu, J.A. Galeras, J. Santos, S. Coll, R.M. Morillas, R. Solà. 2004. Natural history of decompensated hepatitis C virus-related cirrhosis. a study of 200 patients. *Journal of hepatology* **40**(5) 823–830.

Rich, J.D., S.A. Allen, B.A. Williams. 2014. Responding to hepatitis C through the criminal justice system. *N Engl J Med* **370**(20) 1871–1874.

Romeijn, H Edwin, Ravindra K Ahuja, James F Dempsey, Arvind Kumar. 2006. A new linear programming approach to radiation therapy treatment planning problems. *Operations Research* **54**(2) 201–216.

Shechter, S.M., M.D. Bailey, A.J. Schaefer, M.S. Roberts. 2008. The optimal time to initiate HIV therapy under ordered health states. *Operations Research* **56**(1) 20–33.

Silverman, E. 2014. Will Oregon restrict coverage for a pricey gilead pill for hepatitis C? *The Wall Street Journal* Jun 12, 2014.

Snowbeck, C. 2015. Minnesota prison inmates sue to gain access to costly hepatitis C medications. *Star Tribune* May 26, 2015.URL http://www.startribune.com/minnesota-prison-inmates-sue-to-gain-access-to-costly-hepatitis-c-medications/305114781/.

Spaulding, A.C., D.L. Thomas. 2012. Screening for HCV infection in jails. *JAMA* **307**(12) 1259–1260.

Su, X., S. Zenios. 2004. Patient choice in kidney allocation: The role of the queueing discipline. *Manufacturing & Service Operations Management* **6**(4) 280–301.

Su, Xuanming, Stefanos A Zenios. 2005. Patient choice in kidney allocation: A sequential stochastic assignment model. *Operations research* **53**(3) 443–455.

Taskin, Z.C., J.C. S., H.E. R., J.F. Dempsey. 2010. Optimal multileaf collimator leaf sequencing in imrt treatment planning. *Operations Research* **58**(3) 674–690.

Thein, H., Q. Yi, G.J. Dore, Murray D. K. 2008. Estimation of stage-specific fibrosis progression rates in chronic hepatitis C virus infection: a meta-analysis and meta-regression. *Hepatology* **48**(2) 418–431.

Thuluvath, P.J., M.K. Guidinger, J.J. Fung, L.B. Johnson, S.C. Rayhill, S.J. Pelletier. 2010. Liver transplantation in the United States, 1999–2008. *American Journal of Transplantation* **10**(4p2) 1003–1019.

Varan, A.K., D.W. Mercer, M.S. Stein, A.C. Spaulding. 2014. Hepatitis C seroprevalence among prison inmates since 2001: still high but declining. *Public health reports* **129**(2) 187–195.

Wedemeyer, H., G.J. Dore, J.W. Ward. 2015. Estimates on HCV disease burden worldwide–filling the gaps. *Journal of viral hepatitis* **22**(s1) 1–5.

West, H.C., W.J. Sabol, S.J. Greenman. 2010. Prisoners in 2009. *Bureau of justice statistics bulletin* 1–38.

Westbrook, Rachel H, Geoffrey Dusheiko. 2014. Natural history of hepatitis c. *Journal of hepatology* **61**(1) S58–S68.

Whittle, P. 1988. Restless bandits: Activity allocation in a changing world. *Journal of applied probability* 287–298.

Wolfe, R.A., E.C. Roys, R.M. Merion. 2010. Trends in organ donation and transplantation in the United States, 1999–2008. *American journal of transplantation* **10**(4p2) 961–972.

World Bank. 2014. Birth rate, crude (per 1,000 people). URL http://data.worldbank.org/indicator/SP.DYN.CBRT.IN.

Yehia, Baligh R, Asher J Schranz, Craig A Umscheid, Vincent Lo Re III. 2014. The treatment cascade for chronic hepatitis c virus infection in the united states: a systematic review and meta-analysis. *PloS one* **9**(7) e101554.

Zenios, S.A., G.M. Chertow, L.M. Wein. 2000. Dynamic allocation of kidneys to candidates on the transplant waiting list. *Operations Research* **48**(4) 549–569.

Appendix. A. Proofs of Analytical Results

Proof of Lemma 1. First, consider non-candidate health state i. Since we assume no reinfection within prison, once a patient transitions to a non-candidate state, he will always stay at non-candidate states in future periods. Further, when $W \geq 0$, passive is always optimal for all non-candidate states. Therefore, for $W \geq 0$, we have $V_{i,j,k}(W) = \hat{V}_{i,j,k}(W)$. Thus for $0 \leq W_1 \leq W_2$, we have $V_{i,j,k}(W_2) - V_{i,j,k}(W_1) = \hat{V}_{i,j,k}(W_2) - \hat{V}_{i,j,k}(W_1)$.

Now consider candidate health state $i = 7, ..., 11$. Let $V_{i,j,k}(W) \doteq V^r_{i,j,k}(W) + V^w_{i,j,k}(W)$, where $V^r_{i,j,k}(W)$ denotes the pure reward part (excluding passive subsidy) of $V_{i,j,k}(W)$, while $V^w_{i,j,k}(W)$ denotes the subsidy part of $V_{i,j,k}(W)$. Then, we have $V^r_{i,j,k}(W_2) + V^w_{i,j,k}(W_2) \geq V^r_{i,j,k}(W_1) + \frac{W_2}{W_1} V^w_{i,j,k}(W_1)$, because the left hand side is the optimal value function for state (i,j,k) when the subsidy is W_2, while the right hand side is the total reward (including subsidy) when the passive subsidy is W_2 but the optimal policy for W_1 is applied.

Similarly, let $\hat{V}_{i,j,k}(W) \equiv \hat{V}^r_{i,j,k}(W) + \hat{V}^w_{i,j,k}(W)$, where $\hat{V}^r_{i,j,k}(W)$ denotes the pure reward part (excluding passive subsidy) of $\hat{V}_{i,j,k}(W)$, while $\hat{V}^w_{i,j,k}(W)$ denotes the subsidy part of $\hat{V}_{i,j,k}(W)$. Then, by definition of $\hat{V}_{i,j,k}(W)$, $\hat{V}^r_{i,j,k}(W)$ is independent of W, while $\hat{V}^w_{i,j,k}(W)$ is linear in W. Therefore, we have $\hat{V}^r_{i,j,k}(W_2) + \hat{V}^w_{i,j,k}(W_2) = \hat{V}^r_{i,j,k}(W_1) + \frac{W_2}{W_1} \hat{V}^w_{i,j,k}(W_1)$.

Finally, by definition of $V_{i,j,k}(W)$ and $\hat{V}_{i,j,k}(W)$, we clearly have $V_{i,j,k}(W) \geq \hat{V}_{i,j,k}(W)$. Further, by Condition 1, the total pure reward (excluding passive subsidy) is maximized when active action is taken in all candidate states, i.e., $V^r_{i,j,k}(W) \leq \hat{V}^r_{i,j,k}(W)$. Therefore, we must have $V^w_{i,j,k}(W) \geq \hat{V}^w_{i,j,k}(W)$. Then, for any $0 \leq W_1 \leq W_2$:

$$
\begin{aligned}
& V_{i,j,k}(W_2) - V_{i,j,k}(W_1) \\
=& V^r_{i,j,k}(W_2) + V^w_{i,j,k}(W_2) - V^r_{i,j,k}(W_1) - V^w_{i,j,k}(W_1) \\
\geq& V^r_{i,j,k}(W_1) + \frac{W_2}{W_1} V^w_{i,j,k}(W_1) - V^r_{i,j,k}(W_1) - V^w_{i,j,k}(W_1) \\
=& (\frac{W_2}{W_1} - 1) V^w_{i,j,k}(W_1) \\
\geq& (\frac{W_2}{W_1} - 1) \hat{V}^w_{i,j,k}(W_1) \\
=& \hat{V}^r_{i,j,k}(W_1) + \frac{W_2}{W_1} \hat{V}^w_{i,j,k}(W_1) - \hat{V}^r_{i,j,k}(W_1) - \hat{V}^w_{i,j,k}(W_1) \\
=& \hat{V}^r_{i,j,k}(W_2) + \hat{V}^w_{i,j,k}(W_2) - \hat{V}^r_{i,j,k}(W_1) - \hat{V}^w_{i,j,k}(W_1) \\
=& \hat{V}_{i,j,k}(W_2) - \hat{V}_{i,j,k}(W_1),
\end{aligned}
$$

which concludes the proof. □

Proof of Lemma 2. We prove this lemma by induction on j. For $j = 1$ and for any candidate health state i, we have $\hat{V}_{i,1,k}(W_2) - \hat{V}_{i,1,k}(W_1) = r^a_{i,1,k} - r^a_{i,1,k} = 0$, while for any non-candidate health state i, we have $\hat{V}_{i,1,k}(W_2) - \hat{V}_{i,1,k}(W_1) = W_2 + r^b_{i,1,k} - W_1 - r^b_{i,1,k} = W_2 - W_1$. Thus, the equation holds for $j = 1$.

For $j \geq 2$, suppose the equation holds for any i and some $j-1$, i.e.:

$$
\hat{V}_{i,j-1,k+1}(W_2) - \hat{V}_{i,j-1,k+1}(W_1) = (W_2 - W_1) \mathbf{e}_i^T \left(\sum_{j'=1}^{j-1} \beta^{j'-1} (QR)^{j'-1} \mathbf{e}_0 \right).
$$

Then, for any candidate health state $i = 7, ..., 11$, we have:

$$\begin{aligned}
&\hat{V}_{i,j,k}(W_2) - \hat{V}_{i,j,k}(W_1) \\
=&r_{i,k} + \beta \sum_{i'} (QR)_{i,i'} \hat{V}_{i',j-1,k+1}(W_2) - r_{i,k} - \beta \sum_{i'} (QR)_{i,i'} \hat{V}_{i',j-1,k+1}(W_1) \\
=&\beta \sum_{i'} (QR)_{i,i'} (\hat{V}_{i',j-1,k+1}(W_2) - \hat{V}_{i',j-1,k+1}(W_1)) \\
=&\beta (W_2 - W_1) \mathbf{e}_i^T QR \left(\sum_{j'=1}^{j-1} \beta^{j'-1} (QR)^{j'-1} \mathbf{e}_0 \right) \\
=&(W_2 - W_1) \mathbf{e}_i^T \left(\sum_{j'=2}^{j} \beta^{j'-1} (QR)^{j'-1} \mathbf{e}_0 \right) \\
=&(W_2 - W_1) \mathbf{e}_i^T \left(\sum_{j'=1}^{j} \beta^{j'-1} (QR)^{j'-1} \mathbf{e}_0 \right).
\end{aligned}$$

For any non-candidate health state i, we have:

$$\begin{aligned}
&\hat{V}_{i,j,k}(W_2) - \hat{V}_{i,j,k}(W_1) \\
=&r_{i,k} + W_2 + \beta \sum_{i'} R_{i,i'} \hat{V}_{i',j-1,k+1}(W_2) - r_{i,k} - W_1 - \beta \sum_{i'} R_{i,i'} \hat{V}_{i',j-1,k+1}(W_1) \\
=&W_2 - W_1 + \beta \sum_{i'} R_{i,i'} (\hat{V}_{i',j-1,k+1}(W_2) - \hat{V}_{i',j-1,k+1}(W_1)) \\
=&W_2 - W_1 + \beta \sum_{i'} (QR)_{i,i'} (\hat{V}_{i',j-1,k+1}(W_2) - \hat{V}_{i',j-1,k+1}(W_1)) \\
=&W_2 - W_1 + \beta (W_2 - W_1) \mathbf{e}_i^T QR \left(\sum_{j'=1}^{j-1} \beta^{j'-1} (QR)^{j'-1} \mathbf{e}_0 \right) \\
=&W_2 - W_1 + (W_2 - W_1) \mathbf{e}_i^T \left(\sum_{j'=2}^{j} \beta^{j'-1} (QR)^{j'-1} \mathbf{e}_0 \right) \\
=&(W_2 - W_1) \mathbf{e}_i^T \left(\sum_{j'=1}^{j} \beta^{j'-1} (QR)^{j'-1} \mathbf{e}_0 \right),
\end{aligned}$$

which concludes the proof. □

Proof of Lemma 3. For non-candidate states i, the left hand side of the inequality is equal to zero and the inequality holds trivially. Therefore, it is sufficient to focus on candidate health states $i = 7, ..., 11$. Recall that $\mathbf{e}_0$ is a column vector with all entries of non-candidate health states equal to 1 and all entries of candidate health states equal to 0. Then intuitively, $\beta \mathbf{e}_i^T QR \left(\sum_{j'=1}^{j-1} \beta^{j'-1} (QR)^{j'-1} \mathbf{e}_0 \right)$ represents the total units of subsidy collected by a patient if he is always treated whenever at a candidate state, while $1 + \beta \mathbf{e}_i^T R \left(\sum_{j'=1}^{j-1} \beta^{j'-1} (QR)^{j'-1} \mathbf{e}_0 \right)$ represents the total units of subsidy collected by a patient if he is not treated at the current period but is always treated in the following periods whenever at a candidate state. Then, the inequality in Lemma 3 says that the subsidy collected in the latter case is larger.

Recall that $Q_{i,i-5} + Q_{i,i} = 1$. Then, we have:

$$\beta \mathbf{e}_i^T (QR - R) \left(\sum_{j'=1}^{j-1} \beta^{j'-1} (QR)^{j'-1} \mathbf{e}_0 \right)$$

$$
\begin{aligned}
=&\beta\sum_{i'}(Q_{i,i-5}R_{i-5,i'}+Q_{i,i}R_{i,i'}-R_{i,i'})\Bigg(\sum_{j'=1}^{j-1}\beta^{j'-1}(QR)^{j'-1}\mathbf{e}_0\Bigg)_{i'}\\
=&\beta Q_{i,i-5}\sum_{i'}(R_{i-5,i'}-R_{i,i'})\Bigg(\sum_{j'=1}^{j-1}\beta^{j'-1}(QR)^{j'-1}\mathbf{e}_0\Bigg)_{i'}.
\end{aligned}
$$

Since we assume no reinfection within prison, if a patient at state $i=7,...,11$ achieves SVR and transitions to health state $i-5$, he will always stay at non-candidate states. In this case, subsidies will be collected at all of the following periods. Therefore:

$$
\sum_{i'}R_{i-5,i'}\Bigg(\sum_{j'=1}^{j-1}\beta^{j'-1}(QR)^{j'-1}\mathbf{e}_0\Bigg)_{i'}=\sum_{j'=1}^{j-1}\beta^{j'-1}.
$$

On the other hand, if the patient is not treated or does not achieve SVR, he may still be at a candidate state at the next period. In this case, he may only start collecting subsidies after he is treated and achieves SVR or after he transitions to state $i=12,13,14$. To find a lower bound on the total subsidies collected in this case, we only consider the subsidies collected if he is treated and achieves SVR. Then, we have:

$$
\begin{aligned}
&\sum_{i'}R_{i,i'}\Bigg(\sum_{j'=1}^{j-1}\beta^{j'-1}(QR)^{j'-1}\mathbf{e}_0\Bigg)_{i'}\\
\geq&\sum_{j'=2}^{j-1}\beta^{j'-1}\Bigg(Q_{i,i-5}+(1-Q_{i,i-5})Q_{i,i-5}+...+(1-Q_{i,i-5})^{j'-2}Q_{i,i-5}\Bigg)\\
=&Q_{i,i-5}\sum_{j'=2}^{j-1}\beta^{j'-1}\Bigg(1+(1-Q_{i,i-5})+...+(1-Q_{i,i-5})^{j'-2}\Bigg).
\end{aligned}
$$

Therefore, we have:

$$
\begin{aligned}
&\beta Q_{i,i-5}\sum_{i'}(R_{i-5,i'}-R_{i,i'})\Bigg(\sum_{j'=1}^{j-1}\beta^{j'-1}(QR)^{j'-1}\mathbf{e}_0\Bigg)_{i'}\\
\leq&Q_{i,i-5}\sum_{j'=1}^{j-1}\beta^{j'}-Q_{i,i-5}^2\sum_{j'=2}^{j-1}\beta^{j'}\Bigg(1+(1-Q_{i,i-5})+...+(1-Q_{i,i-5})^{j'-2}\Bigg)\\
=&Q_{i,i-5}\beta+Q_{i,i-5}\sum_{j'=2}^{j-1}\beta^{j'}+Q_{i,i-5}(1-Q_{i,i-5})\sum_{j'=2}^{j-1}\beta^{j'}\Bigg(1+(1-Q_{i,i-5})+...+(1-Q_{i,i-5})^{j'-2}\Bigg)\\
&-Q_{i,i-5}\sum_{j'=2}^{j-1}\beta^{j'}\Bigg(1+(1-Q_{i,i-5})+...+(1-Q_{i,i-5})^{j'-2}\Bigg)\\
=&Q_{i,i-5}\beta+Q_{i,i-5}(1-Q_{i,i-5})\sum_{j'=2}^{j-1}\beta^{j'}(1-Q_{i,i-5})^{j'-2}\\
=&Q_{i,i-5}\sum_{j'=1}^{j-1}\beta^{j'}(1-Q_{i,i-5})^{j'-1}\\
\leq&\frac{\beta Q_{i,i-5}}{1-\beta(1-Q_{i,i-5})}\\
\leq&1,
\end{aligned}
$$

which completes the proof. $\square$

Proof of Proposition 1. For $i=7,...,10$, to show $W_{i+1,j,k} \geq W_{i,j,k}$, it is sufficient to show that when the subsidy level is equal to $W_{i,j,k}$, the optimal action for state $(i+1,j,k)$ is active. Similar as before, let $V_{i,j,k}(W)$ be the optimal value function at state (i,j,k) when the passive subsidy is W.

For $j=1$, when passive subsidy is $W_{i,1,k}$, the difference between the rewards of selecting active and passive actions for state $(i+1,1,k)$ is:

$$
\begin{aligned}
& r_{i+1,k}+\beta \sum_{i'}(QR)_{i+1,i'} z_{i',k+1}-r_{i+1,k}-\beta \sum_{i'} R_{i+1,i'} z_{i',k+1}-W_{i,1,k} \\
=&\beta \sum_{i'}(QR)_{i+1,i'} z_{i',k+1}-\beta \sum_{i'} R_{i+1,i'} z_{i',k+1}-W_{i,1,k} \\
=&\beta Q_{i+1,i-4} \sum_{i'} R_{i-4,i'} z_{i',k+1}-\beta(1-Q_{i+1,i+1}) \sum_{i'} R_{i+1,i'} z_{i',k+1}-W_{i,1,k} \\
=&\beta Q_{i+1,i-4} \sum_{i'} R_{i-4,i'} z_{i',k+1}-\beta Q_{i+1,i-4} \sum_{i'} R_{i+1,i'} z_{i',k+1}-W_{i,1,k} \\
\geq&\beta Q_{i,i-5} \sum_{i'} R_{i-4,i'} z_{i',k+1}-\beta Q_{i,i-5} \sum_{i'} R_{i+1,i'} z_{i',k+1}-W_{i,1,k} \\
\geq&\beta Q_{i,i-5} \sum_{i'} R_{i-5,i'} z_{i',k+1}-\beta Q_{i,i-5} \sum_{i'} R_{i,i'} z_{i',k+1}-W_{i,1,k} \\
=&r_{i,k}+\beta \sum_{i'}(QR)_{i,i'} z_{i',k+1}-r_{i,k}-\beta \sum_{i'} R_{i,i'} z_{i',k+1}-W_{i,1,k} \\
=&0,
\end{aligned}
$$

where the first inequality follows from $\sum_{i'} R_{i-4,i'} z_{i',k+1} \geq \sum_{i'} R_{i+1,i'} z_{i',k+1}$ and $Q_{i+1,i-4} \geq Q_{i,i-5}$ (due to Condition 1), the second inequality follows from Inequality 5, and the last equation follows from the definition of Whittle's index at state $(i,1,k)$.

For $j \geq 2$, when passive subsidy is equal to $W_{i,j,k}$, the difference between the rewards of selecting active and passive actions for state $(i+1,j,k)$ is:

$$
\begin{aligned}
& r_{i+1,k}+\beta \sum_{i'}(QR)_{i+1,i'} V_{i',j-1,k+1}(W_{i,j,k})-r_{i+1,k}-\beta \sum_{i'} R_{i+1,i'} V_{i',j-1,k+1}(W_{i,j,k})-W_{i,j,k} \\
=&\beta \sum_{i'}(QR)_{i+1,i'} V_{i',j-1,k+1}(W_{i,j,k})-\beta \sum_{i'} R_{i+1,i'} V_{i',j-1,k+1}(W_{i,j,k})-W_{i,j,k} \\
=&\beta Q_{i+1,i-4} \sum_{i'} R_{i-4,i'} V_{i',j-1,k+1}(W_{i,j,k})-\beta(1-Q_{i+1,i+1}) \sum_{i'} R_{i+1,i'} V_{i',j-1,k+1}(W_{i,j,k})-W_{i,j,k} \\
=&\beta Q_{i+1,i-4} \sum_{i'} R_{i-4,i'} V_{i',j-1,k+1}(W_{i,j,k})-\beta Q_{i+1,i-4} \sum_{i'} R_{i+1,i'} V_{i',j-1,k+1}(W_{i,j,k})-W_{i,j,k} \\
\geq&\beta Q_{i,i-5} \sum_{i'} R_{i-4,i'} V_{i',j-1,k+1}(W_{i,j,k})-\beta Q_{i,i-5} \sum_{i'} R_{i+1,i'} V_{i',j-1,k+1}(W_{i,j,k})-W_{i,j,k},
\end{aligned}
$$

where the inequality follows from $\sum_{i'} R_{i-4,i'} V_{i',j-1,k+1}(W_{i,j,k}) \geq \sum_{i'} R_{i+1,i'} V_{i',j-1,k+1}(W_{i,j,k})$ and $Q_{i+1,i-4} \geq Q_{i,i-5}$ (due to Condition 1). Since we assume no reinfection within prison, a patient at an SVR state will never transition to a candidate state. Therefore, we have $\sum_{i'} R_{i-4,i'} V_{i',j-1,k+1}(W_{i,j,k})=\mathbf{e}_{i-4}^T\left(\sum_{j'=1}^{j-1} \beta^{j'-1} R^{j'} r_{:,k+j'}+\beta^{j-1} R^j z_{:,k+j}\right)$. Further, since $R_i \preceq R_{i+1}$, we have $\sum_{i'} R_{i,i'} V_{i',j-1,k+1}(W_{i,j,k})-\sum_{i'} R_{i+1,i'} V_{i',j-1,k+1}(W_{i,j,k}) \geq (\mathbf{e}_i^T-\mathbf{e}_{i+1}^T) R r_{:,k+1}$. Then:

$$
\beta Q_{i,i-5} \sum_{i'} R_{i-4,i'} V_{i',j-1,k+1}(W_{i,j,k})-\beta Q_{i,i-5} \sum_{i'} R_{i+1,i'} V_{i',j-1,k+1}(W_{i,j,k})-W_{i,j,k}
$$

$$
\begin{aligned}
&\geq \beta Q_{i,i-5} \sum_{i'} R_{i-5,i'} V_{i',j-1,k+1}(W_{i,j,k}) - \beta Q_{i,i-5} \sum_{i'} R_{i,i'} V_{i',j-1,k+1}(W_{i,j,k}) - W_{i,j,k} \\
&= r_{i,k} + \beta \sum_{i'} (QR)_{i,i'} V_{i',j-1,k+1}(W_{i,j,k}) - r_{i,k} - \beta \sum_{i'} R_{i,i'} V_{i',j-1,k+1}(W_{i,j,k}) - W_{i,j,k} \\
&= 0,
\end{aligned}
$$

where the inequality follows from Inequality 5, and the last equation follows from the definition of Whittle's index at state (i, j, k), which completes the proof. □

Proof of Proposition 2. We prove this proposition by induction on j. We first show that for $i = 11, W_{i,2,k} \geq W_{i,1,k}, \forall k$. To do so, it is sufficient to show that when the subsidy level is equal to $W_{i,1,k}$, the optimal action for state $(i, 2, k)$ is active. When the passive subsidy is equal to $W_{i,1,k}$, the difference between the rewards of selecting active and passive actions for state $(i, 2, k)$ is:

$$
\begin{aligned}
&r_{i,k} + \beta \sum_{i'} (QR)_{i,i'} V_{i',1,k+1}(W_{i,1,k}) - r_{i,k} - \beta \sum_{i'} R_{i,i'} V_{i',1,k+1}(W_{i,1,k}) - W_{i,1,k} \\
=&\beta \sum_{i'} (QR)_{i,i'} V_{i',1,k+1}(W_{i,1,k}) - \beta \sum_{i'} R_{i,i'} V_{i',1,k+1}(W_{i,1,k}) - W_{i,1,k} \\
=&\beta Q_{i,i-5} \sum_{i'} R_{i-5,i'} V_{i',1,k+1}(W_{i,1,k}) - \beta(1 - Q_{i,i}) \sum_{i'} R_{i,i'} V_{i',1,k+1}(W_{i,1,k}) - W_{i,1,k} \\
=&\beta Q_{i,i-5} \sum_{i'} R_{i-5,i'} V_{i',1,k+1}(W_{i,1,k}) - \beta Q_{i,i-5} \sum_{i'} R_{i,i'} V_{i',1,k+1}(W_{i,1,k}) - W_{i,1,k}.
\end{aligned}
$$

When the passive subsidy is $W_{i,1,k}$, the optimal action for state $(i, 1, k+1)$ is passive (as we will show in Proposition 3). Therefore, we have:

$$
\begin{aligned}
&\beta Q_{i,i-5} \sum_{i'} R_{i-5,i'} V_{i',1,k+1}(W_{i,1,k}) - \beta Q_{i,i-5} \sum_{i'} R_{i,i'} V_{i',1,k+1}(W_{i,1,k}) - W_{i,1,k} \\
=&Q_{i,i-5}(\mathbf{e}_{i-5}^T - \mathbf{e}_i^T)(\beta R r_{:,k+1} + \beta^2 R^2 z_{:,k+2}) - W_{i,1,k} \\
\geq&Q_{i,i-5}(\mathbf{e}_{i-5}^T - \mathbf{e}_i^T)\beta R z_{:,k+1} - W_{i,1,k} \\
=&\beta Q_{i,i-5} \sum_{i'} R_{i-5,i'} z_{i',k+1} - \beta Q_{i,i-5} \sum_{i'} R_{i,i'} z_{i',k+1} - W_{i,1,k} \\
=&r_{i,k} + \beta \sum_{i'} (QR)_{i,i'} z_{i',k+1} - r_{i,k} - \beta \sum_{i'} R_{i,i'} z_{i',k+1} - W_{i,1,k} \\
=&0,
\end{aligned}
$$

where the inequality follows from Inequality 6, and the last equation follows from the definition of the Whittle's index at state $(i, 1, k)$.

For $j \geq 2$, assume that we have $W_{i,j,k} \geq W_{i,j-1,k}$, we now show that $W_{i,j+1,k} \geq W_{i,j,k}$. To do so, it is sufficient to show that when the subsidy level is equal to $W_{i,j,k}$, the optimal action for state $(i, j+1, k)$ is active. When the passive subsidy is equal to $W_{i,j,k}$, the difference between the rewards of selecting active and passive actions for state $(i, j+1, k)$ is:

$$
\begin{aligned}
&r_{i,k} + \beta \sum_{i'} (QR)_{i,i'} V_{i',j,k+1}(W_{i,j,k}) - r_{i,k} - \beta \sum_{i'} R_{i,i'} V_{i',j,k+1}(W_{i,j,k}) - W_{i,j,k} \\
=&\beta \sum_{i'} (QR)_{i,i'} V_{i',j,k+1}(W_{i,j,k}) - \beta \sum_{i'} R_{i,i'} V_{i',j,k+1}(W_{i,j,k}) - W_{i,j,k} \\
=&\beta Q_{i,i-5} \sum_{i'} R_{i-5,i'} V_{i',j,k+1}(W_{i,j,k}) - (1 - \beta Q_{i,i}) \sum_{i'} R_{i,i'} V_{i',j,k+1}(W_{i,j,k}) - W_{i,j,k} \\
=&\beta Q_{i,i-5} \sum_{i'} R_{i-5,i'} V_{i',j,k+1}(W_{i,j,k}) - \beta Q_{i,i-5} \sum_{i'} R_{i,i'} V_{i',j,k+1}(W_{i,j,k}) - W_{i,j,k}.
\end{aligned}
$$

When the passive subsidy is $W_{i,j,k}$, the optimal action for states $(i,j,k+1)$ and $(i,j-1,k+1)$ are both passive (due to Proposition 3 and induction assumption). Therefore, we have:

$$\begin{aligned}
&\beta Q_{i,i-5}\sum_{i'}R_{i-5,i'}V_{i',j,k+1}(W_{i,j,k})-\beta Q_{i,i-5}\sum_{i'}R_{i,i'}V_{i',j,k+1}(W_{i,j,k})-W_{i,j,k}\\
=&Q_{i,i-5}(\mathbf{e}_{i-5}^T-\mathbf{e}_i^T)\left(\sum_{j'=1}^{j}\beta^{j'}R^{j'}r_{:,k+j'}+\beta^{j+1}R^{j+1}z_{:,k+j+1}\right)-W_{i,j,k}\\
\geq&Q_{i,i-5}(\mathbf{e}_{i-5}^T-\mathbf{e}_i^T)\left(\sum_{j'=1}^{j-1}\beta^{j'}R^{j'}r_{:,k+j'}+\beta^{j}R^{j}z_{:,k+j}\right)-W_{i,j,k}\\
=&\beta Q_{i,i-5}\sum_{i'}R_{i-5,i'}V_{i',j-1,k+1}(W_{i,j,k})-\beta Q_{i,i-5}\sum_{i'}R_{i,i'}V_{i',j-1,k+1}(W_{i,j,k})-W_{i,j,k}\\
=&r_{i,k}+\beta\sum_{i'}(QR)_{i,i'}V_{i',j-1,k+1}(W_{i,j,k})-r_{i,k}-\beta\sum_{i'}R_{i,i'}V_{i',j-1,k+1}(W_{i,j,k})-W_{i,j,k}\\
=&0,
\end{aligned}$$

where the inequality follows from Inequality 6, and the last equation follows from the definition of the Whittle's index at state $(i,1,k)$, which completes the proof. □

Proof of Proposition 3. Clearly, it is sufficient to consider candidate states $i=7,...,11$. For such a state (i,j,k), to show $W_{i,j,k+1}\leq W_{i,j,k}$, it is sufficient to show that when the subsidy level is equal to $W_{i,j,k}$, the optimal action for state $(i,j,k+1)$ is passive. Similar as before, let $V_{i,j,k}(W)$ be the optimal value function at state (i,j,k) when the passive subsidy is W. Then, we have the following result.

LEMMA 4. *For any given passive subsidy W, for $i=2,...,6$ and $i'=i+5,...,14$, $V_{i,j,k}(W)-V_{i',j,k}(W)$ is non-increasing in $k,\forall j$.*

With the above result, we now continue the proof as follows. For $j=1$, when passive subsidy is equal to $W_{i,1,k}$, the difference between the rewards of selecting active and passive actions for state $(i,1,k+1)$ is:

$$\begin{aligned}
&r_{i,1,k+1}^a-r_{i,1,k+1}^b-W_{i,1,k}\\
=&\beta\sum_{i'}(QR)_{i,i'}z_{i',k+2}-\beta\sum_{i'}R_{i,i'}z_{i',k+2}-W_{i,1,k}\\
=&\beta Q_{i,i-5}\sum_{i'}R_{i-5,i'}z_{i',k+2}-\beta(1-Q_{i,i})\sum_{i'}R_{i,i'}z_{i',k+2}-W_{i,1,k}\\
=&\beta Q_{i,i-5}\sum_{i'}R_{i-5,i'}z_{i',k+2}-\beta Q_{i,i-5}\sum_{i'}R_{i,i'}z_{i',k+2}-W_{i,1,k}\\
\leq&\beta Q_{i,i-5}\sum_{i'}R_{i-5,i'}z_{i',k+1}-\beta Q_{i,i-5}\sum_{i'}R_{i,i'}z_{i',k+1}-W_{i,1,k}\\
=&r_{i,1,k}^a-r_{i,1,k}^b-W_{i,1,k}\\
=&0,
\end{aligned}$$

where the inequality follows from $z_{i,k+2}-z_{i',k+2}\leq z_{i,k+1}-z_{i',k+1}$ for $i=2,...,6,i'=i+5,...,14$, and the last equation follows from the definition of the Whittle's index at state $(i,1,k)$.

For $j\geq 2$, when the passive subsidy is equal to $W_{i,j,k}$, the difference between the rewards of selecting active and passive actions for state $(i,j,k+1)$ is:

$$r_{i,k+1}+\beta\sum_{i'}(QR)_{i,i'}V_{i',j-1,k+2}(W_{i,j,k})-r_{i,k+1}-\beta\sum_{i'}R_{i,i'}V_{i',j-1,k+2}(W_{i,j,k})-W_{i,j,k}$$

$$
\begin{aligned}
=&\beta\sum_{i'}(QR)_{i,i'}V_{i',j-1,k+2}(W_{i,j,k})-\beta\sum_{i'}R_{i,i'}V_{i',j-1,k+2}(W_{i,j,k})-W_{i,j,k}\\
=&\beta Q_{i,i-5}\sum_{i'}R_{i-5,i'}V_{i',j-1,k+2}(W_{i,j,k})-(1-\beta Q_{i,i})\sum_{i'}R_{i,i'}V_{i',j-1,k+2}(W_{i,j,k})-W_{i,j,k}\\
=&\beta Q_{i,i-5}\sum_{i'}R_{i-5,i'}V_{i',j-1,k+2}(W_{i,j,k})-\beta Q_{i,i-5}\sum_{i'}R_{i,i'}V_{i',j-1,k+2}(W_{i,j,k})-W_{i,j,k}\\
\leq&\beta Q_{i,i-5}\sum_{i'}R_{i-5,i'}V_{i',j-1,k+1}(W_{i,j,k})-\beta Q_{i,i-5}\sum_{i'}R_{i,i'}V_{i',j-1,k+1}(W_{i,j,k})-W_{i,j,k}\\
=&r_{i,k}+\beta\sum_{i'}(QR)_{i,i'}V_{i',j-1,k+1}(W_{i,j,k})-r_{i,k}-\beta\sum_{i'}R_{i,i'}V_{i',j-1,k+1}(W_{i,j,k})-W_{i,j,k}\\
=&0,
\end{aligned}
$$

where the inequality follows from Lemma 4, and the last equation follows from the definition of the Whittle's index at state (i,j,k), which completes the proof. □

Proof of Lemma 4. We prove the lemma by induction on j. Since for $i=7,...,11$, $r_{i-5,k}-r_{i,k}$ and $z_{i-5,k}-z_{i,k}$ are both non-increasing in k, and for $i=7,...,13$, $r_{i,k}-r_{i+1,k}$ and $z_{i,k}-z_{i+1,k}$ are both non-increasing in k, it is not difficult to check that the result holds for $j=1$. Assume that the result holds for some $j-1$, we now show that it also holds for j. For $i=2,...,6$ and $i'=i+5,...,14$, consider the following two cases:

First, suppose the optimal action at state (i',j,k) is active for given passive subsidy W. Then, we have:

$$
\begin{aligned}
&V_{i,j,k+1}(W)-V_{i',j,k+1}(W)-(V_{i,j,k}(W)-V_{i',j,k}(W))\\
\leq&r_{i,k+1}+W+\sum_{i''}R_{i,i''}V_{i'',j-1,k+2}(W)-r_{i',k+1}-\sum_{i''}(QR)_{i',i''}V_{i'',j-1,k+2}(W)\\
&-r_{i,k}-W-\sum_{i''}R_{i,i''}V_{i'',j-1,k+1}(W)+r_{i',k}+\sum_{i''}(QR)_{i',i''}V_{i'',j-1,k+1}(W)\\
\leq&\sum_{i''}R_{i,i''}V_{i'',j-1,k+2}(W)-\sum_{i''}(QR)_{i',i''}V_{i'',j-1,k+2}(W)\\
&-\sum_{i''}R_{i,i''}V_{i'',j-1,k+1}(W)+\sum_{i''}(QR)_{i',i''}V_{i'',j-1,k+1}(W)\\
\leq&0,
\end{aligned}
$$

where first inequality follows from the fact that $V_{i',j,k+1}(W)\geq r_{i',k+1}+\sum_{i''}(QR)_{i',i''}V_{i'',j-1,k+2}(W)$, the second inequality follows from $r_{i,k+1}-r_{i',k+1}\leq r_{i,k}-r_{i',k}$, and the last inequality follows from induction assumption.

Second, suppose the optimal action at state (i',j,k) is passive for given passive subsidy W. Then, we have:

$$
\begin{aligned}
&V_{i,j,k+1}(W)-V_{i',j,k+1}(W)-(V_{i,j,k}(W)-V_{i',j,k}(W))\\
\leq&r_{i,k+1}+W+\sum_{i''}R_{i,i''}V_{i'',j-1,k+2}(W)-r_{i',k+1}-W-\sum_{i''}R_{i',i''}V_{i'',j-1,k+2}(W)\\
&-r_{i,k}-W-\sum_{i''}R_{i,i''}V_{i'',j-1,k+1}(W)+r_{i',k}+W+\sum_{i''}R_{i',i''}V_{i'',j-1,k+1}(W)\\
\leq&\sum_{i''}R_{i,i''}V_{i'',j-1,k+2}(W)-\sum_{i''}R_{i',i''}V_{i'',j-1,k+2}(W)\\
&-\sum_{i''}R_{i,i''}V_{i'',j-1,k+1}(W)+\sum_{i''}R_{i',i''}V_{i'',j-1,k+1}(W)\\
\leq&0,
\end{aligned}
$$

where first inequality follows from the fact that $V_{i',j,k+1}(W) \geq r_{i',k+1} + W + \sum_{i''} R_{i',i''} V_{i'',j-1,k+2}(W)$, the second inequality follows from $r_{i,k+1} - r_{i',k+1} \leq r_{i,k} - r_{i',k}$, and the last inequality follows from induction assumption, which completes the proof. $\square$

Appendix. B. Parameter Estimation for Index Policies

In this section, we describe the parameter estimation for the computation of index policies. We solve our problem for Genotype 1 male patients, who constitute the vast majority of the U.S. prisoners (Guerino et al. 2011, Nainan et al. 2006). To compute index policies, we first need to estimate the parameters for each bandit/patient. Below, we provide an overview of parameter estimations, and a summary of the parameter values and their sources are included in Table 5.

Transition Probabilities: The transitions of remaining sentence lengths and ages are deterministic and trivial, therefore we describe transitions among different liver health states, which is jointly determined by i) the disease progression matrix R, and ii) the treatment matrix Q. Similar to He et al. (2015), we estimate the disease progression parameters and treatment SVR rates from published studies (Thein et al. 2008, Lok et al. 2009, Fattovich et al. 1997, Planas et al. 2004, Afdhal et al. 2014). We adjust the natural progression matrix R for different ages by considering different background mortality rates, which are estimated from the U.S. life table (Arias 2011).

Rewards: To estimate the rewards for each bandit/patient, it suffices to estimate i) the pre-release QALYs for different health states and ages, and ii) the post-release lumpsum rewards for different health states and ages. The health and age related QALYs are estimated from published studies (Chong et al. 2003, Hanmer et al. 2006). Therefore, the key parameters to estimate are the lumpsum rewards. As noted earlier, we estimate the lumpsum reward for each patient by two components: the cumulative QALYs from the released patient upon release, and the potential reduction in overall QALYs in the population due to possible spread of the disease to the society. First, to estimate the cumulative QALYs from a released patient, it is sufficient to estimate the treatment rate (i.e., how likely he will be treated per year outside prison if he was released untreated but aware of his disease status) and the reinfection rate (i.e., how likely he will be reinfected per year if he was released treated). Based on existing studies, we assume a treatment rate of 0.10 for all released prisoners (Wedemeyer et al. 2015). We assume a zero reinfection rate for non-IDUs, and we estimate the reinfection rate for IDUs from a published study (Currie et al. 2008). Second, we estimate the potential reduction in overall QALYs in the population by the expected number of new infections caused by the released prisoner times the total discounted lifelong loss of QALYs for HCV infection. To estimate the expected number of new infections caused by the released prisoner, it is sufficient to know the infection rate (i.e., how likely the released prisoner will infect others per year if he was released untreated). As there is limited data for HCV transmission, we estimate the infection rates for non-IDUs and IDUs by calibration using the simulation model described in §6. Finally, we estimate the total discounted lifelong loss of QALYs for infection by simulation as the difference of the cumulative QALYs of an average-age uninfected person and that of an average-age F0 patient.

Table 5: Model Parameter Values for the Computation of Index Policies.

Variables	Values	Reference

Disease progression parameters		
F0 to F1	0.117	Thein et al. (2008)
F1 to F2	0.085	Thein et al. (2008)
F2 to F3	0.120	Thein et al. (2008)
F3 to F4	0.116	Thein et al. (2008)
F3 to HCC	0.008	Lok et al. (2009)
F4 to DC	0.039	Fattovich et al. (1997)
F4 to HCC	0.014	Fattovich et al. (1997)
DC to HCC	0.068	Planas et al. (2004)
DC to liver-related death	0.182	Planas et al. (2004)
HCC to liver-related death	0.427	Fattovich et al. (1997)
Background mortality rates	Life table	Arias (2011)
Treatment SVR rates by liver health state		
F0-F4	0.970	Afdhal et al. (2014)
Quality of life		
Health-related quality-of-life weights		
F0-F3	0.930	Chong et al. (2003)
F4	0.900	Chong et al. (2003)
DC	0.800	Chong et al. (2003)
HCC	0.790	Chong et al. (2003)
F0SVR-F4SVR	1.000	Assumption
Age-related quality-of-life weights		Hanmer et al. (2006)
0-29	0.928	
30-39	0.918	
40-49	0.887	
50-59	0.861	
60-69	0.840	
70-79	0.802	
≥80	0.782	
Parameters for estimation of lumpsum rewards		
Treatment rate outside prison for released prisoners	0.010	
Reinfection rate outside prison for non-IDUs	0.000	Assumption
Reinfection rate outside prison for IUDs	0.018	Currie et al. (2008)
Infection rate outside prison for non-IDUs	0.005	Calibration
Infection rate outside prison for IDUs	0.043	Calibration
Discounted lifelong loss of QALYs for infection of HCV	1.320	Simulation

Appendix. C. Parameters for the Simulation Model

Table 6: Baseline Demographics.

Model Parameters	Values	Reference
Population		
Total population	200,000	
Prison population	1,000	
Proportion of inmates	0.5%	Guerino et al. (2011)
Gender (Male)		
In prisons	91%	Guerino et al. (2011)
Outside prisons	52%	Howden and Meyer (2010)
Prevalence of IDUs		
In prisons	26.0%	Mumola et al. (2006)
Outside prisons	1.2%	Brady et al. (2008)
Birth rate		
Number of newborns per 1,000 population per year	14.3	World Bank (2014)

Table 7: Baseline Age Distribution.

Age Category	General Population (Howden and Meyer 2010)		Prison Population (Georgia Department of Corrections 2014)	
	Male (%)	Female (%)	Male (%)	Female (%)
0-4	7.2	6.7	–	–
5-9	7.1	6.6	–	–
10-14	6.8	6.3	–	–
15-19	7.1	6.6	1.5	0.8
20-24	7.3	6.8	15.1	14.3
25-29	7.2	6.7	15.1	14.3
30-34	6.7	6.5	14.8	16.8
35-39	6.3	6.2	14.8	16.8
40-44	6.8	6.7	10.3	11.1
45-49	7.2	7.2	10.3	11.1
50-54	7.1	7.2	6.6	6.0
55-59	6.3	6.5	6.6	6.0
60-64	5.6	5.8	2.0	1.2
65-69	3.7	4.2	2.0	1.2
$\geq$70	–	–	1.1	0.3
70-74	2.8	3.2	–	–
75-79	2.0	2.6	–	–
80-84	1.6	2.2	–	–
$\geq$85	1.2	2.0	–	–

Table 8: Baseline Hepatitis C Distribution.

Model Parameters	Values		Reference
HCV prevalence			
HCV prevalence by age and gender	Male	Female	Denniston et al. (2014)
0-5	0.0093%	0.0093%	
6-20	0.0498%	0.0498%	
20-29	0.1231%	0.0704%	
30-39	1.0523%	0.6023%	
40-49	3.9494%	2.2604%	
50-59	4.3334%	2.4801%	
≥60	0.8069%	0.4618%	
Hazard ratio of HCV among inmates	3.5		Calibration
Hazard ratio of HCV among IDUs	20.0		Calibration
HCV prevalence among new borns	0.0093%		Cottrell et al. (2013)
Chronic hepatitis C disease stage			Davis et al. (2010)
F0	13.7%		
F1	24.6%		
F2	18.7%		
F3	16.7%		
F4	22.9%		
DC	3.1%		
HCC	0.3%		
Disease awareness in community			
Proportion of patients aware of HCV infection	50.0%		Denniston et al. (2012)

Table 9: HCV Disease Progression Probabilities.

Model parameters	Values	Reference
Disease progression parameters		
F0 to F1	0.117	Thein et al. (2008)
F1 to F2	0.085	Thein et al. (2008)
F2 to F3	0.120	Thein et al. (2008)
F3 to F4	0.116	Thein et al. (2008)
F3 to HCC	0.008	Lok et al. (2009)
F4 to DC	0.039	Fattovich et al. (1997)
F4 to HCC	0.014	Fattovich et al. (1997)
F4SVR to DC	0.008	Cardoso et al. (2010)
F4SVR to HCC	0.005	Cardoso et al. (2010)
DC to HCC	0.068	Planas et al. (2004)
DC (first year) to liver transplant in community	0.023	Thuluvath et al. (2010)

DC (subsequent years) to liver-related death	0.112	Thuluvath et al. (2010)
DC to liver-related death	0.182	Planas et al. (2004)
HCC to liver transplant in community	0.040	Lang et al. (2009)
HCC to liver-related death	0.427	Fattovich et al. (1997)
Liver transplant (first year) to liver-related death	0.116	Wolfe et al. (2010)
Liver transplant (subsequent years) to liver-related death	0.044	Wolfe et al. (2010)

Table 10: Transmission-Related Parameters.

Model Parameters	Values	Reference
Transmission related parameters		
Baseline transmission probability	0.006	Calibration
IDU-IDU interaction probability	0.988	Calibration
Hazard ratio of IDU-IDU interaction	10	Calibration
Awareness reduction factor	0.5	Calibration

Table 11: Probability of Diagnosis in Community.

Fibrosis stages	Probabilities of diagnosis Kabiri et al. (2014)
F0	0.037
F1	0.030
F2	0.042
F3	0.046
F4	0.163

Table 12: Probability of Incarceration.

Age at release	Probability by years after release (Langan and Levin 2002)			
	1 year	2 years	3 years	>3 years or never
18-24	0.062	0.049	0.026	0.003
25-29	0.055	0.044	0.023	0.010
30-34	0.054	0.042	0.022	0.012
35-39	0.050	0.039	0.021	0.012
40-44	0.038	0.0307	0.016	0.011
≥45	0.035	0.028	0.015	0.010

Table 13: Length of Sentence.

Length of sentence	Probability Georgia Department of Corrections (2014)
<1 year	0.245
1-1.99 years	0.229
2-2.99 years	0.164
3-3.99 years	0.104
4-4.99 years	0.067
5-5.99 years	0.052
6-6.99 years	0.035
7-7.99 years	0.024
8-8.99 years	0.013
9-9.99 years	0.012
10-10.99 years	0.018
≥11 years	0.038

Table 14: Health and Age Related Quality-Adjusted-Life-Years (QALYs)

Model Parameters	Values		Reference
Health-related quality-of-life weights			
F0-F3	0.93		Chong et al. (2003)
F4	0.90		Chong et al. (2003)
DC	0.80		Chong et al. (2003)
HCC	0.79		Chong et al. (2003)
Liver transplant	0.84		Chong et al. (2003)
F0SVR-F4SVR	1.00		Assumption
Age-related quality-of-life weights	Male	Female	Hanmer et al. (2006)
0-29	0.928	0.913	
30-39	0.918	0.893	
40-49	0.887	0.863	
50-59	0.861	0.837	
60-69	0.840	0.811	
70-79	0.802	0.771	
≥80	0.782	0.724	

Appendix. D. Sensitivity Analysis against Key Model Parameters

In this section, we test the robustness of our proposed index policy against several key model parameters. In particular, we consider the variations of the treatment rate for released prisoners, the infection rate outside prison, reinfection rate for IDUs, and simulation time horizon. We consider a 20% deviation from the base scenario for each of these parameters. Finally, recall that we have assumed no transmission (i.e. zero infection rate) within prison system in the analytical model. In our simulation model, we also relax this assumption and consider the same infection rate as in the community, which is expected to be an upper bound for the infection rate inside the prison system. The ranges of QALY gains of our proposed policy (i.e., the improvement of total QALYs compared with no treatment) are presented in Figure 6.

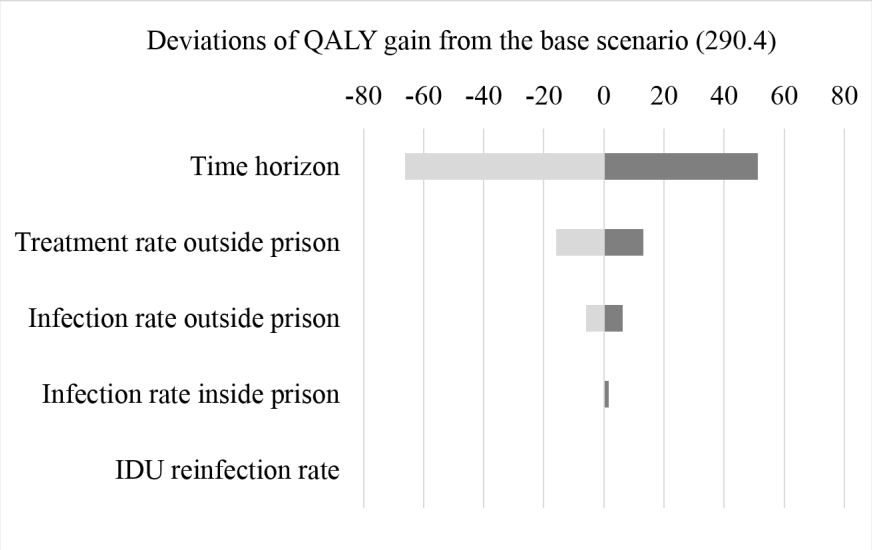


Figure 6 **Deviations of QALY gain of our policy with a variation of 20% for each parameter when treatment capacity within prison is $M = 10$ (in base scenario, QALY gain of our policy is 290.4).**

Note. While we present the results only for treatment capacity $M = 10$, the results are qualitatively similar for other values of M.

From these results, we first observe that time horizon is the most sensitive parameter: When the simulation time horizon ranges from 24 to 36 years (a 20% deviation from 30 years), the QALY gain of our policy falls into a large interval ranging between 224.3 and 341.6 (a deviation between -66.1 and 51.2 from the base scenario). The QALY gain being most sensitive to time horizon is expected, which also demonstrates that the benefit of treatment will accumulate in the long run. On the other hand, the QALY gain of our policy does not vary much when the infection rate inside prison or IDU reinfection rate change.

Further, we remark that since the QALY gains of different policies change in similar directions with respect of variations of parameters (e.g., when the simulation time horizon decreases, the QALY gains of all policies decrease), the percentage improvement of our policy over the benchmark health state policy is very robust. In particular, in all the tested cases, our policy always performs better than the health status policy. Recall

that in the base scenario when $M = 10$, our policy has a 2.9% improvement over the health status policy; in the sensitivity analysis, the worst case occurs when the simulation time horizon becomes short (24 years), which corresponds to a 2.4% improvement over the health status policy. In sum, the performance of our proposed policy is robust and performs better than the benchmark policies in a wide range of parameters.