Optimal Coinsurance Rates for a Heterogeneous Population under Inequality and Resource Constraints

Greggory J. Schell[1], Gian-Gabriel P. Garcia[2], Mariel S. Lavieri[2], Jeremy B. Sussman[3], Rodney A. Hayward[3]

[1] *Center for Naval Analyses, 3003 Washington Blvd, Arlington, VA 22201, USA*

[2] *Industrial and Operations Engineering, University of Michigan, 1205 Beal Ave, Ann Arbor, MI 48109, USA*

[3] *Center for Clinical Management Research, Ann Arbor Veterans Affairs Hospital, 2215 Fuller Road, Ann Arbor, MI 48105, USA*

schellg@cna.org, garciagg@umich.edu, lavieri@umich.edu, jeremysu@med.umich.edu, rhayward@med.umich.edu

The Version of Record of this manuscript has been published and is available in *IISE Transactions* February 6, 2019 https://www.tandfonline.com/doi/abs/10.1080/24725854.2018.1499053

Abstract: While operations research has contributed heavily to the derivation of optimal treatment guidelines for chronic diseases, patient adherence to treatment plans is low and variable. One mechanism for improving patient adherence to guidelines is to tailor coinsurance rates for prescription medications to patient characteristics. We seek to find coinsurance rates which maximize the welfare of the heterogeneous patient population at risk for cardiovascular disease. We analyze the problem as a bilevel optimization model where the lower optimization problem has the structure of a Markov decision process which determines the optimal treatment plan for each patient class. The upper optimization problem is a nonlinear resource allocation problem with constraints on total expenditures and coinsurance inequality. We used dynamic programming with a penalty function for nonseparable constraint violations to derive the optimal coinsurance rates. We parameterized and solved this model by considering patients who are insured by Medicare and are prescribed medications for prevention of cardiovascular disease. We find that optimizing coinsurance rates can be a cost-effective intervention for improving patient adherence and health outcomes, particularly for those patients at high risk for cardiovascular disease.

Keywords: Healthcare, Medical decision making, Cardiovascular disease, Resource allocation, Value-based insurance

1 Introduction

Operations research has contributed heavily to the derivation of optimal treatment guidelines for chronic diseases (Denton et al. 2011). These models provide state-based decision rules for how patients with particular characteristics should be treated over time. While the guidelines have been optimized for individual patients, the performance of these models in practice depends on the patient's adherence to the prescribed treatments. Unfortunately, adherence to treatment in a real-world setting is low and variable (Tamblyn et al. 2010). The reasons for low adherence to treatment include: forgetfullness, side-effects, and cost (Egan et al. 2003). Research in health economics suggests that, as the price of medication decreases, adherence to the medication will increase due to price elasticity. Therefore, insurance providers may be able to directly address low patient adherence by decreasing coinsurance rates and/or copayments (out-of-pocket expenses for patients), while taking into consideration optimal treatment plans prescribed to the patient. Health researchers (Chernew et al. 2010) and insurance providers (Mann et al. 2014) have championed the principle of reducing out-of-pocket expenses for highly valuable clinical services to improve adherence. This concept, coined Value-Based Insurance Design (V-BID), has also gained support in health policy.

In late 2016, Congress signed a defense spending bill which includes a pilot for V-BID within TRICARE, a healthcare program for U.S. service members and their families. The major aims of this pilot program are to evaluate how V-BID can improve medication adherence, quality measures, patient health outcomes, and beneficiary experience (McCain 2016). However, how to best design the out-of-pocket expense reductions for highly valuable clinical services remains an open question.

One such highly valuable clinical service is the prescription of medication for the primary and secondary prevention of cardiovascular disease (CVD), including myocardial infarction (MI) and stroke. Such medications include anti-hypertension and cholesterol-lowering medications, which prevent CVD by treating hypertension and hypercholesterolemia, respectively. The preventative benefits of these medications has generated theoretical (Choudhry et al.

2012) and real-world (Goldman et al. 2006, Gibson et al. 2006b, 2011) studies on the health and cost impact of coinsurance reductions. While the studies have shown the health improvements associated with reduced coinsurance rates, they have not addressed the issue of identifying optimal coinsurance rates.

Given the evidence for the benefit of reducing out-of-pocket expenses, we propose a mathematical model comprised of two optimization problems: (1) optimal preventative treatment planning for medication which maximizes patient health and (2) optimal coinsurance rates for medication which maximize the population's health. We formulate and parameterize these models first in the context of hypertension treatment and then later to cholesterol-lowering medications. The optimal preventative treatment plan is personalized to the CVD characteristics and price elasticity of each patient class (i.e. a subgroup of patients similar in their risk factors and adherence behavior) within the population. The heterogeneity of the risk factors and adherence behavior lead to differences in optimal treatment plans, marginal health benefit of improved adherence, and behavioral response to changes in coinsurance rates. Due to budget constraints, the insurance provider may wish to target coinsurance reductions to those patients who would receive the greatest health benefit. However, such a utilitarian approach may generate inequality in coinsurance rates across the different patient classes which may be politically or legally infeasible for a public health insurer, like the Department of Defense or Medicare. To address these issues, we determine optimal class-specific coinsurance rates, constrained by a budget and a measure of inequality.

Our research provides the following contributions: (1) we develop one of the first models to consider both optimal dynamic preventative treatment plans and strategic policy interventions for improving adherence to those plans, (2) we provide useful insights into the design of optimal coinsurance rates given the characteristics of the patient population and a constrained budget, (3) we analyze the role of coinsurance inequality in the cost-effectiveness of new coinsurance rate plans, and (4) we work closely with clinicians to parameterize our model to determine optimal hypertension and cholesterol-lowering medication coinsurance rates for

patients enrolled in Medicare. While this work is applied to hypertension and cholesterol-lowering medications within Medicare, the formulation, insights and modeling principles are applicable to other chronic diseases for which patients incur out-of-pocket expenses when treating their disease.

In Section 2, we review the relevant academic literature on payment structures, health economics and optimization. In Section 3, we present the mathematical model and discuss its properties. In Section 4, we discuss the penalty function dynamic programming solution methodology. Next, we analyze the numerical findings of our study in Section 5, which initially focuses on the application to hypertension medication and is later applied to cholesterol-lowering medications. Lastly, we discuss managerial insights in Section 6 and discuss study limitations and future research in Section 7.

2 Literature Review

The research domain of our work falls into the fields of medical decision making (which includes optimal treatment planning and resource allocation) and coinsurance rates (which models patient, provider, and payer relationships and the effect of medication cost on patient adherence).

2.1 Medical Decision Making

Denton et al. (2011) provide an overview of optimal treatment planning in healthcare, including Shechter et al. (2008), Zhang et al. (2012a,b), Ayer et al. (2012), Mason et al. (2014), Erenay et al. (2014), Zepeda and Sinha (2016), Skandari et al. (2015), Ibrahim et al. (2016), Ayer et al. (2016), Schell et al. (2016), Khojandi et al. (2018), and Kazemian et al. (2018). Similarly to our work, many of these previous research papers also aim to determine optimal treatment plans to manage a disease or condition. Specifically, Mason et al. (2012, 2014) and Schell et al. (2016) also aim to optimize the management of hypertension and cholesterol. In contrast, our model differs primarily from the existing literature in that our state-space directly includes the patient's adherence to medication so that the optimal preventative treatment plan and corresponding health measures depend on the medication

adherence level. Furthermore, we add complexity to our model by considering the effects of coinsurance changes to this adherence, and ultimately, patient outcomes. Since these coinsurance changes occur under a limited budget and fairness constraints, we also consider resource allocation models in healthcare.

In healthcare, the optimal allocation of limited resources has been considered in liver allocation (Akan et al. 2012), liver cancer (Lee et al. 2015), blood screening (El-Amine et al. 2018), mass-casualty triage (Mills et al. 2013), and diabetes prevention (Mehrotra and Kim 2011). Our problem differs from these previous works since we explicitly consider the effects of allocated resources on patient adherence. We also impose fairness constraints and explicitly consider the effects of allocated resources on patient adherence.

2.2 Determining Coinsurance Rates

Operations research and health economics have aimed to determine optimal coinsurance rates by studying the relationship between providers (e.g. hospital) and payers (e.g. insurance provider) (Sahney and Weddle 1980, Lee and Zenios 2012, Jiang et al. 2012, Yaesoubi and Roberts 2011, Denoyel et al. 2017), the relationship between manufacturers and insurance providers (Zaric et al. 2013), and the relationship between patients and physicians (Calcott 1999, De Jaegher and Jegers 2001, Schottmüller 2013). Our work bears similarities to the research by Jiang et al. (2012), Yaesoubi and Roberts (2011) and Schottmüller (2013) since the decisions made by the insurer and physician ultimately affect patient outcomes. However, all of these works assume that the physician provides services based on cost incentives rather than in an effort to maximize patient health outcomes. Furthermore, our research differs in that we model the relationship between patients and the insurance provider. To this end, Ma and Mcguire (1997) and Ho et al. (2016) have studied the relationship between insurers and patients. Ma and Mcguire (1997) investigate how the insurer can optimally determine contracts which incentivize honest claims reporting from physicians and patients whereas Ho et al. (2016) investigate how the insurer can incentivize patients to obtain influenza vaccines while simultaneously balancing societal benefit and expected cost. We differ from

these works in our objective function, which focuses on maximizing patient health outcomes. This patient-centered approach requires the modeling of disease dynamics over time and an optimization model for deriving dynamic preventative treatment plans. However, we also extend these works by imposing budget and fairness constraints. In particular, this latter constraint on coinsurance inequality differs from most coinsurance optimization research, which instead seeks a purely utilitarian solution.

An important component of our modeling framework is the change in adherence to medication based on changes in coinsurance rates. To this end, the field of health economics has also studied out-of-pocket expense reduction (Choudhry et al. 2008, Mann et al. 2014, Bunting et al. 2008). While the previous research has addressed the economics of out-of-pocket expenses, our work differs from this research in that we employ optimization to determine coinsurance rates which are tailored to each patient class' characteristics and the optimal preventative treatment plan.

To determine the optimal coinsurance rates, we formulate the problem as a bilevel optimization model with a single leader (Medicare) and multiple followers (the different patient classes). While hierarchical models have been applied to domains such as healthcare facilities (Zhang et al. 2010) and option pricing (Kovacevic and Pflug 2014), our model differs due to our constraint on conisurance inequality when determining optimal rates. The direct constraint on inequality allows for the computation of the health cost of coinsurance inequality, as well as a study of how the optimal coinsurance rates change as the inequality constraint is relaxed. We next describe our model formulation and analyze its structural properties.

3 Model

We consider a heterogeneous patient population, which we assume can be clustered into N different classes of patients and receives health insurance from a single provider. Let P denote this finite set of N patient classes. Each patient class $p \in P$ may be defined based on cardiovascular risk factors, e.g. age, sex, diabetes status, (which affect the optimal preventative treatment plan and marginal health benefit of improved adherence) and ad-

herence behavior (which includes behavioral responsiveness to changes in medication cost). Our joint model solves two problems: (1) the optimal preventative treatment policy for each patient class and (2) the optimal coinsurance rate for each patient class. The notation used throughout this section is summarized in Table 1.

Table 1: Notation

$p \in P$	A patient class p from the set of all patient classes P
π_p^*	Optimal MDP-based preventative treatment policy for patient class p
$t \in \{1, ..., T\}$	A decision epoch t from the set of decision epochs T in the MDP
$\mathbf{s}_p$	State corresponding to a patient in class p with risk-factors (e.g., age, sex, SBP), adherence to medication α_p, and health state h_p
$\tau \in \mathcal{A}$, $\mathcal{A}(b_{min}(\mathbf{s}_p))$	Medication type τ from the set of medications $\mathcal{A}$ and the set of medications which can be prescribed to a patient in class p based on a minimum allowable SBP, b_{min}
$d_{\tau,p}(\mathbf{s}_p)$, $q_p(\mathbf{s}_p)$	Disutility associated with prescribing medication τ and the quality of life associated with a patient being in health state $\mathbf{s}_p$, respectively
λ	Discount rate for the MDP
$p_\tau(\mathbf{s}_p, \mathbf{s}'_p)$	Transition probability associated with going from state $\mathbf{s}_p$ to $\mathbf{s}'_p$ after medication τ is prescribed
$V_{t,p}(\mathbf{s}_p)$, $f_{t,p}^\tau(\mathbf{s}_p)$	Optimal value function and state-action pair value function in decision epoch t
$M_p(\alpha_p), \mathbf{E}_p(\alpha_p), V_p(\alpha_p)$	Expected number of medications, number of CV events, and discounted QALYs, respectively, under optimal policy π_p^* for patient class p with adherence α_p
n_p	Expected number of patients in class p at the beginning of the planning horizon
$y_p \in [0, 1]$, $\mathbf{y}$, $\mathbf{y}^i$	Drug coverage for patient class p, vector of drug coverage decisions for each patient class, and current level of drug coverage for each patient class
$g(y_p, \mathbf{s}_p)$	Function describing the change in adherence, α_p, based on the medication coverage y_p and state information $\mathbf{s}_p$
c_τ, $C(\mathbf{y})$	Drug manufacturer's unit cost associated with medication τ and the total drug cost under coverage decision $\mathbf{y}$
B	Total budget for decreasing out-of-pocket costs for the patient population
$W(\mathbf{y})$, w	Measure of inequality based on coverage decisions $\mathbf{y}$ and the maximum allowable level of inequality

3.1 Optimal Treatment Policy

We first considered the optimal preventative treatment policy in the context of hypertension treatment planning. The formulation which follows can be generalized to address other treatment types (such as cholesterol, as we show in Section 5.5) and other chronic diseases. To determine the optimal hypertension treatment policy for a particular patient class, we developed a Markov decision process (MDP) model that is sensitive to the patient class' cardiovascular risk factors and adherence to medication (Schell 2015). The optimal preventative treatment policy π_p^* for patient class p is the type of medication prescribed at each decision epoch t in the planning horizon $t = 1, \ldots, T$ which maximizes the patient class's expected

total discounted QALYs. Let $\mathbf{s}_p$ denote the state of patient class p, which is comprised of the cardiovascular risk factors (e.g. age, diabetes status, SBP), adherence to medication α_p and health state h_p. We consider nine mutually exclusive and exhaustive health states: (1) no history of MI or stroke (healthy); (2) history of MI but no MI event this period; (3) history of stroke but no stroke this period; (4) history of MI and stroke but no adverse event this period; (5) survived a MI event this period; (6) survived a stroke this period; (7) death from a non-CVD related cause; (8) death from MI event this period; (9) death from stroke this period. These health states were developed in close collaboration with clinicians and are consistent with those used in previously published research on personalized hypertension treatment planning (Mason et al. 2012, 2014, Schell et al. 2016). The inclusion of adherence in the state space allows for the derivation of optimal preventative treatment policies which are tailored to the expected SBP reduction from treatment when the patient class is adherent to medication at level α_p. Lower adherence leads to lower SBP reduction from treatment. Hence, our MDP model personalizes the optimal preventative treatment policy based on the patient class's cardiovascular risk factors and adherence to medication.

Let $\mathcal{A}$ denote the set of possible medications which can be prescribed to patients in each patient class. In this framework, we assume that each different dosage level of the same medication will count as a different type of medication. Specifically, for hypertension treatment, $\mathcal{A}$ consists of different dosage levels, i.e., the number, of a generic anti-hypertensive medication. Taking medication $\tau \in \mathcal{A}$ is associated with a quality of life reduction, known as the medication disutility $d_{\tau,p}(\mathbf{s}_p)$. While $\mathcal{A}$ is the same for all patient classes, there are guideline-based restrictions which prohibit the prescription of certain medications. That is, the set of possible medications which can be prescribed in a decision epoch is denoted $\mathcal{A}(b_{min}(\mathbf{s}_p))$ and is constrained by a minimum allowable SBP, b_{min}. This minimum allowable SBP threshold prevents the MDP model from prescribing medications which would bring the SBP too low. While cardiovascular risk decreases as SBP decreases, there is an unsafe range of SBP which is associated with fainting and dizziness.

We also measure the well-being of the patient class given their health state using a quality of life weight $q_p(\mathbf{s}_p)$. For example, suffering a stroke ($h_p = 6$) will have a lower quality of life weight than never having a cardiovascular event ($h_p = 1$).

With a discount rate λ and state transition probabilities $p_\tau(\mathbf{s}_p, \mathbf{s}'_p)$, we can recursively determine the optimal policy and compute the optimal value function $V_{t,p}(\mathbf{s}_p)$ as:

$$\begin{aligned} V_{t,p}(\mathbf{s}_p) &= \max_{\tau \in \mathcal{A}(b_{min}(\mathbf{s}_p))} \left\{ -d_{\tau,p}(\mathbf{s}_p) + \sum_{\mathbf{s}'_p \in \mathcal{S}} p_\tau(\mathbf{s}_p, \mathbf{s}'_p) \left[q_p(\mathbf{s}'_p) + \lambda V_{t+1,p}(\mathbf{s}'_p) \right] \right\} \quad t = 1, \dots T \quad (1) \\ &= \max_{\tau \in \mathcal{A}(b_{min}(\mathbf{s}_p))} \left\{ f^\tau_{t,p}(\mathbf{s}_p) \right\} \quad t = 1, \dots T, \end{aligned}$$

where $f^\tau_{t,p}(\mathbf{s}_p)$ is the state-action pair value function for taking action τ when in state $\mathbf{s}_p$ at time t. It is assumed that the optimal value function and quality of life weight take value of zero when the patient is in a dead health state ($h_p \in \{7, 8, 9\}$).

Given the optimal preventative treatment policy π^*_p, we can directly evaluate the controlled Markov chain to compute the expected total number of medications $M_p(\alpha_p)$, patient class p will take when the adherence is α_p, as well as the vector of expected total number of CV events $\mathbf{E}_p(\alpha_p)$ (comprised of elements for the number of fatal MI events, nonfatal MI events, fatal strokes, and nonfatal strokes).

3.2 Optimal Coinsurance Rate

Define n_p to be the expected number of patients in class p at the beginning of the planning horizon. Let $V_p(\alpha_p)$ be the expected total discounted QALYs over the planning horizon (the value at $t = 1$) for class p patients when their adherence level is α_p and they follow policy π^*_p (as obtained in Section 3.1). For each patient class p, we aim to determine the optimal drug coverage $y_p \in [0, 1]$, defined as the percentage of medication cost covered by the insurance provider. We also assume that each patient class' adherence α_p changes as a function of y_p and state information for class p, i.e., $g(y_p, \mathbf{s}_p) = \alpha_p$. While we do not specify an explicit form $g(\cdot)$ in our model formulation, we assume that α_p is increasing in y_p. We provide an example of such a function in our numerical study, which is covered in Section 5.

Let $\mathbf{y} = [y_1, ..., y_N]$ denote the vector of coverage decisions for all $p \in P$. Given $\mathbf{y}$ and the drug manufacturer's price c_τ (i.e. the total cost charged to the insurance provider for each medication τ before costs are recuperated via coinsurance), we can compute the drug cost $C(\mathbf{y})$ as:

$$C(\mathbf{y}) = \sum_{p \in P} \sum_{\tau \in \mathcal{A}} c_\tau \cdot n_p \cdot y_p \cdot \alpha_p \cdot M_{\tau,p}(\alpha_p). \tag{2}$$

The provider is assumed to have a limited budget for expenditures on decreasing out-of-pocket expenses for its patient population, B, which constrains the total cost-sharing relief available from the provider. We can write this resource constraint as $C(\mathbf{y}) \leq B$.

We further assume that the provider wants to limit the variability of coinsurance rates (which occurs when coinsurance rates are inequal) across the patient classes. To compute inequality, we propose to use a function $W(\mathbf{y})$ paired with a maximum inequality threshold for the population $w \in [0, 1]$. While we do not provide an explicit form for $W(\mathbf{y})$ in the model formulation, we assume that, as the difference between allocation between classes increases, so does the value of $W(\mathbf{y})$. Examples of such a function $W(\mathbf{y})$ include the Gini index and Thiel index (Mussard et al. 2003). A low value of $W(\mathbf{y})$ corresponds to a more equal distribution of coinsurance rates and a high value of $W(\mathbf{y})$ corresponds to more inequal distribution. Therefore, with a maximum inequality threshold for the population w, we can appropriately limit the coinsurance inequality caused by the coverage policy to match the desired level of fairness.

We can formulate the full mathematical program as:

$$Z^* = \max_{\mathbf{y}} \left\{ \sum_{p \in P} n_p \cdot V_p(\alpha_p) \right\} \tag{3}$$

$$\text{s.t.} \quad C(\mathbf{y}) - C(\mathbf{y}^i) \leq B \tag{4}$$

$$W(\mathbf{y}) \leq w \tag{5}$$

$$g(y_p, \mathbf{s}_p) \in [0,1] \text{ for all } p \in P \tag{6}$$

$$y_p \in [0,1] \text{ for all } p \in P, \tag{7}$$

where $\mathbf{y}^i$ denotes the vector of current medication coverage for each patient class.

3.3 Bilevel Optimization Problem

From the formulation above, we see that the optimal preventative treatment policy π_p^* from the MDP model is embedded within constraint (4) since it determines $M_{\tau,p}(\alpha_p)$. Furthermore, the objective function (3) also contains the optimal value function of the MDP for each patient class $V_p(\alpha_p)$. Since the optimal preventative treatment policy derived by solving a MDP specific to class p depends on the adherence level of class p, then π_p^* is a function of the drug coverage decision variable y_p. This mathematical program shows the complex nature of the provider/patient relationship and the necessary hierarchy of the optimization model. The hierarchical structure of the coinsurance problem translates to a bilevel optimization problem (BLP) (Bard 1998). BLPs are characterized by a leader solving an upper level optimization problem and followers solving lower level optimization problems. In a BLP, the follower's lower level optimization problem depends on the solution to the leader's upper level optimization problem. Similarly, the optimal solution to the leader's problem depends on the outcome of the follower's solutions to their optimization problems. In the coinsurance BLP, the leader is the provider and the followers are the patient classes. The leader's optimization problem is to determine drug coverages for each patient class which maximizes total expected discounted QALYs subject to constraints on the coverage budget and inequality. Each follower's optimization problem is to determine the optimal preventative treatment

policy given the drug coverage set by the provider.

3.4 Key Model Properties

We next investigate features of the bilevel optimization model to provide insight into how medication coverage rates should be determined. Throughout this section, we define CV risk to be the sum of 10-year MI risk and 10-year stroke risk. While we tailor the implications of our analytical results to this definition of CV risk, the derivation of key model properties do not depend on this definition. The essence of CV risk is captured as long as the CV risk is greater for patients who are more likely to experience MI, strokes, and/or other related events.

In the following analytical results, we examine the effect of CV risk on the marginal health benefit of adherence and conclude that as CV risk increases, the benefit of improved adherence increases. This implies that patients at higher risk for myocardial infarction and stroke should be prioritized for adherence improving interventions.

Property 1: *If post-treatment CV risk $r_\tau(\boldsymbol{s}_p)$ is decreasing and convex in adherence α and linear in pre-treatment CV risk r, the marginal benefit of adherence $\frac{\partial V_p}{\partial \alpha}$ increases as pre-treatment CV risk r increases, i.e. $\frac{\partial \frac{\partial V_p}{\partial \alpha}}{\partial r} > 0$.*

Proof: If we show that $\frac{\partial \frac{\partial f_\tau^t(\mathbf{s}_p)}{\partial \alpha}}{\partial r} > 0$ for all τ, then we have shown that $\frac{\partial \frac{\partial V_p}{\partial \alpha}}{\partial r} > 0$. By construction,

$$\frac{\partial f_\tau^t(\mathbf{s}_p)}{\partial \alpha} = -\frac{\partial d_{\tau,p}(\mathbf{s}_p)}{\partial \alpha} + \frac{\partial E_{\tau,p}[Q(\mathbf{s}_p)]}{\partial \alpha}. \tag{8}$$

Since $d_{\tau,p}$ does not depend on risk, the partial derivative with respect to r of the partial derivative of $f_\tau^t(\mathbf{s}_p)$ with respect to α is:

$$\frac{\partial \frac{\partial f_\tau^t(\mathbf{s}_p)}{\partial \alpha}}{\partial r} = \frac{\partial \frac{\partial E_{\tau,p}[Q(\mathbf{s}_p)]}{\partial \alpha}}{\partial r}. \tag{9}$$

Taking the partial derivative of the expected cost-to-go function with respect to α:

$$\begin{aligned}\frac{\partial E_{\tau,p}[Q(\mathbf{s}_p)]}{\partial \alpha} &= (1-\rho(\mathbf{s}_p))\left[q(h=\{5,6\})+\lambda V_{t+1,p}(h=\{5,6\})\right]\frac{\partial r_\tau(\mathbf{s}_p)}{\partial \alpha} \\ &\quad -\left[q(h=\{1,2,3,4\})+\lambda V_{t+1,p}(h=\{1,2,3,4\})\right]\frac{\partial r_\tau(\mathbf{s}_p)}{\partial \alpha}.\end{aligned} \tag{10}$$

Then taking the partial derivative with respect to r, we have:

$$\begin{aligned}\frac{\partial \frac{\partial E_{\tau,p}[Q(\mathbf{s}_p)]}{\partial \alpha}}{\partial r} &= (1-\rho(\mathbf{s}_p))q(h=\{5,6\})\frac{\partial \frac{\partial r_\tau(\mathbf{s}_p)}{\partial \alpha}}{\partial r} \\ &\quad +\lambda(1-\rho(\mathbf{s}_p))\left[\frac{\partial \frac{\partial r_\tau(\mathbf{s}_p)}{\partial \alpha}}{\partial r}V_{t+1,p}(h=\{5,6\})+\frac{\partial r_\tau(\mathbf{s}_p)}{\partial \alpha}\frac{\partial V_{t+1,p}(h=\{5,6\})}{\partial r}\right] \\ &\quad -q(h=\{1,2,3,4\})\frac{\partial \frac{\partial r_\tau(\mathbf{s}_p)}{\partial \alpha}}{\partial r}-\lambda V_{t+1,p}(h=\{1,2,3,4\})\frac{\partial \frac{\partial r_\tau(\mathbf{s}_p)}{\partial \alpha}}{\partial r} \\ &\quad -\frac{\partial V_{t+1,p}(h=\{1,2,3,4\})}{\partial r}\frac{\partial r_\tau(\mathbf{s}_p)}{\partial \alpha}.\end{aligned} \tag{11}$$

Since $r_\tau(\mathbf{s}_p)$ is decreasing and convex in α and linear in r, we have that $\frac{\partial \frac{\partial r_\tau(\mathbf{s}_p)}{\partial \alpha}}{\partial r} < 0$. And because $V_{t+1,p}$ is decreasing in r, we have $\frac{\partial V_{t+1,p}}{\partial r} < 0$. Rearranging terms we have:

$$\begin{aligned}\frac{\partial \frac{\partial E_{\tau,p}[Q(\mathbf{s}_p)]}{\partial \alpha}}{\partial r} &= (1-\rho(\mathbf{s}_p))q(h=\{5,6\})\frac{\partial \frac{\partial r_\tau(\mathbf{s}_p)}{\partial \alpha}}{\partial r}-q(h=\{1,2,3,4\})\frac{\partial \frac{\partial r_\tau(\mathbf{s}_p)}{\partial \alpha}}{\partial r} \\ &\quad +\lambda(1-\rho(\mathbf{s}_p))\frac{\partial \frac{\partial r_\tau(\mathbf{s}_p)}{\partial \alpha}}{\partial r}V_{t+1,p}(h=\{5,6\})-\lambda V_{t+1,p}(h=\{1,2,3,4\})\frac{\partial \frac{\partial r_\tau(\mathbf{s}_p)}{\partial \alpha}}{\partial r} \\ &\quad +\lambda(1-\rho(\mathbf{s}_p))\frac{\partial r_\tau(\mathbf{s}_p)}{\partial \alpha}\frac{\partial V_{t+1,p}(h=\{5,6\})}{\partial r}-\frac{\partial V_{t+1,p}(h=\{1,2,3,4\})}{\partial r}\frac{\partial r_\tau(\mathbf{s}_p)}{\partial \alpha}.\end{aligned} \tag{12}$$

Given the following inequalities:

$$(1-\rho(\mathbf{s}_p))q(h=\{5,6\}) < q(h=\{1,2,3,4\}) \tag{13}$$

$$\lambda(1-\rho(\mathbf{s}_p))V_{t+1,p}(h=\{5,6\}) < \lambda V_{t+1,p}(h=\{1,2,3,4\}) \tag{14}$$

$$\lambda(1-\rho(\mathbf{s}_p))\frac{\partial V_{t+1,p}(h=\{5,6\})}{\partial r} < \frac{\partial V_{t+1,p}(h=\{1,2,3,4\})}{\partial r}. \tag{15}$$

We therefore have that $\frac{\partial \frac{\partial E_{\tau,p}[Q(\mathbf{s}_p)]}{\partial \alpha}}{\partial r} > 0$. As this was true for any τ, α and r, then $\frac{\partial \frac{\partial V_p}{\partial \alpha}}{\partial r} > 0$.

■

In addition to prioritizing patients based upon their risk for CVD, the marginal health benefit of adherence also depends on the patient's sensitivity to price, as illustrated by equation (6).

Remark *Ceteris paribus, patient classes with higher sensitivity to changes in medication cost, e.g., price elasticity, have higher priority for coinsurance rate reduction in the optimal medication coverage policy* $\boldsymbol{y}^*$.

Holding all other factors constant, the total health of the population is greater when resources are directed toward those patients whose adherence is more sensitive to medication cost. Therefore, within a risk group, the optimal policy prioritizes those patient classes with higher price elasticities, e.g. low-income patients.

Finally, we analyze how the social welfare function $W(\mathbf{y})$ used to constrain maximum inequality is related to the heterogeneity of the patient classes and the budget for reducing coinsurance rates. Let cardiovascular risk heterogeneity $H(P)$ for patient mix $P = \{p_1, \ldots, p_N\}$ be defined as $\sum_{i=1}^{N} \sum_{j=1}^{N} |r_{p_i} - r_{p_j}|$, i.e. the total absolute difference in cardiovascular risk. As heterogeneity in CV risk increases, the heterogeneity in marginal health benefits of improved adherence also increases. Greater differences in marginal health benefits lead to greater inequality in medication coverage when maximizing the utilitarian objective function (3).

Property 2: *The measure of inequality* $W(\boldsymbol{y}^*)$ *imposed by the optimal medication coverage policy* $\boldsymbol{y}^*$ *is nondecreasing as CV risk heterogeneity* $H(P)$ *increases.*

Proof: From Property 1, we have that $\frac{\partial \frac{\partial V_p}{\partial \alpha}}{\partial r} > 0$, i.e. as CV risk increases, the marginal benefit of improved adherence increases. Therefore, ceteris paribus, the optimal policy will prioritize higher risk patient classes for more coverage allocation y. Hence, if $r_{p_i} > r_{p_j}$ then $y_{p_i} \geq y_{p_j}$. Or more generally, as $|r_{p_i} - r_{p_j}|$ increases, $|y_{p_i} - y_{p_j}|$ is nondecreasing. By assumption on the properties of $W(\mathbf{y})$, as $|y_i - y_j|$ increases, $W(\mathbf{y})$ increases. Then as $H(P)$ increases, $W(\mathbf{y})$ is nondecreasing. ■

The resultant higher inequality implies that the constraint $W(\mathbf{y}) \leq w$ is more likely to be active as $H(P)$ increases. When the $W(\mathbf{y}) \leq w$ constraint is active, there is a health

cost imposed by requiring the inequality to not exceed w. Let the health cost of equality be defined as $\Delta_w(b) = Z^*(w = 1, B = b) - Z^*(w = 0, B = b)$, i.e. the difference in QALYs between the utilitarian policy and the equal coverage policy when the resource budget B is b. The cost of equality depends on the CV risk heterogeneity $H(P)$ as well as the resource budget b.

Property 3: *The marginal benefit of higher maximum allowable inequality w is nondecreasing as cardiovascular risk heterogeneity $H(P)$ increases.*

Proof: Let $H(P'') > H(P')$ and let $\mathbf{y}^*_P$ be the optimal coverage for patient mix P. By the above theorem, $W(\mathbf{y}^*_{P''}) \geq W(\mathbf{y}^*_{P'})$. Then, for a given w, if the constraint $W(\mathbf{y}^*) \leq w$ is active for P', then the constraint is active for P''. However, if the constraint is active for P'', the constraint may not be active for P'. Therefore, there is no marginal benefit to increasing w when the constraint is not active for P', but there may be a marginal benefit for P''. Therefore, the marginal benefit is nondecreasing as $H(P)$ increases. ■

Property 4: *If the optimal policy $\boldsymbol{y}^*$ does not include $y^*_p = 0$ for some $p \in P$, then as the medication coverage budget B increases, the cost of equality $\Delta_w(b)$ is nonincreasing.*

Proof: $\Delta_w(b) = Z^*(w = 1, B = b) - Z^*(w = 0, B = b)$. For $w = 1$, if $y^*_p \neq 0$ for all p, then each additional unit of budget will be allocated toward a patient class $p \in P$ such that $y^*_p < 1$. Due to the ceiling effect of $y \leq 1$, the increase in Z^* from the additional unit of budget is nonincreasing as B increases since low priority patient classes will be allocated resources once high priority classes are at full coverage. For $w = 0$, each additional unit of budget is allocated equally across all classes until all classes have $y^*_p = 1$, at which point $\Delta_w(b) = Z^*(w = 1, B = b) - Z^*(w = 0, B = b) = 0$. Therefore, the difference in objective function values is nonincreasing as B increases. ■

The property implies that one method available to policymakers for addressing concerns about equality is to increase the budget for coinsurance rate reduction. The additional funds will naturally go toward lower priority patients under the utilitarian policy, since higher priority patients will already have full medication coverage. This further implies

that being unequal in coverage will be less costly to total population health if the budget is increased. Consequently, when Medicare is budget-neutral ($B = 0$) and there is no constraint on maximum inequality ($w = 1$), the social welfare function $W(\mathbf{y}^*)$ is maximized. This implies that we can determine the highest w threshold necessary to evaluate in our numerical analysis by first solving the BLP with $B = 0$ and $w = 1$. By determining the highest w first, we can reduce the total number of scenarios required for evaluating the trade-offs between equality and population health.

Property 5: *If the optimal policy* $\boldsymbol{y}^*$ *does not include* $y_p^* = 0$ *for some* $p \in P$, $W(\boldsymbol{y}^*)$ *is maximized when* $B = 0$ *and* $w = 1$.

Proof: As shown in the proof to the prior corollary, if $y_p^* \neq 0$ for all p, then each additional unit of budget will be allocated toward a patient class $p \in P$ such that $y_p^* < 1$. The additional resources reduce the differences in coverages across the patient classes, i.e. $W(\mathbf{y}^*)$ decreases as B increases when $w = 1$. And since $w = 1$ allows for the maximum possible W at a given budget, then $W(\mathbf{y}^*)$ is maximized when $B = 0$ and $w = 1$. ■

4 Solution Methodology

We propose to solve the coinsurance BLP using dynamic programming and penalty functions for nonseparable constraints. Constraints on the action space in a dynamic programming formulation can be difficult to handle depending on their nature. Generally, we can view these constraints as separable and nonseparable constraints. Global resource constraints (i.e. constraint on the total amount of resources spent over the planning horizon) are separable constraints and can be handled via state-space inclusion of a discretized resource parameter, starting with B which decreases at each iteration of the recursive Bellman's equations.

Nonseparable constraints are more challenging since state-space expansion is ill-suited to handle the restrictions placed on the action space. Limiting the maximum allowable inequality as computed by a social welfare function of the drug coverage policy is a nonseparable constraint. The welfare function is nonseparable because each term in the function depends on every element of the policy vector $\mathbf{y}$. One way to handle nonseparable constraints is via

penalty function. Given a penalty weight u, we can penalize the objective function when $W(\mathbf{y}) > w$. By treating the hard constraint of $W(\mathbf{y}) \leq w$ as a soft constraint with penalty weight u in the objective function, we are able to solve the dynamic program quickly. And by using a sufficiently high value for u, we can guarantee that solutions to the dynamic program never violate the constraint $W(\mathbf{y}) \leq w$. We can then formulate and solve the problem by recursively computing $Z(p, b, w)$, which we define as follows:

$$Z(p, b, w) = \max_{\mathbf{y} \in [0,1]^p} \left\{ R(b', p) + Z(b - b', p - 1) - u(W(\mathbf{y}) - w)^+ \right\}. \tag{16}$$

Equation (16) is the optimal value function corresponding to the maximized total expected discounted QALYs when Medicare has b coverage resources to allocate across p patient classes with a maximum allowable inequality of w. We obtain (16) recursively by solving for increasing values of $p = 1, \ldots, N$, where within each iteration we recursively determine the best policy when allocating $b' \leq b$ to patient class p (which returns $R(b', p)$, i.e., the QALYs for class p) and $b - b'$ to patient classes $1, \ldots, p - 1$ (which results in $Z(b - b', p - 1)$, i.e., the total QALYs for classes $1, \ldots, p - 1$) over all $b' \leq b$. When $p = 1$, we determine the best policy when allocating b to the sole patient class. At each recursion, we store the optimal policy $\mathbf{y}^*$ when there are p patient classes and b resources to allocate. This stored optimal policy is used when computing the penalty function $u(W(\mathbf{y}) - w)^+$ of the next recursion.

5 Numerical Study

In our numerical study, we use the Gini index as our measure of inequality, $W(\mathbf{y})$, because it is widely used in economic inequality studies (Cowell 2000). The Gini index is given by

$$W(\mathbf{y}) = \frac{\sum_{i=1}^{P} \sum_{j=1}^{P} |y_i - y_j|}{2P^2 \bar{\mathbf{y}}} \in [0, 1], \tag{17}$$

where $\bar{\mathbf{y}}$ is the average allocation to all patient classes. It is clear that the Gini index satisfies the properties in Section 3.4. Additionally, we use the theory of price elasticities to model the change in adherence for patients in class p as a function of the coverage decision y_p and

state information $\mathbf{s}_p = (\alpha_p^i, \delta_p^\alpha, y_p^i)$:

$$g(y_p, \mathbf{s}_p) := \alpha_p = \alpha_p^i + \delta_p^\alpha \cdot (y_p - y_p^i) \in [0, 1], \tag{18}$$

where α_p^i is the initial adherence, δ_p^α is the price elasticity of adherence, and y_p^i is the current coverage level. While other functional forms may be used to model adherence changes as a function of coverage, we use (18) due to its ease of interpretation.

Since the minimum enrollment age for Medicare is 65 (Med), we parameterized the model using a population of patients age 65 and older who were randomly sampled from the National Health and Nutrition Examination Survey (NHANES). NHANES is unique in that it is nationally-representative, large, and includes information from surveys, clinic data, and laboratory information. It has been the centerpiece of numerous United States policy-based simulation models, including the classic CHD Policy Model (Weinstein et al. 1987), and other recent CHD policy studies (Timbie et al. 2010, Bibbins-Domingo et al. 2010, Sussman et al. 2011, 2013, Moran et al. 2015). We specifically parameterize our models using NHANES III, which consists of survey data collected between 1988-1994 when anti-hypertension medication and cholesterol-lowering statins were scarcely used, allowing us to model the role of these medications on cardiovascular risk reduction compared to non-use. Based on NHANES' prevalence and importance in many United States-based CHD studies, we believe this sample population is a good representation of patients who are enrolled in Medicare. In Table 2, we summarize the relevant characteristics of the data sample used in our numerical analysis.

We assumed a 10 year planning horizon for treatment and drug coverage decisions. In our analysis, we considered 4 patient classes based on cardiovascular risk and sensitivity to medication cost: high risk, high sensitivity; low risk, high sensitivity; high risk, low sensitivity; and low risk, low sensitivity. To identify high risk and low risk patients, we computed the pre-treatment CV risk for each patient as the sum of the 10-year MI and stroke risks based on the calculation by Anderson et al. (1991). Patients with CV risk less than the 25^{th} percentile were assumed to be low risk while those above the 75^{th} percentile

Table 2: Summary of NHANES III data sample used in numerical analysis.

Variable	
Sample size	54887
	n (%)
White	35270 (64.26)
Male	25479 (46.42)
CVD History	3655 (6.66)
Diabetic	6449 (11.75)
Smoker	9622 (17.53)
	Mean (SD)
Age	73.73 (5.99)
Total Cholesterol (mg/dL)	224.75 (42.65)
HDL (mg/dL)	52.71 (16.19)
SBP (mmHg)	145.98 (20.87)
DBP (mmHg)	76.81 (11.27)

were assumed to be high risk. Based on health economics literature, we assumed high (low) sensitivity patient classes have an initial adherence of 38% (34%) (Medicare Payment Advisory Commission and others 2014) and a price elasticity of adherence of 0.21 (0.29) (Manning et al. 1987). Currently under Medicare, 75% of the cost of generic medication is covered for all patients. Table 3 summarizes the key inputs to the BLP.

Table 3: Coinsurance BLP inputs and data sources. All costs reported are in 2015 USD after using appropriate cumulative inflation rates, e.g. 9.9% for 2008 to 2015.

Input	**Value**	**Source**
Patient data	NHANES III	(Timbie et al. 2010)
Current coverage $\mathbf{y}^i$	75%	(Klees et al. 2010)
Initial adherence for high SES patients α_p^i	38%	(Medicare Payment Advisory Commission and others 2014)
Initial adherence for medium SES patients α_p^i	34%	(Medicare Payment Advisory Commission and others 2014)
Price elasticity of adherence for high SES patients δ_p^α	0.21	(Manning et al. 1987)
Price elasticity of adherence for medium SES patients δ_p^α	0.29	(Manning et al. 1987)
Risk calculator	Framingham	(Anderson et al. 1991)
Treatment benefit	Meta-analysis of RCTs	(Law et al. 2009)
Mortality and fatality likelihoods	CV event and other-cause	(Koek et al. 2006, Arias 2007)
Cost of nonfatal MI	\$18,202	(Choudhry et al. 2008)
Cost of fatal MI	\$15,929	(Choudhry et al. 2008)
Cost of nonfatal stroke	\$15,251	(Choudhry et al. 2008)
Cost of fatal stroke	\$10,345	(Choudhry et al. 2008)
Cost of non-CV death	\$11,059	(Choudhry et al. 2008)
Annual cost of generic hypertension medication	\$233	(Choudhry et al. 2008)
Discount rate λ	3%	
Hypertension medication disutility per drug $d_\tau(\mathbf{s})$	0.001	(Sussman et al. 2013)

5.1 Optimal Prescription Medication Coverage

For the BLP model parameterized as described above, we solved the BLP using the solution methodology discussed in Section 4. Figure 1 shows the optimal BLP prescription medication coverage policy at different budget levels when there is no constraint on inequality, i.e. $w = 1$. We find that the optimal BLP policy matches the policy derived analytically in Section 3.4. When Medicare has been allocated no additional funds by Congress to reduce out-of-pocket medication expenses for its patients, i.e. budget-neutral or $B = 0$, the BLP policy prioritizes high risk patients over low risk patients for coinsurance rate reduction. The high risk patients have QALYs which are more sensitive to adherence to medication and therefore yield a greater return for each invested dollar of coinsurance reduction. Within each risk group (high and low), the BLP policy prioritizes high sensitivity over low sensitivity patients. When budget-neutral, the BLP policy provides free medication to high risk patients, but must decrease coverage for low risk patients to recuperate the costs of free medication for high risk patients. As the total budget increases, the additional resources are allocated to low risk, high sensitivity patients until their coverage is 100%. Then, all remaining resources are allocated to the lowest priority class: low risk, low sensitivity patients.

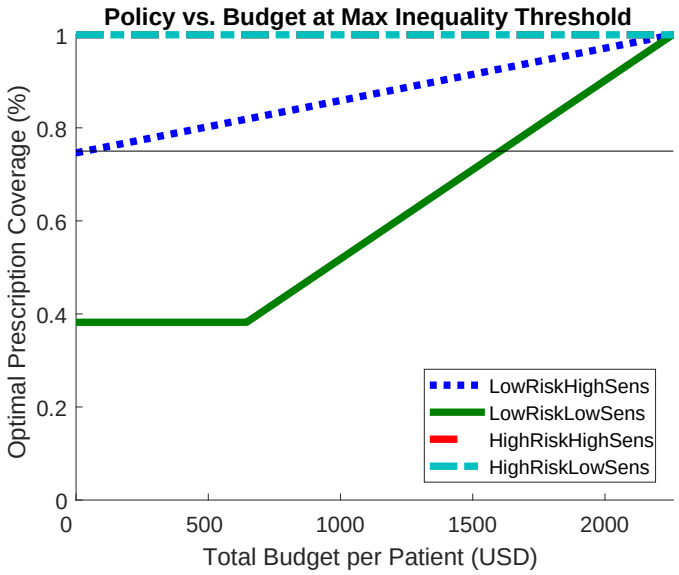


Figure 1: Optimal coverage decisions at different budget levels assuming no constraint on inequality. Current Medicare coverage (75%) is denoted by the horizontal bar.

To evaluate whether the rule of prioritizing CV risk before cost sensitivity is robust to the price elasticity of adherence, we varied the ratio of the price elasticity of adherence of high sensitivity to low sensitivity patient classes. Under the baseline ratio of 1.377 (= 0.292/0.212), the optimal BLP policy prioritizes coverage based on CV risk first and cost sensitivity second. Our analysis reveals that over a 300% increase in the ratio ($\delta_p^\alpha = 1.30$ for high sensitivity and 0.212 for low sensitivity) results in the same optimal BLP policy. This result suggests that the marginal health benefit of improved adherence for high CV risk patients strongly dominates the marginal health benefit for low CV risk patients such that more than a 4-fold increase in price elasticity for high sensitivity patients is insufficient to warrant prioritization of coverage based on cost sensitivity before CV risk.

5.2 Health Performance of Optimal BLP Policy

Next, we evaluate the health performance of the optimal BLP policy when there is no constraint on inequality. For the BLP policies described in Section 5.1, we computed the change in the objective function value, the expected discounted QALYs per 1000 patients, with respect to the current 75% coverage for all patients at each budget level. Figure 2 shows that the QALYs saved per 1000 patients under the BLP policies (when there is no constraint on inequality) range from 13.6 QALYs when budget-neutral to 24.9 QALYs when the budget is high enough for all patient classes to have zero out-of-pocket expenses for medication. This result implies that allocating enough resources to Medicare so that all patients receive free medication would save an additional 11.3 QALYs per 1000 patients. The health performance of the BLP policy when Medicare is budget-neutral indicates that the BLP policy is able to improve total population health without increasing resources. This improvement, however, requires the ability to tailor coinsurance rates to the characteristics of each patient class which naturally generates inequality in coverage.

5.3 Health Cost of Equality

Given the potential importance of equality in coverage for a government-run public health insurance provider like Medicare, we next compute the health cost of equality: $\Delta_w(b) =$

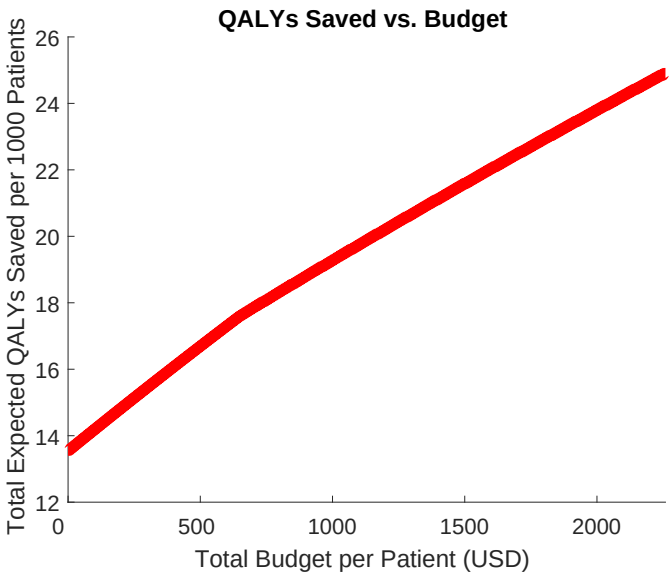
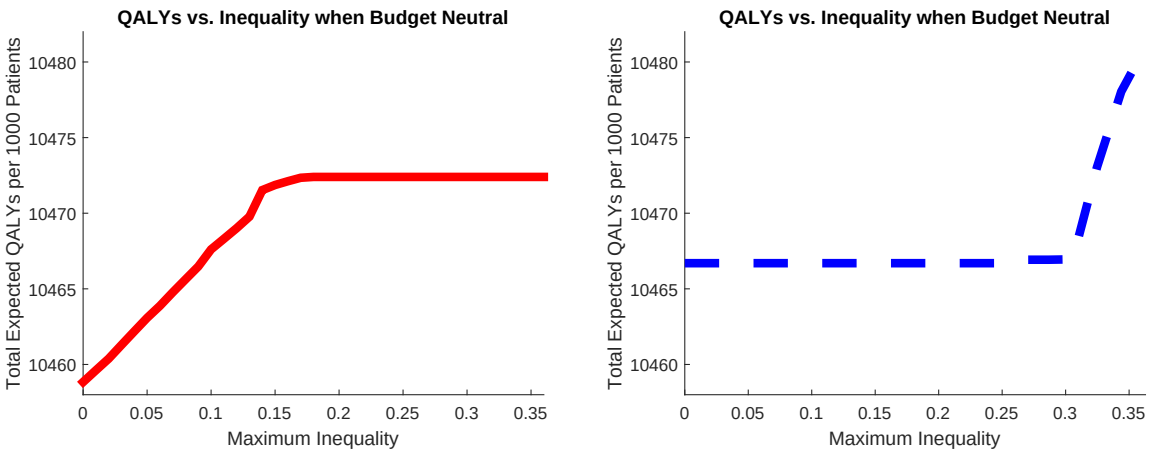


Figure 2: Quality-adjusted life years (QALYs) saved with respect to current practice at each budget level when Medicare has no constraint on inequality.

$Z^*(w = 1, B = b) - Z^*(w = 0, B = b)$. The maximum cost of equality occurs when Medicare is budget-neutral. Figure 3a shows the expected QALYs per 1000 patients as a function of the maximum inequality threshold w when $B = 0$. Comparing the QALYs when $w = 0$ to when $w = 1$, we can compute the cost of equality $\Delta_w(0)$ as approximately 13.6 QALYs. Note that the difference in QALYs is more drastic when comparing lower w values than higher w values. This difference occurs because at higher w, the changes in the optimal policy are with respect to the low risk classes (high risk classes have full coverage in the higher w range). The low risk classes are less sensitive to changes in adherence and therefore do not contribute heavily to changes in QALYs once high risk classes have full coverage. We observe that the gain in QALYs is much larger for the initial deviations from equal coverage (small w) than when the maximum inequality gets closer to 1. This indicates that even small flexibility in tailoring coinsurance rates based on patient class can have significant effect on the health outcomes of the Medicare population. It is not necessary to have full tailoring of coinsurance rates to accrue significant improvements in QALYs, and small deviations from equal coverage may be politically feasible for Congress.

As we observed, the health cost of equality is substantial and should be considered when



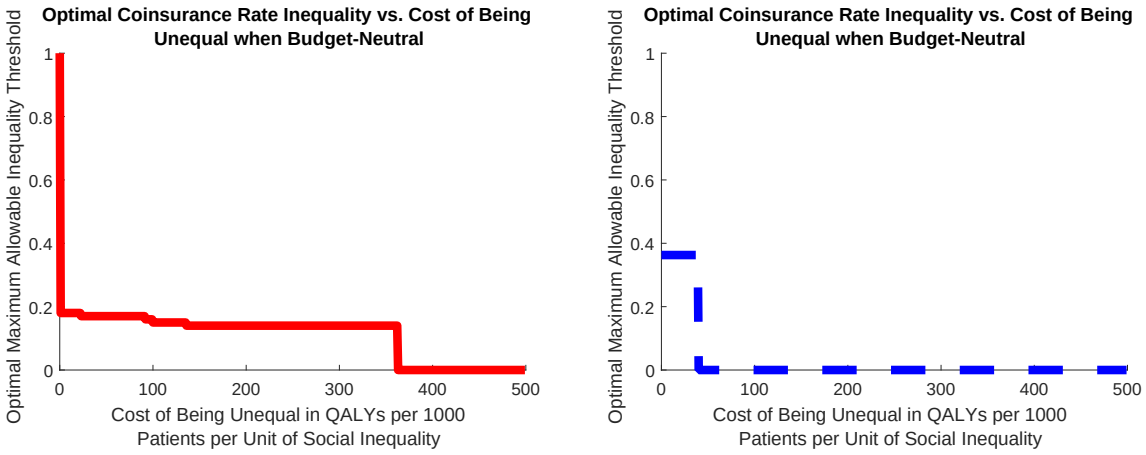
(a) Coverage inequality

(b) Health inequality

Figure 3: Quality-adjusted life years (QALYs) at each maximum inequality threshold when Medicare is budget-neutral.

determining how strictly equality of coverage should be enforced in public policy. One method to determine the w to use when designing policy is to consider the trade-off of maximized population health with a cost associated with being unequal in coverage. Consider a parameter β which measures how many QALYs per 1000 patients Medicare is willing to forgo per unit of social inequality as measured by $W(\mathbf{y})$. A high value of β indicates that the cost of being unequal is high to Medicare (e.g. political pressure for equal coinsurance rates is high). Then for a given budget B, we seek the w^* which maximizes $Z^*(B, w) - \beta \cdot W(\mathbf{y}^*)$ where $Z^*(B, w)$ is the maximized QALYs per 1000 patients under the optimal policy $\mathbf{y}^*$ when the maximum inequality threshold is w. Figure 4a plots w^* as a function of β when Medicare is budget-neutral. We observe that the optimal w^* is highly sensitive to β, particularly at low costs of being unequal. This indicates that the optimal maximum inequality threshold depends strongly on how important equality of coverage is to Medicare, as measured by β.

We also considered an alternative definition of inequality: inequality in health outcomes associated with the new coverage policy. We first define the health change for patient class p as $X_p = V_p(\alpha_p) - V_p(\alpha_p^i)$, i.e., the change in expected QALYs under the new coinsurance rate. Let $\mathbf{X} = [X_1, ..., X_N]$ denote the vector containing X_p for all $p \in P$. Then, our measure of health inequality is given by $W(\mathbf{X})$, where $W(\cdot)$ is the Gini index defined in (17). The health cost of equality is naturally the same for both definitions of inequality. However,



(a) Coverage inequality

(b) Health inequality

Figure 4: Optimal maximum inequality threshold w^* at different costs of being unequal parameters β when $B = 0$.

the change in QALYs at each value of w differs greatly. Figure 3b plots the health cost of equality under the alternative definition of health outcome inequality.

For coverage inequality, there is a sharp rise in QALYs at low levels of w since these values of w allow the prioritization of high risk patients over low risk patients. For health inequality, there is little increase in QALYs at low levels of w because all health outcomes are constrained to be nearly equal. As medium levels of w increase toward the maximum w, the change in QALYs is small for inequality based on coverage, whereas it is large for inequality based on health outcomes. We also find differences in the optimal maximum inequality threshold. Figure 4b shows that the optimal maximum inequality threshold goes to 0 very quickly as the QALY cost of being unequal increases. In fact, it drops from the maximum value of 0.36 to 0 almost instantaneously when $\beta = 40$, compared to the step-wise drop under coverage inequality. Under health inequality, the optimal BLP policy is able to prioritize coverage for high risk patients only when w is sufficiently high. This policy translates to an optimal w under health inequality that decreases to zero much faster than the optimal w under the coverage inequality definition as the β parameter for the QALY trade-off increases.

5.4 Cost-Effectiveness of BLP Policy

Next, we consider the change in total monetary costs relative to the current Medicare coverage of 75%. The costs we consider when computing total change in costs are medication costs and CV event costs. As prescription medication coverage $\mathbf{y}$ increases, the medication costs $C(\mathbf{y})$ should increase. However, as coverage increases, the corresponding improvement in medication adherence should lead to fewer CV events $\mathbf{E}_p(\alpha_p)$ over the planning horizon. With the cost of each CV event (both fatal and nonfatal) dwarfing the cost of each medication, the preventative nature of hypertension medication can lead to cost savings.

At low budget levels, the BLP policy is cost-saving (negative cost change). As the budget increases, the change in total costs also increases with small amounts of sensitivity to the maximum inequality threshold. The changes in total cost eventually remains near constant for each fixed budget level even as the maximum inequality threshold becomes very large. This phenomenon is due to being able to prioritize high risk patients to 100% coverage when w is increased slightly. After full coverage for high risk patients, the changes in coverage for the lower risk patients has small effects on costs. Furthermore, as the budget increases, the BLP policy goes from cost-saving to cost-increasing. This result implies that at higher budgets, the change in medication costs outpaces the cost reduction from fewer CV events. This implication is especially true when increasing the budget once high risk patients are at full coverage. The increase in costs may suggest that the preventative benefits of hypertension medication are not sufficiently clinically beneficial to warrant coinsurance rate reductions for low risk patients.

To appropriately evaluate the economic value of the BLP policies, we next compute the incremental cost-effectiveness ratio (ICER) at all budget and inequality thresholds. ICER is defined as the ratio of change in costs to change in QALYs. When the change in costs is low and the change in QALYs is high, the ICER measure is small and indicates a more cost-effective policy. Paired with a cost-effectiveness threshold, the ICER can be used to determine if a policy is cost-effective. While a threshold of $50,000 per QALY is most com-

monly used in the United States (Grosse 2008), critics have suggested that thresholds ranging between $50,000 and $200,000 should be considered (Neumann et al. 2014). Therefore, in our numerical analysis, we vary this threshold from $50,000 to $200,000. Figure 5 maps the cost-effectiveness of the BLP policies for each cost-effectiveness threshold under each combination of B and w. Each shade indicates the smallest cost-effectiveness threshold at which a specific combination of B and w is cost-effective. We find that every parameter combination is cost-effective at a threshold of $100,000 per QALY. When budget neutral, the BLP policy is cost-saving, and thus cost-effective regardless of maximum inequality. At small to medium budgets, the BLP policy is not cost-effective at $50,000 per QALY when coverage must be equal across all patients (i.e. $w = 0$). However, when the maximum inequity threshold is increased to allow for prioritization of patient classes and the tailoring of coinsurance rates, the BLP policy becomes cost-effective. As the budget increases such that all patients receive free medications, the BLP policy becomes cost-ineffective requires higher cost-effectiveness thresholds to justify the additional cost of covering medications for low-risk patients. By the ICER measure, our results indicate that providing free medication to all low risk patients is not cost-effective at $50,000 per QALY.

We performed multiple sensitivity analyses on the cost of medication and CV events, the price elasticity of adherence, and the current Medicare coverage for each patient class to determine if the cost-effectiveness strongly depends on these parameters. In this analysis, we fixed the cost-effectiveness threshold at $50,000 per QALY since this threshold is widely used in assessing the cost-effectiveness of healthcare expenditures in the United States (Grosse 2008). First, we performed sensitivity analysis of the cost-effectiveness of full coverage for each patient class for different medication c and CV event $\mathbf{e}$ costs. We performed two-way sensitivity analysis of $y_p = 1$ under $c \pm 100\%$ and $\mathbf{e} \pm 100\%$ (no cost to double cost) and plotted the cost-effectiveness at $50,000 per QALY in Figure 6. At full coverage ($y_p = 1$), we observe that no patient class is cost-effective at current costs. Therefore, the costs must be subsidized by reducing coverage for other classes. Additionally, the cost-effectiveness varies

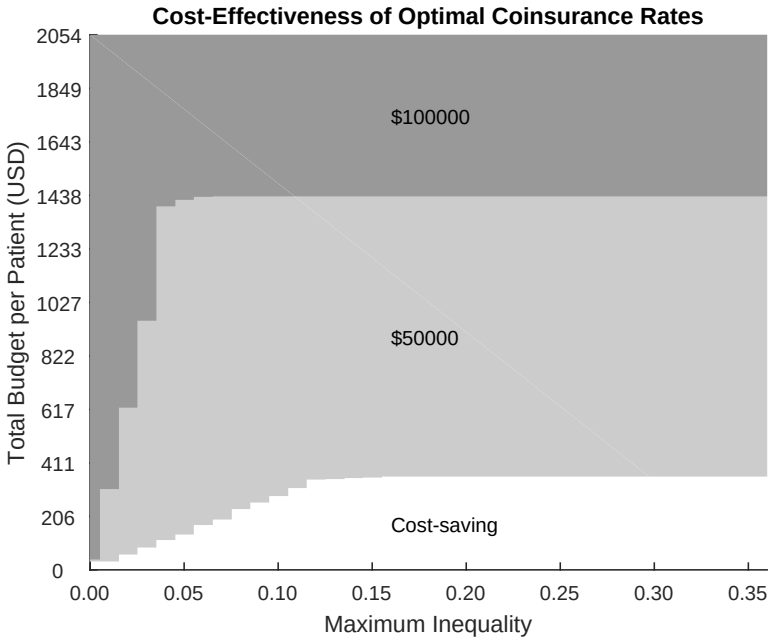


Figure 5: Cost-effectiveness of BLP policy along different cost-effectiveness thresholds (ICER $\leq \$50,000 - \$200,000$ per QALY) at each budget and maximum inequality threshold. Each shade indicates the smallest cost-effectiveness threshold at which a specific parameter combination is cost-effective. Cost-saving implies that QALYs are gained at no additional cost.

greatly based on CV risk. For example, at current event costs, full coverage becomes cost-effective for the high risk patient classes when medication costs decrease by 40%. However, for low risk patients, full coverage becomes cost-effective if medication costs decrease by 80%. For fixed medication costs, the high risk patient classes become more cost-effective as event costs increase since the increase in adherence due to free medications leads to improved health outcomes.

Finally, we performed sensitivity analysis of the cost-effectiveness of full coverage for each patient class for different price elasticities of adherence δ_p^α and current Medicare coverage y_p^i. We performed two-way sensitivity analysis of $y_p = 1$ under $\delta_p^\alpha \in [0, 1]$ and $y_p^i \in [0, 100\%]$ and plotted the cost-effectiveness at \$50,000 per QALY in Figure 7. We find that full coverage becomes cost-effective for low risk patients when the price elasticity of adherence is 0.83 at Medicare's current 75% coverage rate. For high risk patients, full coverage becomes cost-ineffective when the price elasticity drops below 0.2.

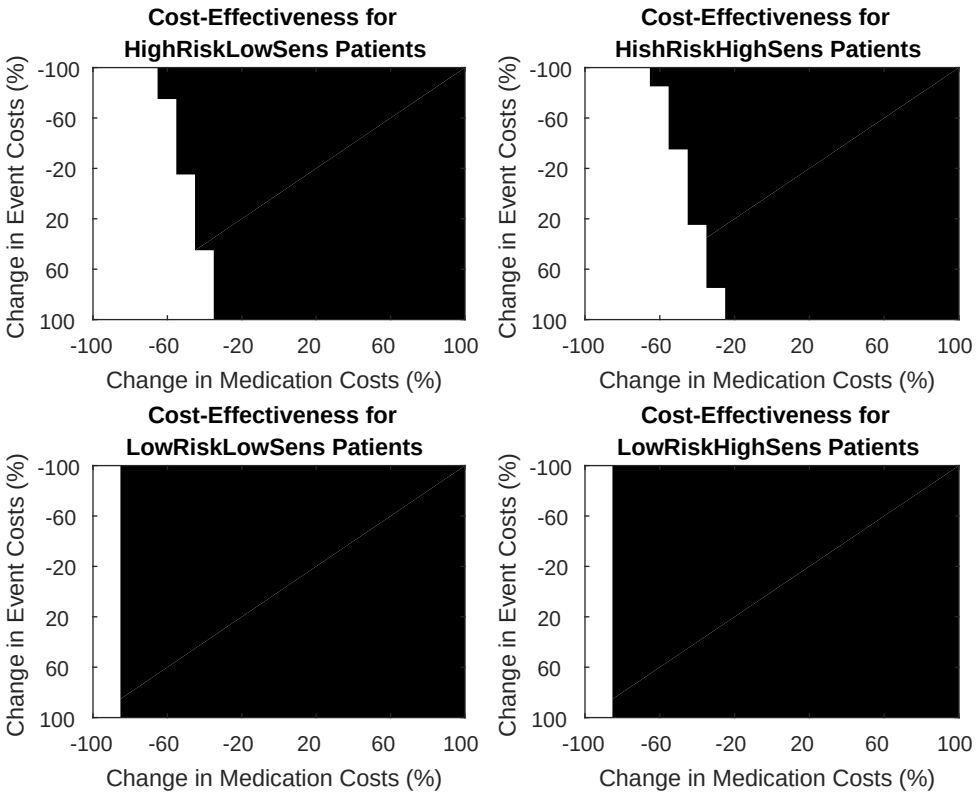


Figure 6: Cost-effectiveness of full coverage for each patient class (ICER $\leq$ \$50,000 per QALY) at each level of medication cost and CV event cost. White regions indicate cost-effective coverage.

5.5 Additional Applications

Much like anti-hypertension drugs, cholesterol-lowering medications are prescribed to reduce the risk of CVD events. However, compared to hypertension drugs, cholesterol-lowering medications are generally more expensive on a per-pill basis. We performed analysis on two additional applications under the same Medicare setting. In the first application, we considered medium-intensity and high-intensity statins as preventative treatment options. The second application is identical to the first except we included the PCSK9 inhibitor as an additional treatment option. The PCSK9 inhibitor has been found to be a very effective treatment option to lower cholesterol, but it comes at a significantly higher price than statins (Kazi et al. 2016). The model formulation to determine the optimal cholesterol-lowering preventative treatment policy was identical to the formulation in Section 3.1 except the action space consisted of the type of medication prescribed instead of the number of medications. Furthermore, the type of medication prescribed was not constrained by $b_{min}(\mathbf{s}_p)$. The parameters

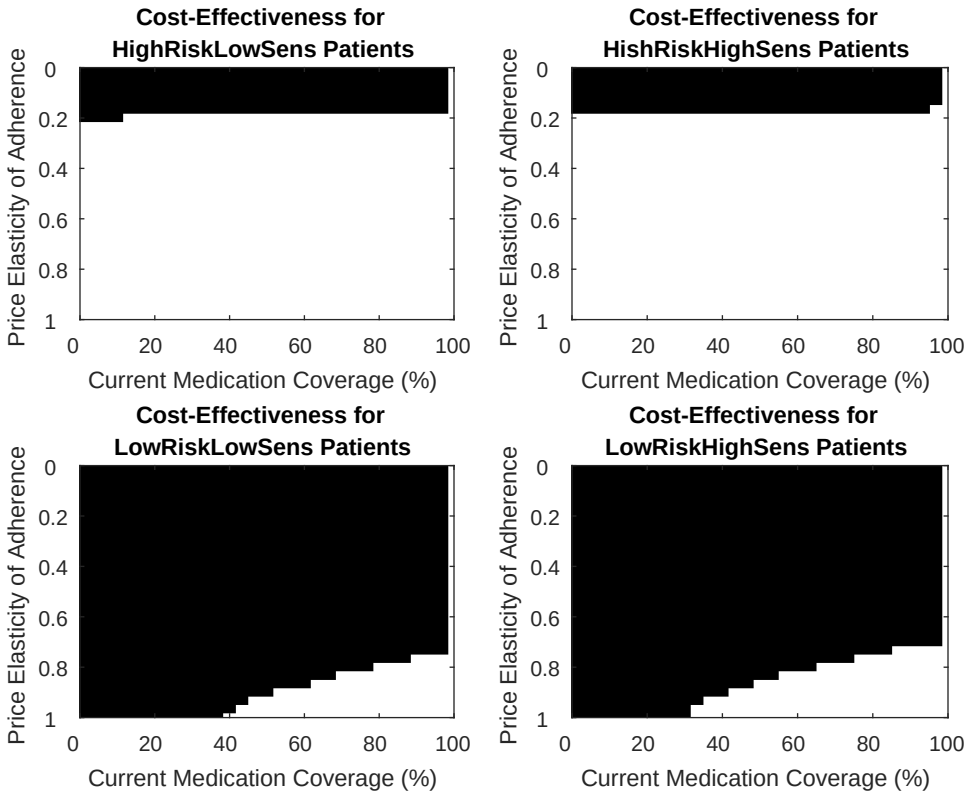
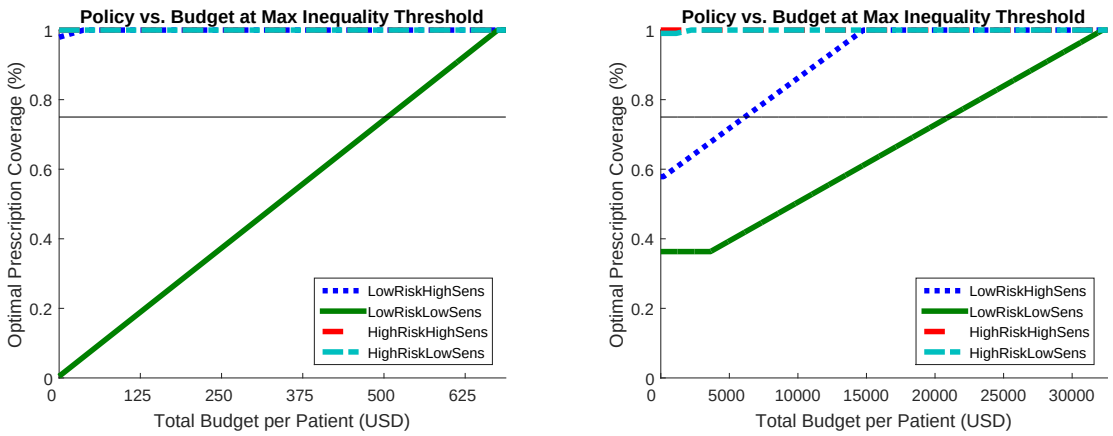


Figure 7: Cost-effectiveness of full coverage for each patient class (ICER $\leq$ \$50,000 per QALY) at each level of price elasticity and current coverage. White regions indicate cost-effective coverage.

for both of these applications are provided in Appendix A.

Figure 8(a) and Figure 8(b) show the optimal coverage policy when there is no constraint on inequality for the first application (i.e., medium- and high-intensity statins) and second application (i.e., statins and PCSK9 inhibitor), respectively. The optimal coverage policy for both applications is similar to that in Section 5.1. That is, it is also optimal to prioritize first by CV risk then by cost sensitivity. For the first application, it is optimal to provide little to no coverage for low risk, low sensitivity patients when no additional budget is given. Such a policy may be infeasible to implement in practice and suggests the importance of the coverage inequality constraint. Additionally, due to the price of PCSK9 inhibitors, the required budget increase to ensure full medication for all patient classes in the second application is more than 20 times the additional budget needed when only statins are considered. The magnitude of this difference raises an important question about whether the gain in QALYs is worth the cost.



(a) Medium-intensity statins and high-intensity statins (b) Medium-intensity statins, high-intensity statins, and PCSK-9 inhibitor

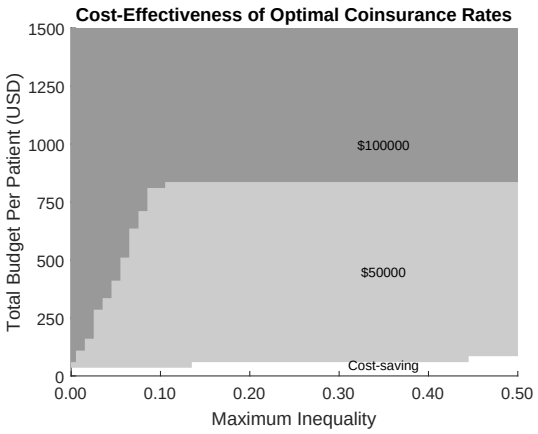
Figure 8: Optimal coverage decisions for cholesterol-lowering medications at different budget levels assuming no constraint on inequality. Current Medicare coverage (75%) is denoted by the horizontal bar.

The cost-effectiveness of optimal coverage policies for the first application and second application are shown in Figure 9(a) and Figure 9(b), respectively. In both additional applications, the cost-effectiveness of the optimal coverage policies bears similarities to the cost-effectiveness results in Section 5.4 — the optimal coverage policy becomes less cost-effective as the budget increases and more cost-effective as the maximum inequality threshold increases. Compared to the results in Section 5.4 and the application which only considers statins, the application with PCSK9 can tolerate a higher total budget increase before it becomes cost-ineffective. In fact, at modest levels of maximum allowable inequality (e.g., $w = 0.15$), the application with PCSK9 remains cost-saving even at higher budget levels. However, relative to the budget required to achieve full coverage for all patient classes, the application for PCSK9 becomes cost-ineffective much more quickly compared to the other applications, even at a cost-effectiveness threshold of $200,000 per QALY.

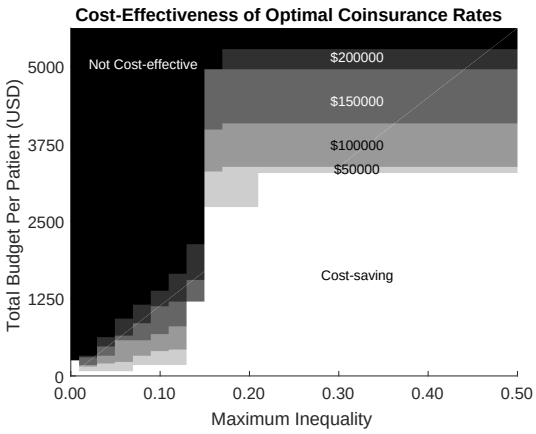
6 Managerial Insights

From the analytical and numerical results, we have found the following managerial insights:

Prioritize First on CV Risk, Then On Cost Sensitivity. Our analytical and numerical findings revealed that patients at higher risk for cardiovascular disease events should be prioritized for coinsurance rate reduction due to their higher marginal health benefit of in-



(a) Medium-intensity and high-intensity statins



(b) Medium-intensity and high-intensity statins and PCSK-9 inhibitor

Figure 9: Cost-effectiveness of BLP policy for cholesterol-lowering medications along different cost-effectiveness thresholds (ICER $\leq \$50,000 - \$200,000$ per QALY) at each budget and maximum inequality threshold. Each shade indicates the smallest cost-effectiveness threshold at which a specific parameter combination is cost-effective. Cost-saving implies that QALYs are gained at no additional cost and Not Cost-effective implies that ICER $\geq \$200,000$.

creased adherence. When applied to both hypertension medication and cholesterol-lowering medication, we found that the utilitarian policy (i.e. a policy without concern for equality of coverage) increases coinsurance rates for lower risk patients in order to cover the cost of free medications for higher risk patients, unless the budget is sufficiently high. However, as the budget increases, coverage is increased for the lower risk patients since any reduction in coinsurance rate is associated with an increase in expected QALYs. The improvement in QALYs observed from the numerical study suggests that adherence to medication is a vitally important issue in managing cardiovascular disease. The specific improvement in health (13.6 to 24.9 QALYs per 1000 patients) in our numerical study of hypertension medication is a reflection of the price elasticity of adherence δ_p^α found in the health economics literature and the current 75% coverage. If patients were more sensitive to the price of medications and current coverage were lower, the health gains would increase accordingly. The importance of price elasticity is also evident in the optimal BLP policy: within each risk strata, prioritize patients with higher price elasticities.

Higher Budgets May Decrease Cost-Effectiveness. If Congress allocates additional funding to Medicare, the coinsurance rates of the Medicare population will decrease and the

expected QALYs saved will increase. Yet, these additional resources are not guaranteed to lead to cost-effective coinsurance rates. The additional funds needed to provide free medication to high risk patients is cost-effective when coinsurance rates for low risk patients are sufficiently increased. However, budget increases to decrease coinsurance rates for low risk patients reduces this cost-effectiveness.

Cost-Effectiveness Requires Personalized Coinsurance Rates. As the maximum inequality threshold decreases, the restriction on tailoring coinsurance rates based on the marginal health benefit of the particular patient class leads to an inefficient allocation of resources and cost-ineffective policies. To further understand how the maximum inequality threshold affects population health performance, we computed the health cost of equality and found the cost to be substantial. The health cost of equality further illustrates the benefit of being able to tailor coinsurance rates to the characteristics of each patient class, rather than treating all patients as monolithic with the same marginal health benefits. We found that allowing small deviations from equal coverage can lead to significant health gains.

Adherence to Preventative Medication Should be Improved. We found that increasing adherence to anti-hypertension and cholesterol-lowering medication resulted in improved population health and potentially cost-saving coinsurance rates. While our analysis has been with respect to coinsurance rates for these specific medications, the findings are applicable to other cardiovasuclar disease (preventative) treatments and other chronic diseases where adherence to medication is low. For example, adherence improving interventions for HIV medication can be modeled in a similar fashion to the BLP model described in Section 3. The interventions should be targeted to those patients who have the greatest marginal health benefit of improved adherence and those patients whose adherence will improve the most from the interventions. As observed from our sensitivity analyses, the cost-effectiveness of the intervention will depend on the elasticity of adherence, the cost of medications, the cost of adverse health events, the current coverage rate, and the change in QALYs caused by improvements in adherence. And similarly to cardiovascular disease, the

health cost of equality will depend on the difference in marginal health benefits across the set of patient classes. As heterogeneity in risk for HIV complications and progression within the patient population increases, the cost of equality will also increase.

7 Study Limitations and Future Research

Ideally, treatment plans and coinsurance rates would be determined at the individual patient level. However, implementing such policies would be administratively challenging and burdensome. Thus, in our numerical analysis, we considered a case-study of four patient classes based on combinations of risk and cost sensitivity. While our analysis using four classes captures the important prioritization scheme of the optimal coverage policy, the full Medicare population is comprised of a richer set of classes with more detailed differences in risk factors (e.g. diabetes status, smoking status, sex, and race) and adherence-influencing characteristics (e.g. education and income). Further stratification of the four patient classes would allow for additional tailoring of the optimal coverage policy, but would require more data for price elasticities and additional administrative oversight for implementation in practice. However, our theoretical and numerical findings provide the key insights necessary for the design of optimal coverage policies, even under a strict requirement on equality of coverage.

In our numerical study, we assumed that adherence to medication as a function of coinsurance rates was reflected by (18). However, we understand that the linear form of this function may oversimplify the relationship between adherence and medication cost. Thus, future formulations of this model may consider other functions to model this relationship. For instance, in Gibson et al. (2006a), the authors developed a logistic regression model to describe a patient's adherence to statins as a function of his or her demographic information, income type, cardiovascular risk, and statin copayment.

When computing ICERs, we considered changes in total costs based on medication and CV event costs. Other costs associated with CV disease include hospital visits and emergency department (ED) visits. Research indicates that as adherence to anti-hypertension medication improves, hospital visits and ED visits decrease (Choudhry et al. 2008). This

finding suggests that our analysis conservatively estimates the cost-effectiveness of the BLP policy. Future research may investigate how cost savings in hospital and ED visits can be utilized when determining medication coinsurance rates and when computing the ICER.

In our analysis, we restricted our attention to financial mechanisms for improving adherence in the form of medication coinsurance rate reductions. However, the insights derived from analyzing coinsurance rates also affect the optimal design of other adherence improving programs. Such programs include pharmacist intervention (Santschi et al. 2012), cultural education (Beune et al. 2014), and registering patients with a text-message system to remind them to take their medication (Ho et al. 2014). Our modeling framework is amenable to extending the decision space to include a portfolio of mechanisms (e.g. $\mathbf{y}$ for coinsurance, $\mathbf{x}$ for text-messaging) where Medicare can select the intensity of the mechanism (e.g. the number of text messages to send) for each patient class. The impact of the portfolio of mechanisms on the patient's adherence can then be modeled using correlated elasticities of adherence for each mechanism.

Lastly, we assumed perfect information on the part of Medicare. In practice, Medicare may not know exactly how patients will react to changes in copayments, both in terms of the new adherence rate and the resulting optimal preventative treatment policy (or the degree to which physicians prescribe medications optimally). Future work in this area may include modeling the price elasticity of adherence as an unknown variable within an uncertainty set. Such a formulation lends itself to solutions via robust optimization paired with BLP techniques.

8 Conclusion

In conclusion, we have developed a bilevel optimization model to determine coinsurance rates for individual patient classes within a heterogeneous patient population. We parameterized and solved this model by considering patients who are insured by Medicare and are prescribed hypertension or cholesterol-lowering medications for prevention of cardiovascular disease. Using the theory of price elasticity, we have modeled the relationship between coin-

surance rates and the patient's adherence to medication. Each patient class is unique in its cardiovascular risk characteristics and cost sensitivity, and receives a different clinical benefit from improved adherence to medication. With the knowledge obtained from studying the optimal and personalized preventative treatment policies generated by the patient-level MDP for hypertension and cholesterol treatment planning, we developed a dynamic programming approach with a penalty function to derive optimal coinsurance rates for different patient classes within the population. The class-specific marginal health benefit of adherence informs the model's determination of the coinsurance rate that should be paid when Medicare faces a budgetary constraint, as well as a constraint on the total coinsurance inequality caused by the optimal coverage policy. Overall, we find that reducing barriers to high clinical value services can be beneficial to population health and the insurance provider's finances, depending on the budget available and the required level of coinsurance fairness. While our work is based upon Medicare coinsurance rates for anti-hypertension medication and cholesterol-lowering treatments, we believe the insights and modeling framework are applicable to other prevalent chronic diseases, adherence improving interventions, and private insurance providers.

9 Acknowledgments

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE1256260. Dr. Lavieri would like to acknowledge NSF CAREER (CMMI-1552545) for their funding support. Dr. Sussman was supported by the Department of Veterans Affairs CDA 13-021 and IIR 15-432. We are grateful to the anonymous referees and the Editors for their helpful comments and suggestions.

10 Notes on Contributors

Greggory J. Schell is a Research Scientist in Data Science at the Center for Naval Analyses. He holds a Bachelor's in Industrial Engineering from the University of Pittsburgh, a Master's in Statistics and Industrial and Operations Engineering from the University of Michigan, and his Ph.D. in Industrial and Operations Engineering. His doctoral research focused on optimal

sequential medical decision making for glaucoma and cardiovascular disease. His research has won the Medical Decision Making Lee Lusted Award, the INFORMS Doing Good with Good OR Award, and the IBM Research Service Science Best Student Paper Award. His primary research interests are in machine learning and optimization to derive data-driven policy recommendations to improve the Navy's readiness across its manning, training, and equip functions.

Gian-Gabriel P. Garcia is a Ph.D. Candidate in the Industrial and Operations Engineering Department at the University of Michigan. He holds a Bachelor's degree in Industrial Engineering from the University of Pittsburgh and a Master's degree in Industrial and Operations Engineering from the University of Michigan. His primary research interest is in improving medical decision making through the development and analysis of models which incorporate optimization under uncertainty, stochastic modeling, game theory, and predictive modeling. His most recent work includes applications to concussion, glaucoma, and cardiovascular disease. Gian is the recipient of the National Science Foundation Graduate Research Fellowship, Rackham Merit Fellowship, and first prize at the INFORMS Minority Issues Forum Poster Competition. He has also received honorable mention for the Ford Foundation Pre-doctoral Fellowship.

Mariel S. Lavieri is an Associate Professor in the Industrial and Operations Engineering Department at the University of Michigan. She has Bachelor's degrees in Industrial and Systems Engineering and Statistics and a minor in String Bass Performance from the University of Florida. She holds a Master's and Ph.D. in Management Science from the University of British Columbia. In her work, she applies operations research to healthcare topics. Among others, she has developed dynamic programming, stochastic control, and continuous, partially observable state space models to guide screening, monitoring, and treatment decisions of chronic disease patients. She has also created models for health workforce and capacity planning. She is the recipient of the Willie Hobbs Moore Aspire, Advance, Achieve Mentoring Award, the National Science Foundation CAREER Award, the International Conference

on Operations Research Young Participant with Most Practical Impact Award, and the Bonder Scholarship. She has also received the Pierskalla Best Paper Award, and an honorary mention in the George B. Dantzig Dissertation Award. She has guided work that won the Medical Decision Making Lee Lusted Award, the INFORMS Doing Good with Good OR Award, the IBM Research Service Science Best Student Paper Award and the Production and Operations Management Society College of Healthcare Operations Management Best Paper Award.

Jeremy B. Sussman is an Assistant Professor in Internal Medicine at the University of Michigan and a research scientist in the Center for Clinical Management Research at VA Ann Arbor Healthcare System, where he is also a primary care physician. He has a Bachelor's degree in Biology from Amherst College, an MD from the University of California, San Francisco, and Master's degrees from the University of California, Berkeley, and the University of Michigan as part of the Robert Wood Johnson Clinical Scholars Program. He performed an internship and residency in Internal Medicine at Yale-New-Haven Hospital. His primary research interests are in improving the personal tailoring of medical decisions, particularly in cardiovascular disease and diabetes. His work uses risk prediction, decision analysis, and implementation science. He has been published in top medical journals, including JAMA Internal Medicine, Circulation, and BMJ.

Rodney A. Hayward is a Professor of Public Health and Internal Medicine at the University of Michigan and Co-Director of the Center for Practice Management and Outcomes Research at the Ann Arbor VA HSR&D. He received his training in health services research as a Robert Wood Johnson Clinical Scholar at UCLA and at the RAND Corporation, Santa Monica. His current and past work includes studies examining measurement of quality, costs and health status, environmental and educational factors affecting physician practice patterns, quality improvement, and physician decision making. His current work focuses on quality measurement and improvement for chronic diseases, such as diabetes, hypertension and heart disease.

References

Original Medicare (Part A and B) Eligiblity and Enrollment. https://www.cms.gov/Medicare/Eligibility-and-Enrollment/OrigMedicarePartABEligEnrol/index.htm. Accessed: 2018-02-27.

Akan, M., Alagoz, O., Ata, B., Erenay, F. S., and Said, A. (2012). A Broader View of Designing the Liver Allocation System. *Operations Research*, 60(4):757–770.

Anderson, K. M., Odell, P. M., Wilson, P. W., and Kannel, W. B. (1991). Cardiovascular disease risk profiles. *American Heart Journal*, 121(1):293–298.

Arias, E. (2007). United states life tables, 2004. *National Vital Statistics Reports*, 56(9):1–40.

Ayer, T., Alagoz, O., and Stout, N. K. (2012). Pomdp approach to personalize mammography screening decisions. *Operations Research*, 60(5):1019–1034.

Ayer, T., Alagoz, O., Stout, N. K., and Burnside, E. S. (2016). Heterogeneity in Women's Adherence and Its Role in Optimal Breast Cancer Screening Policies. *Management Science*, 62(5):1339–1362.

Bard, J. F. (1998). *Practical bilevel optimization: algorithms and applications*, volume 30. Springer Science & Business Media.

Beune, E. J., van Charante, E. P. M., Beem, L., Mohrs, J., Agyemang, C. O., Ogedegbe, G., and Haafkens, J. A. (2014). Culturally adapted hypertension education (cahe) to improve blood pressure control and treatment adherence in patients of african origin with uncontrolled hypertension: Cluster-randomized trial. *PloS one*, 9(3):e90103.

Bibbins-Domingo, K., Chertow, G. M., Coxson, P. G., Moran, A., Lightwood, J. M., Pletcher, M. J., and Goldman, L. (2010). Projected Effect of Dietary Salt Reductions on Future Cardiovascular Disease. *New England Journal of Medicine*, 362(7):590–599.

Bunting, B. A., Smith, B. H., and Sutherland, S. E. (2008). The asheville project: clinical and economic outcomes of a community-based long-term medication therapy management program for hypertension and dyslipidemia. *Journal of the American Pharmacists Association*, 48(1):23–31.

Calcott, P. (1999). Demand inducement as cheap talk. *Health Economics*, 8(8):721–733.

Chernew, M. E., Juster, I. A., Shah, M., Wegh, A., Rosenberg, S., Rosen, A. B., Sokol, M. C., Yu-Isenberg, K., and Fendrick, A. M. (2010). Evidence that value-based insurance can be effective. *Health Affairs*, 29(3):530–536.

Choudhry, N. K., Fischer, M. A., Avorn, J. L., Lee, J. L., Schneeweiss, S., Solomon, D. H., Berman, C., Jan, S., Lii, J., Mahoney, J. J., et al. (2012). The impact of reducing cardiovascular medication copayments on health spending and resource utilization. *Journal of the American College of Cardiology*, 60(18):1817–1824.

Choudhry, N. K., Patrick, A. R., Antman, E. M., Avorn, J., and Shrank, W. H. (2008). Cost-effectiveness of providing full drug coverage to increase medication adherence in post–myocardial infarction medicare beneficiaries. *Circulation*, 117(10):1261–1268.

Cowell, F. A. (2000). Measurement of inequality. In *Handbook of Income Distribution*, volume 1, chapter 2, pages 87–166.

De Jaegher, K. and Jegers, M. (2001). The physician-patient relationship as a game of strategic information transmission. *Health Economics*, 10(7):651–668.

Denoyel, V., Alfandari, L., and Thiele, A. (2017). Optimizing healthcare network design under reference pricing and parameter uncertainty. *European Journal of Operational Research*, 263(3):996–1006.

Denton, B. T., Alagoz, O., Holder, A., and Lee, E. K. (2011). Medical decision making: open research challenges. *IIE Transactions on Healthcare Systems Engineering*, 1(3):161–167.

Egan, B. M., Lackland, D. T., and Cutler, N. E. (2003). Awareness, knowledge, and attitudes of older americans about high blood pressure: implications for health care policy, education, and research. *Archives of Internal Medicine*, 163(6):681–687.

El-Amine, H., Bish, E. K., and Bish, D. R. (2018). Robust Postdonation Blood Screening Under Prevalence Rate Uncertainty. *Operations Research*, 66(1):1–17.

Erenay, F. S., Alagoz, O., and Said, A. (2014). Optimizing Colonoscopy Screening for Col-

orectal Cancer Prevention and Surveillance. *Manufacturing & Service Operations Management*, 16(3):381–400.

Gibson, T. B., Mahoney, J., Ranghell, K., Cherney, B. J., and McElwee, N. (2011). Value-based insurance plus disease management increased medication use and produced savings. *Health Affairs*, 30(1):100–108.

Gibson, T. B., Mark, T. L., Axelsen, K., Baser, O., Rublee, D. A., and McGuigan, K. A. (2006a). Impact of statin copayments on adherence and medical care utilization and expenditures. *The American Journal of Managed Care*, 12(special issue):SP11–SP19.

Gibson, T. B., Mark, T. L., McGuigan, K. A., Axelsen, K., and Wang, S. (2006b). The Effects of Prescription Drug Copayments on Statin Adherence. *The American Journal of Managed Care*, 12(9):509–517.

Goldman, D. P., Joyce, Geoffrey, F., and Karaca-Mandic, P. (2006). Varying Pharmacy Benefits With Clinical Status: The Case of Cholesterol-lowering Therapy. *American Journal of Managed Care*, 12(1):21–28.

Grosse, S. D. (2008). Assessing cost-effectiveness in healthcare: history of the $50,000 per QALY threshold. *Expert Review of Pharmacoeconomics & Outcomes Research*, 8(2):165–178.

Ho, P. M., Lambert-Kerzner, A., Carey, E. P., Fahdi, I. E., Bryson, C. L., Melnyk, S. D., Bosworth, H. B., Radcliff, T., Davis, R., Mun, H., et al. (2014). Multifaceted intervention to improve medication adherence and secondary prevention measures after acute coronary syndrome hospital discharge: a randomized clinical trial. *JAMA internal medicine*, 174(2):186–193.

Ho, T.-Y., Fishman, P. A., and Zabinsky, Z. B. (2016). Designing And Analyzing Healthcare Insurance Policies To Reduce Cost and Prevent The Spread of Seasonal Influenza. In *Proceedings of the 2016 Winter Simulation Conference*, pages 2018–2029.

Ibrahim, R., Kucukyazici, B., Verter, V., Gendreau, M., and Blostein, M. (2016). Designing

Personalized Treatment: An Application to Anticoagulation Therapy. *Production and Operations Management*, 25(5):902–918.

Jiang, H., Pang, Z., and Savin, S. (2012). Performance-based contracts for outpatient medical services. *Manufacturing & Service Operations Management*, 14(4):654–669.

Kazemian, P., Helm, J. E., Lavieri, M. S., Stein, J. D., and Van Oyen, M. P. (2018). Dynamic Monitoring and Control of Irreversible Chronic Diseases With Application to Glaucoma. (Under review).

Kazi, D. S., Moran, A. E., Coxson, P. G., Penko, J., Ollendorf, D. A., Pearson, S. D., Tice, J. A., Guzman, D., and Bibbins-Domingo, K. (2016). Cost-effectiveness of PCSK9 Inhibitor Therapy in Patients With Heterozygous Familial Hypercholesterolemia or Atherosclerotic Cardiovascular Disease. *Journal of the American Medical Association*, 316(7):743–753.

Khojandi, A., Maillart, L. M., Prokopyev, O. A., Roberts, M. S., and Saba, S. F. (2018). Dynamic Abandon/Extract Decisions for Failed Cardiac Leads. *Management Science*, 64(2):633–651.

Klees, B., Wolfe, C., and Curtis, C. (2010). Brief summaries of medicare and medicaid: Title xviii & title xix of the social security act as of november 1, 2010. *Retrieved May*, 26:2011.

Koek, H. L., de Bruin, A., Gast, F., Gevers, E., Kardaun, J. W., Reitsma, J. B., Grobbee, D. E., and Bots, M. L. (2006). Short-and long-term prognosis after acute myocardial infarction in men versus women. *The American journal of cardiology*, 98(8):993–999.

Kovacevic, R. M. and Pflug, G. C. (2014). Electricity swing option pricing by stochastic bilevel optimization: A survey and new approaches. *European Journal of Operational Research*, 237(2):389–403.

Law, M., Morris, J., and Wald, N. (2009). Use of blood pressure lowering drugs in the prevention of cardiovascular disease: meta-analysis of 147 randomised trials in the context of expectations from prospective epidemiological studies. *British Medical Journal*, 338:b1665.

Lee, D. K. and Zenios, S. A. (2012). An evidence-based incentive system for medicare's end-stage renal disease program. *Management Science*, 58(6):1092–1105.

Lee, E., Lavieri, M. S., Volk, M. L., and Xu, Y. (2015). Applying reinforcement learning techniques to detect hepatocellular carcinoma under limited screening capacity. *Health Care Management Science*, 18(3):363–375.

Ma, B. C.-t. A. and Mcguire, T. G. (1997). Optimal Health Insurance and Provider Payment. *The American Economic Review*, 87(4):685–704.

Mann, B. S., Barnieh, L., Tang, K., Campbell, D. J., Clement, F., Hemmelgarn, B., Tonelli, M., Lorenzetti, D., and Manns, B. J. (2014). Association between drug insurance cost sharing strategies and outcomes in patients with chronic diseases: A systematic review. *PloS one*, 9(3):e89168.

Manning, W. G., Newhouse, J. P., Duan, N., Keeler, E. B., and Leibowitz, A. (1987). Health insurance and the demand for medical care: evidence from a randomized experiment. *The American economic review*, pages 251–277.

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

Mason, J. E., England, D. A., Denton, B. T., Smith, S. A., Kurt, M., and Shah, N. D. (2012). Optimizing statin treatment decisions for diabetes patients in the presence of uncertain future adherence. *Medical Decision Making*, 32(1):154–166.

McCain, J. (2016). Fiscal Year 2017 National Defense Authorization Bill.

Medicare Payment Advisory Commission and others (2014). Report to the congress: Medicare and the health care delivery system. 2013.

Mehrotra, S. and Kim, K. (2011). Outcome Based State Budget Allocation for Diabetes Prevention Programs Using Multi-criteria Optimization with Robust Weights. *Health Care Management Science*, 14(4):324–337.

Mills, A. F., Argon, N. T., and Ziya, S. (2013). Resource-Based Patient Prioritization in Mass-Casualty Incidents. *Manufacturing & Service Operations Management*, 15(3):361–377.

Moran, A. E., Odden, M. C., Thanataveerat, A., Tzong, K. Y., Rasmussen, P. W., Guzman, D., Williams, L., Bibbins-Domingo, K., Coxson, P. G., and Goldman, L. (2015). Cost-Effectiveness of Hypertension Therapy According to 2014 Guidelines. *New England Journal of Medicine*, 372(5):447–455.

Mussard, S., Seyte, F., and Terraza, M. (2003). Decomposition of gini and the generalized entropy inequality measures. *Economics Bulletin*, 4(7):1–6.

Neumann, P. J., Cohen, J. T., and Weinstein, M. C. (2014). Updating Cost-Effectiveness — The Curious Resilience of the $50,000-per-QALY Threshold. *New England Journal of Medicine*, 371(9):796–797.

Sahney, V. K. and Weddle, T. (1980). A Model for Sensitivity Analysis of Reimbursement in Hospitals. *A I I E Transactions*, 12(2):133–139.

Santschi, V., Chiolero, A., Paradis, G., Colosimo, A. L., and Burnand, B. (2012). Pharmacist interventions to improve cardiovascular disease risk factors in diabetes a systematic review and meta-analysis of randomized controlled trials. *Diabetes care*, 35(12):2706–2717.

Schell, G. J. (2015). *Personalized Medicine in Chronic Disease Managment*. PhD thesis, University of Michigan.

Schell, G. J., Marrero, W. J., Lavieri, M. S., Sussman, J. B., and Hayward, R. A. (2016). Data-Driven Markov Decision Process Approximations for Personalized Hypertension Treatment Planning. *MDM Policy & Practice*, 1(1):238146831667421.

Schottmüller, C. (2013). Cost incentives for doctors: A double-edged sword. *European Economic Review*, 61:43–58.

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

Skandari, M. R., Shechter, S. M., and Zalunardo, N. (2015). Optimal Vascular Access Choice for Patients on Hemodialysis. *Manufacturing & Service Operations Management*, 17(4):608–619.

Sussman, J., Vijan, S., and Hayward, R. (2013). Using benefit-based tailored treatment to improve the use of antihypertensive medications. *Circulation*, 128(21):2309–2317.

Sussman, J. B., Vijan, S., Choi, H., and Hayward, R. A. (2011). Individual and Population Benefits of Daily Aspirin Therapy: A Proposal for Personalizing National Guidelines. *Circulation: Cardiovascular Quality and Outcomes*, 4(3):268–275.

Tamblyn, R., Reidel, K., Huang, A., Taylor, L., Winslade, N., Bartlett, G., Grad, R., Jacques, A., Dawes, M., Larochelle, P., et al. (2010). Increasing the detection and response to adherence problems with cardiovascular medication in primary care through computerized drug management systems: a randomized controlled trial. *Medical Decision Making*, 30(2):176–188.

Timbie, J. W., Hayward, R. A., and Vijan, S. (2010). Variation in the net benefit of aggressive cardiovascular risk factor control across the us population of patients with diabetes mellitus. *Archives of internal medicine*, 170(12):1037–1044.

Weinstein, M. C., Coxson, P. G., Williams, L. W., Pass, T. M., Stason, W. B., and Goldman, L. (1987). Forecasting coronary heart disease incidence, mortality, and cost: The coronary heart disease policy model. *American Journal of Public Health*, 77(11):1417–1426.

Yaesoubi, R. and Roberts, S. D. (2011). Payment contracts in a preventive health care system: A perspective from operations management. *Journal of health economics*, 30(6):1188–1196.

Zaric, G. S., Zhang, H., and Mahjoub, R. (2013). Modeling risk sharing agreements and patient access schemes. In *Operations Research and Health Care Policy*, pages 295–310. Springer.

Zepeda, E. D. and Sinha, K. K. (2016). Toward an effective design of behavioral health care delivery: An empirical analysis of care for depression. *Production and Operations Management*, 25(5):952–967.

Zhang, J., Denton, B. T., Balasubramanian, H., Shah, N. D., and Inman, B. A. (2012a).

Optimization of Prostate Biopsy Referral Decisions. *Manufacturing & Service Operations Management*, 14(4):529–547.

Zhang, J., Denton, B. T., Balasubramanian, H., Shah, N. D., and Inman, B. A. (2012b). Optimization of PSA Screening Policies. *Medical Decision Making*, 32(2):337–349.

Zhang, Y., Berman, O., Marcotte, P., and Verter, V. (2010). A bilevel model for preventive healthcare facility network design with congestion. *IIE Transactions*, 42(12):865–880.