Optimal Health Expenditures under Uncertainty in Patient Numbers *

Dilan Su Alpergin †

Abstract

Healthcare providers face unexpected changes in patient numbers as diseases and injuries are unpredictable. This paper studies socially optimal health expenditures and treatments by drawing attention to this quantity uncertainty in healthcare markets. This paper is the first to approach this problem from the norms of welfare economics. I first introduce a model that characterizes the socially optimal setting, then provide a model for healthcare providers and compare the outcomes from the two models. The results show that healthcare providers' investments in treatment resources differ from socially optimal levels. Based on this finding, I argue that healthcare providers fail to bring socially optimal treatment to society when there are unexpected surges in hospital patient flows.

1 Introduction

One of the fundamental problems hospitals face is the uncertainty over demand for medical services. Patient admission rates are unexpected as diseases and injuries are unpredictable. The existence of this uncertainty in healthcare markets has important consequences both on healthcare providers and societies. How societies deal with the problem of uncertainty in patient numbers requires studying the socially optimal treatment and health expenditures.

*I am grateful to Steven Slutsky for his helpful comments and guidance along the way. I would also like to thank Perihan Saygin, Hector Sandoval, and W. Bruce Vogel for their support. Finally, I thank Mark Rush and Fatma Gunay for their comments and suggestions.

†Department of Economics, University of Florida, Gainesville, USA

This paper provides a model for a stylized society in which there is an illness that brings different numbers of ill individuals to society. Due to this quantity uncertainty, society must make some decisions in advance and others after learning about the number of patients. Real-world examples show that societies under uncertainty make decisions in this way. For instance, the recent COVID-19 outbreak required governments and healthcare providers to make many short-term decisions to manage the unexpected increase in patient demand. Hospitals, for example, made short-term alterations to patient areas to accommodate caring for an increased number of patients. In addition, they increased the number of short-term travel nurses to fill the staff gaps. The recent COVID-19 pandemic was a good example of understanding the importance of uncertainties in healthcare markets. Therefore, I believe analyzing how societies under uncertainty optimally cope with health care problems is fundamental. To this end, this paper studies the socially optimal treatment and health care expenditures under uncertainty in patient numbers.

Some variations in hospital admission rates, such as weekday and weekend inflows or seasonal fluctuations, are regular and predictable. However, some variations are stochastic, as diseases and injuries are unpredictable. Because of this random demand component, hospitals need to choose some inputs before the realization of demand. In contrast, they must adjust some other inputs ex-post. Therefore, the literature suggests two types of factors of production for hospitals; fixed and variable. Fixed inputs are immobile for long periods, such as the number of beds. However, variable inputs are adjustable in response to demand in the short run, such as the number of labor and supplies. The stochastic component of hospital admission rates has several implications for hospitals. For instance, it changes the structure of hospital costs. Gaynor and Anderson (1995), estimating a translog cost function for hospitals facing uncertain demand, show that the uncertainty in patient admission rates affects hospitals' input choices because hospitals hold standby capacity to avoid turning patients away during unexpected surges in demand. Therefore, when there is higher uncertainty, they face higher costs due to attempts to hedge more capacity. Keeler and Ying (1996) also assume that hospitals hold standby capacity. The authors, analyzing a family of short-run cost functions, find that standby capacity brings inefficiencies into healthcare markets. According to the authors, technological advances and increased competition in healthcare markets created excess capacity in hospitals by decreasing capital utilization rates. They argue that excess capacity is a real cost burden for hospitals, and one of the ways to reduce these costs is to find a way to bring occupancy rates to optimal levels.

Stigler (1939) studies production and distribution theory in the short run. He defines the short run as the period in which there are both fixed and variable costs. He argues that entrepreneurs face an ongoing problem of maximizing net returns due to uncertain input price and demand conditions in the short run. According to his theory, the price of each factor of production tends to be equal to its marginal product. Therefore, understanding how entrepreneurs deal with uncertainty problems requires analyzing the effect of each factor of production on output in the relevant economic period. According to Stigler, there should be short-term alterations of the productive services because of unexpected movements in prices and outputs. Stigler mainly focuses on price uncertainties and the short-term alterations of a fixed plant. The main point of his work is to understand how entrepreneurs change the allocation of productive services in the short run by altering their allowances when input price movements occur.

This paper builds on Stigler's work and studies how governments change the allocation of resources in societies when there is an unexpected increase in demand for health care services. However, the focus of this paper is uncertainty in patient numbers rather than uncertainty in prices, as studied by Stigler (1939). I approach the quantity uncertainty problem from the norms of welfare economics by modeling a socially optimal treatment and expenditure system under uncertainty in patient numbers. By comparing socially optimal outcomes with the outcomes of healthcare providers, I find that healthcare providers fail to provide socially optimal levels of treatment. By assuming that the government can ex-post judge the allocation of resources and can take initiatives to change it by using subsequent transfers, I investigate further to decide if socially optimal outcomes are achievable through government intervention.

The rest of the paper proceeds as follows. Section 3 provides background on the Medicaid and Medicare reimbursement systems. Section 4 outlines the model and presents the social problem and the hospital's problem. Section V presents the results. Finally, Section 6 summarizes the findings and provides policy implications.

2 Background

The healthcare providers in this paper are considered nonprofit hospitals. Nonprofit hospitals, or not-for-profit (NFP), are community-oriented institutions driven by the mission of benefiting the communities they serve rather than shareholder returns. These institutions neither seek profits nor desire to retain administrative control in the market. Instead, they are driven by altruistic motives, such as improving community well-being and providing higher quality and access to medical services. According to the Internal Revenue Code, 501(c), NFP hospitals are exempt from income and property taxes. In return, they must provide community benefits and charity care to justify their favored tax status. NFP hospitals are funded mainly by charity, religion, research, and educational funds.

One of the primary funding sources of NFP hospitals is the Medicare reimbursement system. *Medicare* is a federal medical insurance program that serves people aged 65 years and older of all income levels. On the other hand, *Medicaid* is a federal-state assistance program for low-income people of every age. Medicare and Medicaid differ in terms of how they reimburse hospitals. Medicaid makes payments to healthcare providers based on a fee-for-service agreement or managed care arrangement. However, Medicare reimburses hospitals and doctors depending on the services provided to Medicare patients. In addition to these payments, Medicare provides payments to hospitals through various payment channels such as add-on payments and pay-for-performance programs, explained further in the next section.

There are many different views in the health care literature about the behavior of nonprofit hospitals and the objective function they maximize. According to Newhouse (1970), nonprofit hospitals maximize the utility of hospital administrators by choosing the quantity and quality of health care services, subject to a break-even constraint. On the other hand, Pauly and Redisch (1937) focus more on the role of physicians in hospital decisions. The authors suggest that nonprofit hospitals maximize utility from physician income, subject to a zero-profit constraint. Finally, Deneffe and Masson (2002) argue that nonprofit hospitals do not have any incentives to ration Medicare and Medicaid patients. Therefore, they maximize a general utility function that includes profits and the number of all patients by choosing the number of the private patient.

This paper assumes that a nonprofit hospital maximizes the aggregate average quality of treatment provided to society, subject to a zero-profit constraint. In other words, NFP hospitals choose treatment inputs by minimizing the harm of the illness. Also, I assume that more patients bring more revenues to hospitals. Due to the nonprofit nature of the hospitals, the extra revenues are used to invest more in resources, such as providing more equipment for medical staff and having more nursing staff.

2.1 Medicaid reimbursement system

The U.S. federal government and states jointly fund the Medicaid program. Each state is allowed to establish its own Medicaid reimbursement rates under general federal requirements. Therefore, Medicaid reimbursement rates differ from state to state. However, they generally pay for services through fee-for-service or managed care arrangements (Medicaid.gov, 2021).

The fee-for-service rate reimburses healthcare providers for specific services provided to each patient. Under this system, each state establishes specific payments for each service category and reimburses providers for providing that service to a patient. In other words, this system pays providers by the volume of services they provide to patients. For example, office visits and tests are considered two separate services. Therefore, if a provider needs to run some tests for a patient visiting the doctor's office, then the provider is reimbursed separately for each of these services. However, since the reimbursements depend on the total quantity of care provided to a patient, this system may incentivize healthcare providers to ask for unnecessary treatments.

Due to the risks of overutilization of health care services under the fee-for-service system, many states use the managed care model instead. Under this model, states set arrangements with managed care organizations[1] (MCOs) that receive a set payment from states to provide specific services to enrolled Medicaid beneficiaries. In return for providing those services, managed care organizations are paid monthly capitation payments (U.S. Department of Health and Human Services, 2021). This model is sometimes called comprehensive or capi-

[1]MCO is a health care provider or a group or organization of medical service providers who offer managed care health plans.

tated care because of the fixed per-person costs. In addition, the managed care model aims to eliminate unnecessary procedures because patient care is considered as a whole rather than individual services under this model. In other words, Medicaid pays a fixed amount of money per patient regardless of the services provided to the patient. The total amount then is divided according to the number of services given to the patient.

2.2 Medicare Prospective Payment System

The Centers for Medicare & Medicaid Services (CMS) is the largest payer for health care in the United States. When Medicare was first established in 1965, hospitals were reimbursed retrospectively for capital and labor costs associated with the care of Medicare patients. However, under this cost-based reimbursement system, health care spending increased dramatically. In response to increasing costs, policymakers established Medicare Prospective Payment System (PPS) in 1983 to control costs and encourage more efficient uses of medical technology. PPS has a partial cost reimbursement feature because only capital input expenses are reimbursed.

In contrast, labor input expenses are covered by a fixed price paid per case, regardless of the actual cost incurred for that patient (Acemoglu & Finkelstein, 2008). Under this system, inpatient admission cases are divided into categories called Diagnosis-Related Groups (DRG). Under the DRG reimbursement method, Medicare pays hospitals a fixed rate for each charge of inpatient hospital care to reward efficient hospitals and encourage inefficient hospitals to become more efficient (U.S. Department of Health and Human Services, 2001).

There are over 740 DRG categories, each of which has a unique DRG weight assigned by CMS. Each DRG weight represents the average level of resources required for an average Medicare patient in the DRG relative to the average level of resources for all Medicare patients. The differences between the DRG weights reflect the cost variations between different types of treatments. The variation in weights ranges from 0.2 to 39.85 depending on the treatment. For example, more costly treatments like heart transplants have higher DRG weights. In addition, The DRG weights are adjusted by the wage index and the cost of living adjustment factor to account for local differences in wages and prices. In addition to these adjustments, the DRG weights can be higher if the hospital treats a high percentage of Medicare patients.

In this situation, the hospital receives a percentage add-on payment known as the Disproportionate Share Hospital (DSH) adjustment, applied to the DRG-adjusted base payment rate. Also, if the hospital is a teaching hospital, it receives a percentage add-on payment, the Indirect Medical Education (IME) adjustment, applied to the DRG-adjusted base payment rate. Furthermore, if the hospital provides treatments for outlier cases, it receives an additional payment. CMS designed all these add-on payments under the Hospital Quality Initiatives to protect hospitals from significant financial losses and prevent a possible quality deterioration of care (U.S. Centers for Medicare & Medicaid Services, 2020).

2.3 Hospital Quality Initiatives

Quality health care is a high priority for the Centers for Medicare & Medicaid Services. CMS implements various quality incentives, including pay for reporting, public reporting, and add-on payments to achieve the high-quality care goals: effective, safe, efficient, patient-centered, equitable, and timely care. In addition to these incentives, CMS has other financial incentives to support better care for individuals, better health for society, and lower costs. Making additional payments to healthcare providers under the Pay for Performance programs, also known as the Value-Based Programs, is one of the most promising ones. Value-based programs reward (penalize) healthcare providers when they meet (fail to achieve) pre-defined targets for quality indicators and specific performance measures. Compared to the previously mentioned incentives, these programs are seen as more comprehensive quality strategies to reform how health care is delivered and paid for. CMS claims that "value-based programs are important because they help us move toward paying providers based on the quality, rather than the quantity of care they give patients" (U.S. Centers for Medicare & Medicaid Services, 2020). There are four value-based programs for hospitals, which are summarized below.

- End-Stage Renal Disease Quality Incentive Program (ESRD QIP):
  Launched on January 1, 2012, this is the first mandatory federal pay-for-performance program. Under this program, CMS penalizes renal dialysis[2] facilities by reducing their

[2]Renal dialysis is the process of removing excess water, solutes, and toxins from the blood in people whose kidneys can no longer perform these functions naturally.

total reimbursement payments up to 2 percent depending on their performance on quality of care measures. These quality measures are categorized into four groups: clinical care measure, care coordination measure, safety measure, and a measure of patients' perspective. Facilities are scored by CMS based on their performances on these measures and penalized if specific performance standards are not met. This program aims to promote high-quality services in renal dialysis facilities.

- Hospital Value-Based Purchasing (VBP) Program:
  Under the VBP program, CMS rewards acute care hospitals with incentive payments based on their performance on specific measures such as mortality and complications, healthcare-associated infections, patient safety and experience, and efficiency and cost reductions (U.S. Centers for Medicare & Medicaid Services, 2021). Hospitals receive bonus payments under the Inpatient Prospective Payments System, depending on their performance on these quality measures.

- Hospital Readmission Reduction Program (HRRP):
  HRRP is established to encourage hospitals to reduce avoidable readmissions by improving communication and care coordination between patients and caregivers regarding post-discharge planning. Under HRRP, CMS penalizes hospitals for having higher than expected readmission rates for acute myocardial infarction, heart failure, and pneumonia (AHA Fact Sheet: Hospital Readmissions Reduction Program, 2015).

- Hospital-Acquired Conditions (HAC) Reduction Program:
  Hospital-acquired conditions are undesirable and avoidable conditions that a patient experiences while staying in a hospital but being treated for something else. These conditions include pressure ulcers, hip fractures after falls in the hospital, and postoperative sepsis. An IBM Watson Health states, "hospital-acquired conditions contribute over $2 billion in excess hospital costs annually, add an average of eight days to the patient length of stay and increase mortality risk by 72 percent" (IBM Watson Health, 2018). HAC program is established to reduce these undesirable costs and to improve patients' safety by encouraging hospitals to implement best practices. Under the HAC

reduction program, CMS penalizes hospitals with a HAC score higher than 75 percent of all total HAC scores, i.e., the lowest-performing 25 percent of all hospitals, by reducing their payments by 1 percent.

The above incentives indicate that governments are responsible for advancing public health. In other words, better care for individuals and better health for society are the main objectives of government incentives. Therefore, it is assumed throughout the paper that the government can intervene in the healthcare market if the market outcomes provide unsatisfactory results.

3 The model

The society I consider is composed of three parties: a representative government, a non-profit hospital, and N number of ex-ante identical individuals. It is assumed that society goes through an illness stage, considered an epidemic. Some individuals contract the illness, while others do not get it. Therefore, the total number of individuals in society, represented by N, is equal to the sum of the total number of healthy, denoted by N^H, and the total number of sick individuals, N^S. Also, I assume that all sick individuals get the same illness. However, there is uncertainty about its severity. In other words, the level of the sickness may be more or less severe, depending on nature's choice of probability level, denoted by θ^j. This probability of the sickness specifies different states of the illness: a good and a bad state, where $j = \{G,B\}$. I assume this probability is higher in the bad state, θ^B, than in the good state, θ^G. Therefore, the bad state brings more sick individuals into society than the good state. Assuming N is large and there is independence in the probability of individuals becoming sick, then $\theta^B N > \theta^G N$. This specification also implies that the total number of healthy individuals in society is higher in the good state than in the bad state, $(1-\theta^G)N > (1-\theta^B)N$.

Society is assumed to have only two inputs; variable and fixed. The fixed, x, and the variable, y^j, inputs have two different uses in society. First, they are used in the treatment of the disease. The role of the hospital is to provide treatment for the disease by using the treatment resources. Since the fixed input is immobile in the short run, the hospital decides on its level before the realization of the sickness severity. On the other hand, the hospital adjusts the variable input in response to the changing numbers of sick individuals in society brought by the different states of the illness. Second, y^j and x are used to produce a single consumption good, which is assumed to be the individuals' single source of utility. Since the production of the consumption good and the treatment use the same inputs, the aggregate consumption in society, C^j, differs depending on the state of the illness where $j = \{G, B\}$. In other words, C^j represents the aggregate consumption left in society after levels of the treatment resources are decided. Moreover, healthy and sick individuals are assumed to have different levels of private consumption, denoted by C^{jH} and C^{jS}, respectively. With these specifications, the aggregate consumption in society is

$$C^j = N^{jH} C^{jH} + N^{jS} C^{jS}, \quad j = \{G, B\}. \tag{1}$$

The government's allocation of the inputs between treatment and consumption determines the level of aggregate consumption. In other words, the government has the below resource constraint

$$N^{jH} C^{jH} + N^{jS} C^{jS} + x + y^j = R, \quad j = \{G, B\}, \tag{2}$$

where R is some fixed number representing the government's aggregate resources. Eq. 2 shows that the government has a trade-off between consumption and treatment. In other words, the more the inputs are used in the treatment, the less the aggregate consumption good is left to the individuals.

The healthy and sick individuals have different utility functions given below, respectively:

$$U^{jH} = U^{jH}(C^{jH}) \tag{3}$$

$$U^{jS} = U^{jS}\left(C^{jS}, S^j\left(x, y^j, N^{jS}\right)\right), \quad j = \{G, B\}, \tag{4}$$

where $S^j\left(x, y^j, N^{jS}\right)$ is the level of the sickness remaining after the treatment. Equations 3 and 4 are concave and increasing functions of the consumption good. However, Eq. 4 decreases with the increase in the level of sickness remaining after the treatment. That is either because of the discomfort it gives to a sick individual or because of the loss of productive time during the illness. Therefore, $S^j\left(x, y^j, N^{jS}\right)$ is characterized as the harm of the sickness to a sick individual. These relationships of the functions are summarized as follows:

$$\frac{\partial U^{jH}}{\partial C^{jH}} > 0, \qquad \frac{\partial U^{jS}}{\partial C^{jS}} > 0, \qquad \frac{\partial U^{jS}}{\partial S^{j}} < 0.$$

The sickness level remained after the treatment, $S^j\left(x, y^j, N^{jS}\right)$, decreases with the increase in the level of the treatment resources. In contrast, it increases when the number of sick individuals is higher. In other words, the higher the investments made in the treatment, the lower the sickness level remained after the treatment. On the other hand, having more sick individuals spreads the treatment resources across more sick individuals and hence makes the harm worse.

$$\frac{\partial S^{j}}{\partial x} < 0, \qquad \frac{\partial S^{j}}{\partial y^{j}} < 0, \qquad \frac{\partial S^{j}}{\partial N^{jS}} > 0.$$

3.1 The social problem

The social problem is considered a two-stage game. In the first stage, the representative government chooses the level of the fixed input, x, before the realization of sickness severity. After the sickness severity is observed, the representative government allocates resources optimally in the second stage. In other words, the second stage refers to the government's resource allocation, including allocating the consumption levels between the healthy and the sick individuals and choosing how much of the variable input goes into the treatment. The latter allocation implicitly determines the aggregate consumption left to society. The optimal levels of C^{jH}, C^{jS}, y^j are derived from the maximization of the social welfare function subject to the resource constraint, Eq. 2. The social welfare function includes the ex-post individual utilities. Therefore, the utilities are a function of the level of the fixed input chosen in the first stage and the realized level of the sickness severity. The ex-post social welfare function is:

$$W^j\left(U^{jH}(C^{jH}),\, U^{jS}\left(C^{jS},\, S^j\left(x, y^j, N^{jS}\right)\right)\right), \quad j = \{G, B\}. \tag{5}$$

Social welfare increases with an increase in the healthy and sick individuals' utility functions. Therefore, higher levels of aggregate consumption increase the well-being of society, whereas higher levels of sickness decrease it. Then, there are five variables that change the well-being of society: C^{jH}, C^{jS}, x, y^j, and N^{jS}. Since individuals derive utility from consumption, the first two variables increase social well-being. On the other hand, since the harm of the illness decreases in the treatment resources and increases in the number of sick individuals, the third and the fourth variables increase the well-being of society, whereas the last variable decreases it. In other words, the total number of sick individuals imposes a negative externality on society.

The social problem is constructed as the maximization of the social welfare function subject to the resource constraint defined in Eq. 2. I assume that the representative government's preferences derive from a weighted utilitarian welfare function. Then, after obtaining the level of x from the first stage, the representative government's problem in the second stage is:

$$\max_{C^{jH},C^{jS},y^j} \alpha(1-\theta^j)NU^{jH}(C^{jH}) + \beta\theta^j NU^{jS}\left(C^{jS},\, S^j\left(x,y^j,\theta^j N\right)\right) \tag{6}$$

s.t.

$$(1-\theta^j)NC^{jH} + \theta^j NC^{jS} + x + y^j = R, \qquad j = \{G,B\}, \tag{7}$$

where α is the welfare weight for the healthy and β is the welfare weight for the sick individuals. The optimal values of the government's allocation in the second stage are functions of the chosen level of x and the realized sickness severity.

The first-order conditions are derived by forming the below Lagrangian expression with a multiplier λ^j.

$$L^j = \alpha(1-\theta^j)NU^{jH}(C^{jH}) + \beta\theta^j NU^{jS}\left(C^{jS},\, S^j\left(x,y^j,\theta^j N\right)\right) - \lambda^j\left((1-\theta^j)NC^{jH} + \theta^j NC^{jS} + x + y^j - R\right) \tag{8}$$

The first-order conditions are

$$\frac{\partial L^j}{\partial C^{jH}}: \quad \alpha(1-\theta^j)N\,U_1^{jH}(C^{jH}) - \lambda^j\left((1-\theta^j)N\right) = 0 \tag{9}$$

$$\frac{\partial L^j}{\partial C^{jS}}: \quad \beta\theta^j N\,U_1^{jS}\left(C^{jS},\, S^j\left(x,y^j,\theta^j N\right)\right) - \lambda^j\left(\theta^j N\right) = 0 \tag{10}$$

$$\frac{\partial L^j}{\partial y^j}: \quad \beta\theta^j N\,U_2^{jS}\left(C^{jS},\, S^j\left(x,y^j,\theta^j N\right)\right) S_2^j\left(x,y^j,\theta^j N\right) - \lambda^j = 0 \tag{11}$$

$$-\frac{\partial L^j}{\partial \lambda^j}: \quad (1-\theta^j)NC^{jH} + \theta^j NC^{jS} + x + y^j - R = 0 \tag{12}$$

There are two equations derived from the first-order conditions:

$$\alpha\,U_1^{jH}(C^{jH}) = \beta\,U_1^{jS}\left(C^{jS},\, S^j\left(x,y^j,\theta^j N\right)\right) \tag{13}$$

$$U_1^{jS}\left(C^{jS},\, S^j\left(x,y^j,\theta^j N\right)\right) = \theta^j N\,U_2^{jS}\left(C^{jS},\, S^j\left(x,y^j,\theta^j N\right)\right) S_2^j\left(x,y^j,\theta^j N\right). \tag{14}$$

By simplifying the equations above, the equilibria below are obtained.

$$\frac{U_1^{jH}(C^{jH})}{U_1^{jS}\left(C^{jS},\, S^j\left(x, y^j, \theta^j N\right)\right)} = \frac{\beta}{\alpha} \tag{15}$$

$$\frac{U_1^{jS}\left(C^{jS},\, S^j\left(x, y^j, \theta^j N\right)\right)}{U_2^{jS}\left(C^{jS},\, S^j\left(x, y^j, \theta^j N\right)\right)} = S_2^j\left(x, y^j, \theta^j N\right)\theta^j N, \qquad j = \{G, B\} \tag{16}$$

The left-hand side of Eq. 15 is the ratio of a healthy individual's marginal utility of consumption to a sick individual's marginal utility of consumption in the good and bad state of the illness. In other words, Eq. 15 is the marginal rate of substitution between the consumption of healthy and sick individuals in society. The numerator represents how much a change in the utility of a healthy individual affects social welfare when his private consumption changes. On the other hand, the denominator demonstrates how much a change in the utility of a sick individual affects social welfare when his private consumption changes. The right-hand side of Eq. 15 is the ratio of the government's welfare weights for the sick and healthy individuals. For example, if the government is a utilitarian that equally weights the individuals' utilities, then this ratio is simply 1. That means that at the optimum, the government balances the effect of the healthy individuals' marginal utility of consumption and the effect of the sick individuals' marginal utility of consumption on social welfare. However, if the government is more like a Rawlsian who gives a higher welfare weight for individuals in an inferior position, then the welfare weight to the sick individuals would be higher than the healthy. In that case, the ratio on the right-hand side of Eq. 15 would be greater than 1.

The left-hand side of Eq. 16 represents a sick individual's marginal rate of substitution between consumption and sickness in the good and bad state of the illness. This ratio is negative since there is the disutility of sickness to a sick individual. On the other hand, the right-hand side of Eq. 16 describes the overall reduction in the level of sickness remaining after treatment as the level of the variable input increases by one unit. In other words, the right-hand side of the equation shows how much sickness can be prevented if the government allocates an additional unit of resources for the treatment rather than the consumption of a sick individual. Therefore, Eq. 16 characterizes the total cost of increasing the sick individual's private consumption by one unit on society's welfare. In other words, it represents the government's trade-off between the welfare of sick individuals and the overall welfare of society. This trade-off raises some critical questions when discussing optimal health policies. For instance, what is the government's optimal way of distributing resources between treatment and the overall well-being of individuals in society?

3.2 The hospital's problem

The nonprofit hospital's problem is considered a two-stage problem. In the first stage, the hospital chooses the level of the fixed treatment resource x before the realization of sickness severity. Then, after observing the sickness severity, the hospital's problem in the second stage is to choose the highest possible level of the variable input that ensures the zero-profit constraint in the good and bad state of the illness.

The hospital is assumed to be a price taker in the input markets that faces a wage rate of w per unit of the variable resource and a cost of r per unit of the fixed resource. I simply assume that the input prices are $w = r = 1$. Then, the equation below demonstrates the total cost of the hospital in each state of the illness.

$$TC^j\,(x^P, y^{jP}) = x^P + y^{jP}, \qquad j = \{G, B\}, \tag{17}$$

where x^P is the level of the fixed treatment resource chosen by the hospital in the first stage and y^{jP} is the variable input level chosen by the hospital in the second stage. The level of the fixed treatment resource is constant in the second stage. On the other hand, decisions on the level of the variable resource depend on the probability of sickness severity. I assume that the hospital provides the socially optimal level of the fixed treatment resource in the first stage. In other words, it is assumed that $x^P = x$. Then, the total cost of the hospital is a function of the optimal level of the variable input and the chosen level of the treatment resources.

This paper assumes that the government reimburses the hospital through Medicare reimbursement payments. It also assumes that the reimbursement payment is the only source of revenue for the hospital. Therefore, the hospital's revenue function is defined based on the Medicare reimbursement system. Under the Medicare reimbursement system, the government pays hospitals a flat rate per patient depending on the diagnosis-related group (DRG) weights. Along with these payments, the government makes additional payments (or deductions) to hospitals as part of the quality initiatives, which constitute the second revenue source for hospitals. For instance, add-on payments are one of the quality incentive tools used by CMS for achieving the goal of high-quality care. Hospitals receive a percentage add-on payment applied to the DRG-adjusted base payment rate under specific circumstances. For example, if the hospital treats a high percentage of low-income patients or if the hospital is a teaching hospital. In addition to the existing specific circumstances, CMS may establish new add-on payments if necessary for quality of care. For example, CMS established a New COVID-19 Treatments Add-on Payment (NCTAP) in 2020, which increased the DRG-adjusted base payment rate by 20%. This new add-on payment aims to mitigate potential financial disincentives for hospitals to provide up-to-date COVID-19 treatments to patients.

Based on the Medicare reimbursement system, the revenue function for the hospital per patient is defined as follows:

$$R^j(x, y^{jP}, \theta^j), \qquad j = \{G, B\}. \tag{18}$$

Eq. 18 depends on the hospital's investments in the variable resource, the optimal level of the fixed resource, and the probability of the sickness severity. It is assumed that $R^j(x, y^{jP}, \theta^j)$ is increasing in x, y^{jP}, and θ^j. There are two reasons to assume that the hospital is reimbursed more for investing more in the treatment resources. First, under the Medicare reimbursement system, CMS reimburses hospitals depending on the average resources used in treatment. Therefore, the hospital gets higher payments when it invests more in treatment resources. Second, investing more in treatment resources increases the quality of the treatment provided to sick individuals. Higher quality of care brings more revenues to the hospital under the Medicare reimbursement system. Thus, I assume that the hospital's revenue function is increasing in the fixed and variable resources used by the hospital.

On the other hand, based on the Medicare reimbursement system, I assume that the hospital's revenue function increases in the probability of the sickness severity. The DRG rates are higher when more intensive procedures are performed to treat more severe cases, whereas it is lower for the less severe cases requiring less intensive procedures. For example, coronary artery bypass graft (CABG) surgery and angioplasty are two major cardiac procedures that are medically substitutable for moderately ill patients. The former procedure is recommended only when the patient is severely ill, whereas the latter treatment is recommended when the illness is less severe. Hospitals using the former treatment are paid more than those using the latter treatment because CABG has a higher DRG weight than angioplasty. Based on this reimbursement system, the hospital payment is assumed to increase with the probability of the sickness severity. In other words, the hospital receives higher payments in the bad state than in the good state of the illness because it provides high-intensity medical treatment to severely sick patients in the bad state. In contrast, payments are lower in the good state because the hospital provides less-intensity medical treatment to treat moderately sick patients.

Then, the hospital's profit in each state of the illness is defined by the following equation.

$$\Pi^j(x, y^{jP}, \theta^j) = \theta^j N R^j(x, y^{jP}, \theta^j) - x - y^{jP}, \qquad j = \{G, B\} \tag{19}$$

The hospital uses higher revenues to invest more in treatment resources due to its nonprofit character. Then, the hospital's problem in the second stage is to provide the highest possible level of the variable input y^{jP} that ensures zero profit in each state of the illness. The following equation characterizes this choice

$$\theta^j N R^j(x, y^{jP}, \theta^j) - x - y^{jP} = 0 \tag{20}$$

$$y^{jP} = \theta^j N R^j(x, y^{jP}, \theta^j) - x, \qquad j = \{G, B\}, \tag{21}$$

where y^{jP} is the level of the variable resource that the hospital uses in the treatment. The variable input level is determined by the hospital's budget constraint. In other words, the level of y^{jP} is the remainder after the expenses on the fixed input chosen in the first stage is subtracted from the hospital's total revenues. Thus, its value depends on the chosen level of the fixed treatment resource and the realization of the sickness severity.

Since this paper assumes that the hospital provides the socially optimal level of the fixed treatment resource in the first stage, the question in the second stage is whether the hospital provides the socially optimal level of the variable input to society. The above equations show that the hospital's total revenue in each state of the illness is the main determinant of the hospital's outcome. If the hospital provides $y^{jP} = y^j$ in both states of the illness, then the hospital's expenditures on treatment resources are optimal. Therefore, the hospital provides socially optimal treatment to society.

4 Results

The below results are obtained by solving the social problem defined in Eq. 6. Thus, they characterize the government's optimal allocation of resources when the severity of the sickness increases.

Proposition 1: *When the probability of the severity of the sickness increases from θ^G to θ^B, society's total consumption decreases at the optimum.*

Both the consumption of the sick and healthy individuals decreases as θ^j increases. In other words, the sick and healthy individuals consume less when the sickness worsens.

Proof. See Appendix

Corollary 1: *If the government is utilitarian that equally weights the utilities of the healthy and sick individuals, $\alpha = \beta$, then the consumption of the sick and healthy individuals decreases the same amount when the sickness severity increases.*

Proof. From the comparative statics, the change in the consumption of the healthy and sick individuals are given as below:

$$\frac{\partial C^{jH}}{\partial \theta^j} = -\frac{mN\left((\theta^j N)V - y^j\right)}{4k^2\alpha(C^{jS}-S^j)^2(y^j)^3 x}\cdot\frac{1}{|H_L|}$$

$$\frac{\partial C^{jS}}{\partial \theta^j} = -\frac{mN\left((\theta^j N)V - y^j\right)}{4k^2\beta(C^{JS}-S^j)^{1/2}(C^{jH})^{3/2}(y^j)^3 x}\cdot\frac{1}{|H_L|},$$

where $\beta(C^{jH})^{1/2} = \alpha(C^{jS}-S^J)^{1/2}$ from Eq. 45 and $x(y^j)^2 = (\theta^j N)^2$ from Eq. 48 presented in Appendix. Then,

$$\frac{\partial C^{jH}}{\partial \theta^j} = \frac{\partial C^{jS}}{\partial \theta^j} = -\frac{mN\left((\theta^j N)V - y^j\right)}{4k^2\alpha(C^{jS}-S^j)^2(y^j)^3 x}\cdot\frac{1}{|H_L|},$$

where $\theta^j NV - y^j \leq 0$ from Eq. 77 and $|H_L| < 0$ from Eq. 61.

Proposition 2: *When the probability of the severity of the sickness increases from θ^G to θ^B, the variable input level increases at the optimum.*

Proof. See Appendix

5 Interpretation of the results and conclusion

This paper provides a model for a stylized society with three parties; a representative government, a non-profit hospital, and N number of ex-ante identical individuals. Society is assumed to go through an illness stage, but there is uncertainty about the severity of the illness. Therefore, two states are defined for the sickness; good and bad. There are fewer sick individuals in society in the good state than in the bad state. Since society does not know what state of the illness will occur, it must make some decisions ex-ante and others ex-post. Thus, the problem is considered a two-stage game. In the first stage, the government and the hospital decide on the fixed input level before the sickness severity is observed. After the realization of the sickness severity, in the second stage, the government chooses the consumption of healthy and sick individuals as well as the optimal level of the variable input. On the other hand, the role of the hospital in the second stage is to use the highest possible level of the variable input in the treatment that ensures zero profit. Also, the hospital in this paper is assumed to be reimbursed through Medicare.

The results show that at the optimum, the consumption levels of the healthy and sick individuals decrease, whereas the variable input level increases when the sickness worsens. Since the variable input is used in the production of the consumption good and the treatment, fewer consumption results in more variable input used in the treatment. In other words, the government allocates resources less to the individuals and more to treatment as the sickness worsens. In short, optimally allocating resources in the second stage requires increasing the level of the variable input used in the treatment by reducing the consumption of individuals in society when the sickness worsens. Also, Corollary 1 shows that if the government is utilitarian, then optimally increasing the variable input level requires reducing the consumption levels of the sick and healthy individuals by an equal amount. In other words, treating sick individuals for society's overall well-being creates an externality for healthy individuals.

The Medicare insurance program has two parts; hospital insurance (Part A) and medicare insurance (Part B). Medicare Part A covers hospitalization and some nursing home and home health care services. Medicare Part B, on the other hand, covers doctor visits and other outpatient services, such as lab tests and diagnostic screenings. There are monthly fees that Medicare beneficiaries pay for Part A and Part B insurance programs. Most people do not have to pay a premium for Part A because beneficiaries age 65 or older and who worked and paid Medicare taxes for at least ten years are exempt from Part A premium.[3] However, everyone must pay for Part B premium if they want their insurance to cover medically necessary services and preventive services[4] that are not covered by Medicare Part A. In addition to the premiums, Part A and Part B programs have deductibles and co-insurance that the beneficiaries pay out-of-pocket.[5]

Each year, CMS releases the premiums, deductibles, and co-insurance rates for the Medicare Part A and Part B programs. CMS increases the Medicare Part A and Part B premiums, deductibles, and co-insurance to finance increased Medicare spending. For example, the standard premium for Medicare Part B was $134 per month for 2017 and 2018, $135.50 for 2019, $144.60 for 2020, and $148.50 for 2021. In addition, the annual deductible for all Medicare Part B beneficiaries was $198 for 2020 and $203 for 2021. On the other hand, the Medicare Part A inpatient hospital deductible was $1,408 in 2020 and $1,484 in 2021. (U.S. Centers for Medicare & Medicaid Services, 2021).

Suppose the government in this paper is considered to increase the Medicare premiums in the bad state of the illness to transfer resources from individuals to the health care provider. In that case, individuals' private consumption decreases due to increased expenses on insurance premiums; hence more resources will go into the treatment. This scenario is consistent with the findings in this paper. However, the government also increases out-of-pocket fees, such as deductibles and co-insurance rates, to offset the high health care costs. In other words, the real-world scenario shows that elderly individuals who consume healthcare services have higher expenses than healthy elderly individuals. However, Corollary 1 illustrates that sick and healthy individuals' consumption must decrease by an equal amount to increase the treatment resources optimally when the sickness worsens. Therefore, collecting more money from sick elderly individuals than the healthy does not satisfy the optimum conditions I found in this paper. Therefore, I suspect the government's transfer payments from individuals to

[3]About 99 percent of Medicare beneficiaries do not have Part A premium since they have at least 40 quarters of Medicare-covered employment.

[4]Preventive services include exams, shots, lab tests, and screenings.

[5]Deductible is a payment that beneficiaries must pay for covered health care services before their insurance plan begins paying their health costs. Co-insurance is a percentage of a medical charge that beneficiaries pay which applies after the amount of deductible is met.

treatment resources will be optimal. Since the government's reimbursement payment to the hospital is the primary determinant in providing the optimal level of the variable input into the treatment, I suspect the hospital will provide the optimal treatment to society. Based on this finding, healthcare providers fail to bring socially optimal treatment to society when the sickness brings more sick individuals into society.

This paper shows that the current trends in healthcare markets of utilizing the Medicare reimbursement system are appropriate as long as the government can find a way for optimal distribution. Since healthcare providers have very little control over the choice of treatment, the optimal treatment requires the government to increase the money collected from healthy elderly individuals. One possible way is to impose a tax on their consumption. However, the existing real-world scenario is not sufficient to bring optimal treatment to society.

6 Future work

There are two stages of the problem. This paper so far focuses only on the second stage by assuming that the health care provider is choosing the optimal level of the fixed input for the treatment in the first stage.

The hospital's problem in the first stage is to choose the level of the fixed input x to maximize the expected average quality of the treatment. The quality of the treatment per patient is defined based on the following logic:

$$S\left(x, y^j, N^{jS}\right) - S, \qquad j = \{G, B\}, \tag{22}$$

where $S(x, y^j, N^{jS})$ is the level of the treatment remaining after the treatment, i.e., harm of the sickness to a sick individual, and S is some fixed number representing the level of harm from the illness if there were no resource constraints. The smaller Eq. 22, the higher the quality of the treatment. Therefore, the hospital's problem in the first stage is to minimize Eq. 22 for each patient subject to the budget constraint. In other words, the hospital maximizes the aggregate quality of the treatment subject to the zero-profit constraint:

$$\theta^j N\left[S - S\left(x, y^j, N^{jS}\right)\right], \qquad j = \{G, B\}. \tag{23}$$

Since there is uncertainty about the number of sick individuals in the first stage, the hospital's problem in the first stage is to choose the level of x to maximize the expected average quality of the treatment subject to the zero-profit constraint. That is,

$$\max_{x} \quad \pi\, \theta^G N \left[S - S^G\left(x, y^G(x,\theta^G), \theta^G N\right)\right] + (1-\pi)\, \theta^B N \left[S - S^B\left(x, y^B(x,\theta^B), \theta^B N\right)\right] \tag{24}$$

s.t.

$$\theta^G N\, R^G(x, y^G(x,\theta^G), \theta^G) + N\, R^B(x, y^B(x,\theta^B), \theta^B) - x - y^G(x,\theta^G) - y^B(x,\theta^B) = 0. \tag{25}$$

Similarly, the government's problem in the first stage is to choose the level of x to maximize the expected social welfare of society subject to the resource constraint:

$$\max_{x}\ \pi\, W^G\left[U^{GH}\left(C^{GH}(x,\theta^G)\right), U^{GS}\left(C^{GS}(x,\theta^G), S^G(x, y^G(x,\theta^G), \theta^G N\right)\right] + \tag{26}$$

$$(1-\pi)\, W^B\left[U^{BH}\left(C^{BH}(x,\theta^B)\right), U^{BS}\left(C^{BS}(x,\theta^B), S^B(x, y^B(x,\theta^B), \theta^B N\right)\right] \tag{27}$$

s.t.

$$(1-\theta^G)\, N\, C^{GH}(x,\theta^G) + \theta^G\, N\, C^{GS}(x,\theta^G) + x + y^G(x,\theta^G) = R_G \tag{28}$$

$$(1-\theta^B)\, N C^{BH}(x,\theta^B) + \theta^B\, N C^{BS}(x,\theta^B) + x + y^B(x,\theta^B) = R_B, \tag{29}$$

where R_G and R_B are some fixed numbers representing the government's aggregate resources in the good and bad state, respectively.

Appendices

A First and second-order conditions

The social welfare maximization from Eq. 6 gives the below problem:

$$\max_{C^{jH},C^{jS},y^j} \alpha(1-\theta^j)NU^{jH}(C^{jH}) + \beta\theta^j NU^{jS}\left(C^{jS},\, S^j\left(x,y^j,\theta^j N\right)\right) \tag{30}$$

s.t.

$$(1-\theta^j)NC^{jH} + \theta^j NC^{jS} + x + y^j = R, \qquad j = \{G,B\}. \tag{31}$$

The first-order conditions are derived by forming the below Lagrangian expression with a multiplier λ^j.

$$L = \alpha(1-\theta^j)NU^{jH}(C^{jH}) + \beta\theta^j NU^{jS}\left(C^{jS},\, S^j\left(x,y^j,\theta^j N\right)\right) - \lambda^j\left((1-\theta^j)NC^{jH} + \theta^j NC^{jS} + x + y^j - R\right) \tag{32}$$

For the simplicity of notations, call $m = \alpha\,(1-\theta^j)N$ and $k = \beta\,\theta^j N$. Then, the above expression becomes:

$$L = mU^{jH}(C^{jH}) + kU^{jS}\left(C^{jS},\, S^j\left(x,y^j,\theta^j N\right)\right) - \lambda^j\left(\frac{m}{\alpha}C^{jH} + \frac{k}{\beta}C^{jS} + x + y^j - R\right). \tag{33}$$

The first-order conditions from the above Lagrangian are:

$$\frac{\partial L}{\partial C^{jH}}: \quad m\,U_1^{jH}(C^{jH}) - \frac{m}{\alpha}\lambda^j = 0 \tag{34}$$

$$\frac{\partial L}{\partial C^{jS}}: \quad k\,U_1^{jS}\left(C^{jS},\, S^j\left(x,y^j,\theta^j N\right)\right) - \frac{k}{\beta}\lambda^j = 0 \tag{35}$$

$$\frac{\partial L}{\partial y^j}: \quad k\,U_2^{jS}\left(C^{jS},\, S^j\left(x,y^j,\theta^j N\right)\right) S_2^j\left(x,y^j,\theta^j N\right) - \lambda^j = 0 \tag{36}$$

$$\frac{\partial L}{\partial \lambda^j}: \quad -\frac{m}{\alpha}C^{jH} - \frac{k}{\beta}C^{jS} - x - y^j + R = 0. \tag{37}$$

By taking the Hessian of the Lagrangian, the second-order conditions are derived as follows:

$$
\begin{aligned}
H_L &= \begin{bmatrix} \frac{\partial^2 L}{\partial (C^{jH})^2} & \frac{\partial^2 L}{\partial C^{jH}\partial C^{jS}} & \frac{\partial^2 L}{\partial C^{jH}\partial y^j} & \frac{\partial^2 L}{\partial C^{jH}\partial \lambda^j} \\ \frac{\partial^2 L}{\partial C^{jS}\partial C^{jH}} & \frac{\partial^2 L}{\partial (C^{jS})^2} & \frac{\partial^2 L}{\partial C^{jS}\partial y^j} & \frac{\partial^2 L}{\partial C^{jS}\partial \lambda^j} \\ \frac{\partial^2 L}{\partial y^j\partial C^{jH}} & \frac{\partial^2 L}{\partial y^j\partial C^{jS}} & \frac{\partial^2 L}{\partial (y^j)^2} & \frac{\partial^2 L}{\partial y^j\partial \lambda^j} \\ \frac{\partial^2 L}{\partial \lambda^j\partial C^{jH}} & \frac{\partial^2 L}{\partial \lambda^j\partial C^{jS}} & \frac{\partial^2 L}{\partial \lambda^j\partial y^j} & \frac{\partial^2 L}{\partial (\lambda^j)^2} \end{bmatrix} \qquad (38) \\
&= \begin{bmatrix} mU_{11}^{jH} & 0 & 0 & -\frac{m}{\alpha} \\ 0 & kU_{11}^{jS} & kU_{12}^{jS} S_2^j & -\frac{k}{\beta} \\ 0 & kU_{21}^{jS} S_2^j & k\left(U_2^{jS} S_{22}^j + U_{22}^{jS}(S_2^j)^2\right) & -1 \\ -\frac{m}{\alpha} & -\frac{k}{\beta} & -1 & 0 \end{bmatrix} \qquad (39)
\end{aligned}
$$

The following functional forms are assigned to the utility functions and the sickness remaining after the treatment to derive the results:

$$U^{jH} = (C^{jH})^{1/2} \qquad (40)$$

$$U^{jS} = \left(C^{jS} - S^j\right)^{1/2} \qquad (41)$$

$$S^j = \frac{\theta^j N}{x y^j} \qquad (42)$$

The first and second-order derivatives from the above functions are given below.

$$
\begin{aligned}
U_1^{jH}(C^{jH}) &= \frac{1}{2}(C^{jH})^{-1/2} & U_{11}^{jH}(C^{jH}) &= -\frac{1}{4}(C^{jH})^{-3/2} \\
U_1^{jS}(C^{jS}, S^j) &= \frac{1}{2}(C^{jS} - S^J)^{-1/2} & U_{11}^{jS}(C^{jS}, S^j) &= -\frac{1}{4}(C^{jS} - S^J)^{-3/2} \\
U_{12}^{jS}(C^{jS}, S^j) &= \frac{1}{4}(C^{jS} - S^J)^{-3/2} & U_2^{jS}(C^{jS}, S^j) &= -\frac{1}{2}(C^{jS} - S^J)^{-1/2} \\
U_{22}^{jS}(C^{jS}, S^j) &= -\frac{1}{4}(C^{jS} - S^J)^{-3/2} & U_{21}^{jS}(C^{jS}, S^j) &= \frac{1}{4}(C^{jS} - S^J)^{-3/2} \\
S_2^j(x, y^j, \theta^j) &= -\frac{\theta^j N}{x(y^j)^2} & S_{22}^j(x, y^j, \theta^j) &= \frac{2\theta^j N}{x(y^j)^3} \\
S_3^j(x, y^j, \theta^j) &= \frac{1}{x y^j} & S_{23}^j(x, y^j, \theta^j) &= -\frac{1}{x(y^j)^2}.
\end{aligned}
$$

The utility functions defined in Eq. 40 and Eq. 41 satisfy the basic properties: monotonically increasing in the consumption good with a diminishing marginal utility. That is,

$$\begin{aligned} U_1^{jH}(C^{jH}) &= \frac{1}{2}(C^{jH})^{-1/2} > 0 \\ U_1^{jS}(C^{jS}, S^j) &= \frac{1}{2}(C^{jS} - S^J)^{-1/2} > 0 \\ U_{11}^{jH}(C^{jH}) &= -\frac{1}{4}(C^{jH})^{-3/2} < 0 \\ U_{11}^{jS}(C^{jS}, S^j) &= -\frac{1}{4}(C^{jS} - S^J)^{-3/2} < 0. \end{aligned}$$

For simplicity, the below notations are assigned to denote first and second-order conditions.

$A = U_{11}^{jS}$	$D = S_2^j$	$G = U_{22}^{jS}$	$R = U_{11}^{jH}$
$E = S_3^j$	$H = S_{22}^j$	$C = U_{12}^{jS}$	$F = U_2^{jS}$
$J = S_{23}^j$	$V = (C^{jH} - C^{jS})$		

From the first-order conditions, equations 10 and 11, I have the below equality:

$$\lambda^j = \beta U_1^{jS}\left(C^{jS}, S^j\left(x, y^j, \theta^j N\right)\right) = k\, U_2^{jS}\left(C^{jS}, S^j\left(x, y^j, \theta^j N\right)\right) S_2^j\left(x, y^j, \theta^j N\right) \tag{43}$$

Replacing the above functional forms, equations 40, 41, and 42 and $k = \beta\,\theta^j N$ in the above equality gives:

$$\beta\frac{1}{2}(C^{jS} - S^J)^{-1/2} = \beta\theta^j N\left(-\frac{1}{2}(C^{jS} - S^J)^{-1/2}\right)\cdot\left(-\frac{\theta^j N}{x(y^j)^2}\right) \tag{44}$$

$$x(y^j)^2 = (\theta^j N)^2. \tag{45}$$

Similarly, replacing the above functional forms in Eq. 15 gives me the following condition.

$$\frac{U_1^{jH}(C^{jH})}{U_1^{jS}\left(C^{jS}, S^j\left(x, y^j, \theta^j N\right)\right)} = \frac{\beta}{\alpha} \tag{46}$$

$$\frac{\frac{1}{2}(C^{jH})^{-1/2}}{\frac{1}{2}(C^{jS} - S^J)^{-1/2}} = \frac{\beta}{\alpha} \tag{47}$$

$$\beta(C^{jH})^{1/2} = \alpha(C^{jS} - S^J)^{1/2} \tag{48}$$

Since the government is assumed to be a utilitarian, Eq. 48 implies the below condition:

$$C^{jH} = (C^{jS} - S^J). \tag{49}$$

A.1 Bordered Hessian

With the help of the notations above, a bordered Hessian from Eq. 39 is given below:

$$|\overline{H}| = \begin{vmatrix} mR & 0 & 0 & -\frac{m}{\alpha} \\ 0 & kA & kCD & -\frac{k}{\beta} \\ 0 & kCD & k(FH+GD^2) & -1 \\ -\frac{m}{\alpha} & -\frac{k}{\beta} & -1 & 0 \end{vmatrix}. \tag{50}$$

Its bordered principal minors are defined as:

$$|\overline{H_2}| = \begin{vmatrix} kA & kCD & -\frac{k}{\beta} \\ kCD & k(FH+GD^2) & -1 \\ -\frac{k}{\beta} & -1 & 0 \end{vmatrix} \tag{51}$$

$$|\overline{H_3}| = \begin{vmatrix} mR & 0 & 0 & -\frac{m}{\alpha} \\ 0 & kA & kCD & -\frac{k}{\beta} \\ 0 & kCD & k(FH+GD^2) & -1 \\ -\frac{m}{\alpha} & -\frac{k}{\beta} & -1 & 0 \end{vmatrix} \tag{52}$$

with $|\overline{H_n}| = |\overline{H_3}| = |\overline{H}|$ where n is the number of choice variables. Then, the sufficient condition for negative-definiteness of the objective function requires the bordered principal minors to alternate in sign.

$$|\overline{H_2}| = \frac{k\left(\beta^4x^2\left(y^j\right)^4 - 2\beta^2k^2x\left(y^j\right)^2 + 4\beta k^3\left(C^{jS}-S^j\right)xy^j + k^4\right)}{4\beta^4x^2\left(y^j\right)^4\left(C^{jS}-S^j\right)^{3/2}} \tag{53}$$

$$= \frac{k\beta^4\left(x^2(y^j)^4 - 2(\theta^jN)^2x(y^j)^2 + (\theta^jN)^4 + 4(\theta^jN)^3(C^{jS}-S^j)xy^j\right)}{4\beta^4x^2\left(y^j\right)^4\left(C^{jS}-S^j\right)^{3/2}} \tag{54}$$

$$= \frac{k\left(\left(x(y^j)^2-(\theta^jN)^2\right)^2 + 4(\theta^jN)^3(C^{jS}-S^j)xy^j\right)}{4x^2\left(y^j\right)^4\left(C^{jS}-S^j\right)^{3/2}} \tag{55}$$

$$= \frac{\beta\theta^jN\left(x(y^j)^2-(\theta^jN)^2\right)^2}{4x^2\left(y^j\right)^4\left(C^{jS}-S^j\right)^{3/2}} + \frac{\beta(\theta^jN)^4}{(C^{jS}-S^j)^{1/2}x(y^j)^3} \tag{56}$$

Replacing the equality obtained in Eq. 45 in the above determinant gives the following:

$$|\overline{H_2}| = \frac{\beta(\theta^jN)^4}{(C^{jS}-S^j)^{1/2}x(y^j)^3} > 0. \tag{57}$$

$$|\overline{H_3}| = -\frac{km\left[(C^{jS}-S^j)^{1/2}(\theta^jN)^2\alpha^2\left((\theta^jN)^2+x(y^j)^2\left(\frac{x(y^j)^2-2(\theta^jN)^2}{(\theta^jN)^2}\right)\right)\right]}{16\alpha^2(C^{jS}-S^j)^2(C^{jH})^{3/2}(y^j)^4x^2} \tag{58}$$

$$-\frac{4km(\theta^jN)\left[km(C^{jH})^{3/2}+(\theta^jN)^2\alpha^2(C^{jS}-S^j)^{3/2}\right]}{16\alpha^2(C^{jS}-S^j)^2(C^{jH})^{3/2}(y^j)^3x} \tag{59}$$

$$= -\frac{km\left(x(y^j)^2-(\theta^jN)^2\right)^2}{16(C^{jS}-S^j)^{3/2}(C^{jH})^{3/2}(y^j)^4x^2} - \frac{4km(\theta^jN)\left[km(C^{jH})^{3/2}+(\theta^jN)^2\alpha^2(C^{jS}-S^j)^{3/2}\right]}{16\alpha^2(C^{jS}-S^j)^2(C^{jH})^{3/2}(y^j)^3x} \tag{60}$$

Again, by replacing the equality obtained in Eq. 45 in the above determinant, the below result is obtained.

$$|\overline{H_3}| = -\frac{4km(\theta^jN)\left[km(C^{jH})^{3/2}+(\theta^jN)^2\alpha^2(C^{jS}-S^j)^{3/2}\right]}{16\alpha^2(C^{jS}-S^j)^2(C^{jH})^{3/2}(y^j)^3x} < 0 \tag{61}$$

According to the above results, the sufficient condition for a constrained local maximum is satisfied.

B Comparative statics

The comparative statics aims to illustrate how a change in the probability of the sickness severity affects the consumption of healthy and sick individuals and the level of the variable input. Findings from the comparative statics are addressed below.

B.1 Proofs of Propositions 1 and 2

Using the Hessian of the Lagrangian by replacing the defined functional forms and dividing the rows by k, the below equation is obtained.

$$H_L = \begin{bmatrix} \frac{m}{k}R & 0 & 0 & \frac{m}{\alpha k} \\ 0 & A & CD & 1/\beta \\ 0 & CD & FH+GD^2 & 1/k \\ \frac{m}{\alpha k} & 1/\beta & 1/k & 0 \end{bmatrix} \cdot \begin{bmatrix} \frac{\partial C^{jH}}{\partial \theta^j} \\ \frac{\partial C^{jS}}{\partial \theta^j} \\ \frac{\partial y^j}{\partial \theta^j} \\ \frac{-\partial \lambda^j}{\partial \theta^j} \end{bmatrix} = \begin{bmatrix} 0 \\ -CNE \\ -\left(\frac{FD}{\theta^j}+FJN+DGEN\right) \\ \frac{N}{k}V \end{bmatrix} \tag{62}$$

$$\frac{\partial C^{jH}}{\partial \theta^j} = \frac{\begin{vmatrix} 0 & 0 & 0 & \frac{m}{\alpha k} \\ -CNE & A & CD & 1/\beta \\ -\left(\frac{FD}{\theta^j}+FJN+DGEN\right) & CD & FH+GD^2 & 1/k \\ \frac{N}{k}V & 1/\beta & 1/k & 0 \end{vmatrix}}{|L|} \tag{63}$$

$$= -\frac{mN\left((\theta^j N)V - y^j\right)}{4k^2\alpha(C^{jS}-S^j)^2(y^j)^3 x} \cdot \frac{1}{|H_L|}, \qquad |H_L| < 0 \tag{64}$$

$$\frac{\partial C^{jS}}{\partial \theta^j} = \frac{\begin{vmatrix} \frac{m}{k}R & 0 & 0 & \frac{m}{\alpha k} \\ 0 & -CNE & CD & 1/\beta \\ 0 & -\left(\frac{FD}{\theta^j}+FJN+DGEN\right) & FH+GD^2 & 1/k \\ \frac{m}{\alpha k} & \frac{N}{k}V & 1/k & 0 \end{vmatrix}}{|H_L|} \tag{65}$$

$$= \frac{mN\left[(\theta^j N)V - y^j\right]\left[-4\,(\theta^j N)\,(C^{JS}-S^J)xy^j + x(y^j)^2 - (\theta^j N)^2\right]}{16k^3(C^{JS}-S^j)^{3/2}(C^{jH})^{3/2}(y^j)^4 x^2} \cdot \frac{1}{|H_L|} \tag{66}$$

$$= \frac{mN\left[(\theta^j N)V - y^j\right]\left[-4\,(\theta^j N)\,(C^{JS}-S^J)xy^j\right]}{16k^3(C^{JS}-S^j)^{3/2}(C^{jH})^{3/2}(y^j)^4 x^2} \cdot \frac{1}{|H_L|} \tag{67}$$

$$= -\frac{mN\left((\theta^j N)V - y^j\right)}{4k^2\beta(C^{JS}-S^j)^{1/2}(C^{jH})^{3/2}(y^j)^3 x} \cdot \frac{1}{|H_L|}, \qquad |H_L| < 0 \tag{68}$$

$$\frac{\partial y^j}{\partial \theta^j} = \frac{\begin{vmatrix} \frac{m}{k}R & 0 & 0 & \frac{m}{\alpha k} \\ 0 & A & -CNE & 1/\beta \\ 0 & CD & -\left(\frac{FD}{\theta^j}+FJN+DGEN\right) & 1/k \\ \frac{m}{\alpha k} & 1/\beta & \frac{N}{k}V & 0 \end{vmatrix}}{|H_L|} \tag{69}$$

$$= \left\{ \frac{mN\left[(\theta^j N) - V(y^j)x\right]\left[x(y^j)^2 - (\theta^j N)^2\right]}{16(C^{jS}-S^j)^{3/2}(C^{jH})^{3/2}k^3(y^j)^3x^2} \right. \tag{70}$$

$$\left. - \frac{4mN}{16(C^{jS}-S^j)^{1/2}k^2(y^j)^2 x}\left[\frac{(\theta^j N)^2}{k(C^{jH})^{3/2}} + \frac{m}{(C^{jS}-S^j)^{3/2}\alpha^2}\right]\right\} \cdot \frac{1}{|H_L|} \tag{71}$$

$$= -\frac{mN\left[(C^{jS}-S^j)^{3/2}\alpha^2(\theta^j N)^2 + mk(C^{jH})^{3/2}\right]}{4(C^{jS}-S^j)^2(C^{jH})^{3/2}k^3\alpha^2(y^j)^2 x} \cdot \frac{1}{|H_L|}, \qquad |H_L| < 0 \tag{72}$$

$$= \frac{y^j}{k^4\theta^j} \ \geq\ 0 \tag{73}$$

From the resource constraint defined in Eq. 2, where $N^{jH} = (1-\theta^j)N$ and $N^{jS} = \theta^j N$, we have the below condition.

$$NC^{jH} - \theta^j NC^{jH} + \theta^j NC^{jS} + x + y^j = R \tag{74}$$

$$NC^{jH} + x - R = \theta^j N(C^{jH} - C^{jS}) - y^j \tag{75}$$

$$NC^{jH} + x - R = \theta^j NV - y^j \tag{76}$$

It is plausible enough to assume that the left-hand side of Eq. 76 is negative. In other words, the aggregate resources in society are greater than the sum of the level of the fixed input and some level of consumption. Based on this assumption,

$$\theta^j NV - y^j \leq 0. \tag{77}$$

Then, using the information obtained above and in Eq. 61, the signs of Equations 64 and 68 are determined as follows.

$$\frac{\partial C^{jH}}{\partial \theta^j} \leq 0 \tag{78}$$

$$\frac{\partial C^{jS}}{\partial \theta^j} \leq 0 \tag{79}$$

B.2 Comparative statics of the sickness

$$\frac{\partial S^j}{\partial \theta^j} = \frac{Nxy^j - \theta^j Nx\frac{\partial y^j}{\partial \theta^j}}{(xy^j)^2} \tag{80}$$

$$= \frac{Nxy^j - \theta^j Nx\frac{y^j}{k^4\theta^j}}{(xy^j)^2} \tag{81}$$

$$= \frac{Nxy^j - \frac{Nxy^j}{k^4}}{(xy^j)^2} \tag{82}$$

$$= \frac{N}{xy^j}\left(1 - \frac{1}{k^4}\right) \geq 0 \tag{83}$$

The above result is not surprising. When the probability of sickness severity increases, the sickness remaining after the treatment increases. In other words, the sickness causes more harm to a sick individual in the bad state than in the good state of the illness.

References

[1] Daron Acemoglu and Amy Finkelstein. Input and technology choices in regulated industries: Evidence from the health care sector. *Journal of Political Economy*, 116(5):837–880, 2008.

[2] Zubin J. Eapen Colleen K. McIlvennan and Larry A. Allen. Hospital readmissions reduction program. *Circulation*, 131(20):1796–1803, 2015.

[3] Daniel Deneffe and Robert T Masson. What do not-for-profit hospitals maximize? *International Journal of Industrial Organization*, 20(4):461–492, 2002.

[4] Centers for Medicare & Medicaid Services. 2021 medicare parts a & b premiums and deductibles. Retrieved from: `https://www.cms.gov/newsroom/fact-sheets/2021-medicare-parts-b-premiums-and-deductibles`, November 2020. (Accessed on 11/10/2021).

[5] Centers for Medicare & Medicaid Services. Acute inpatient pps. Retrieved from: `https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/AcuteInpatientPPS`, February 2020. (Accessed on 11/10/2021).

[6] Centers for Medicare & Medicaid Services. What are the value-based programs? Retrieved from: `https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/Value-Based-Programs`, January 2020. (Accessed on 11/10/2021).

[7] Centers for Medicare & Medicaid Services. Hospital value-based purchasing program. Retrieved from: `https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/Hospital-Value-Based-Purchasing-`, February 2021. (Accessed on 11/10/2021).

[8] Martin Gaynor and Gerard F Anderson. Uncertain demand, the structure of hospital costs, and the cost of empty hospital beds. *Journal of health economics*, 14(3):291–317, 1995.

[9] Medicaid. gov. Financial management. Retrieved from: `https://www.medicaid.gov/medicaid/financial-management/index.html`, 2016. (Accessed on 11/10/2021).

[10] IBM Watson Health. Hospital-acquired conditions lead to avoidable cost and excess deaths. Retrieved from: `https://www.ibm.com/downloads/cas/X97QXLER`, October 2018. (Accessed on 11/10/2021).

[11] Theodore E Keeler and John S Ying. Hospital costs and excess bed capacity: A statistical analysis. *The Review of Economics and Statistics*, pages 470–481, 1996.

[12] Joseph P Newhouse. Toward a theory of nonprofit institutions: An economic model of a hospital. *The American economic review*, 60(1):64–74, 1970.

[13] Mark Pauly and Michael Redisch. The not-for-profit hospital as a physicians' cooperative. *The American Economic Review*, 63(1):87–99, 1973.