An Analytical Model of Demand for Hospital Inpatient Care

Pius Eze

Abstract

This paper presents an analytical model of demand for curative care such as hospital inpatient care. In this essay, a person’s illness is used as a measure of the person’s health status. Medical care and non-medical consumption are the choice variables in a consumer’s constrained utility-maximization problem. Methodologically, this model of demand for medical care does not rely on capital theory, common in the literature. All else equal, a person’s demand for hospital inpatient care depends on how ill the person is. That is, the demand is a function of the severity of the person’s illness. The model is static and deterministic, and is based on the fact that, by definition, only sick persons demand curative medical care. One implication of distinguishing curative care from other types of care is the possibility of satiation when the person is healthy. Satiation diminishes differences in medical care expenditures between wealthy and poor persons and can help explain the empirically observed low income elasticity of demand for hospital inpatient care.

1. Introduction

Health care expenditures have been rising in many countries and this trend has continued since the end of World War II. Hospital expenditures are the largest and fastest growing portion of health care expenditures. In response, governments and other payers increasingly attempt to control or stem these trends (Aaron 1981, Culyer and Newhouse 2000). The prospective hospital reimbursement system instituted by the United States government Medicare programme is an example of a government effort (Eze and Wolfe, 1993). This paper presents a concept that can aid policy efforts by with an analytical model of demand for hospital inpatient care in order to highlight the curative aspects. Conceptually, illness imposes costs on the ill, and these costs, psychic, financial and time costs, as well as risk of death depend on how ill the person is (Arrow 1963). Also, the notion of cure of an illness presupposes a level of health to which a sick person’s health is to be elevated for the person to be considered cured. The severity of a person’s illness is a measure of how much current health has fallen from its healthy level.

The model presented in this paper is constructed partly following Arrow’s (1963) suggestion that the purpose of curative medical care is to cure illness. Only sick persons utilize curative care. Curative care is distinguished from other types of medical care in that a person derives utility from it only if and when sick. This implies satiation in the utilization of curative care. Satiation suggests a state-dependent utility function with marginal utility of hospital care non-positive when the person is healthy and non-negative when the person is sick (iatrogenic problems apart). Also, satiation diminishes differences in medical care expenditures between wealthy and poor persons and can help explain the empirically observed low income elasticity of demand for hospital inpatient care (Phelps and Newhouse, 1974).

The demand for medical care, like the demand for other goods and services, is influenced by the opportunity costs of alternative uses of resources. As such, the demand for medical care is not perfectly inelastic, in general. In a sense, medical care is a normal good (Newhouse 1981). As a commodity, medical care is special but not unique. That is the consensus in the literature. Medical care is special because of, among other things, its importance in the maintenance of life; but so also is food, for example. As with the demand for many other goods and services such as smoke detectors, the demand for medical care is a derived demand. That is, medical care is not demanded for its own sake but rather for the health it yields (Arrow 1963; Grossman 1972). But not only is the demand for medical care derived from the demand for health, but the demand for health is itself a derived demand given that what a person really wants is not health per se but rather the quality of life and longevity health helps make possible. One could go further with this trend of reasoning. For example, longevity is desirable because, all else equal, longevity increases lifetime utility by lengthening the time over which instantaneous utilities are
summed. The longevity motive is to prolong life or to postpone death. On the other hand, quality of life is desired because it represents higher levels of instantaneous utilities. The quality of life motive is related to the avoidance or reduction of various costs and inconveniences imposed by illness.

The relationship between illness and curative care utilization is similar to the relationship between hunger and food consumption. Both illness and hunger occur prior to the demand (for medical care or food), and medical care can cure illness just like food can quench hunger. Also, the utilization of care and the consumption of food are contemporaneous with their respective accruing benefits, for the most part. Given these similarities, it seems appropriate to analyze curative medical care with the same standard model of consumer choice used to analyze food.

This paper shows that many of the results from human capital models of demand for health care are also derivable from standard consumer models. The remainder of the paper is as follows. Section 2 reviews the relevant literature. Section 3 derives a relationship between illness and medical care utilization. Section 4 derives comparative static analyses as well as the implications of the model are presented in section 5. The conclusion is in section 6.

2. The Literature

The model of demand for medical care presented in this paper (a) is a standard consumer model; (b) it distinguishes curative medical care from other types of care and (c) distinguishes the illness process from aging. As such, the model differs methodologically from existing models of demand for medical care. Existing models are based, more or less, on the human capital model, specialized by Grossman (1972) as well as its variations and extensions (for example, Muurinen 1982, Cropper 1981). According to Muurinen (1982) "Grossman's work is the first truly economic explanation of health decision-making." But, in keeping with the human capital doctrine, Grossman's model emphasizes the importance of time as an economic resource, such that a person desires to be healthy to avoid decreases in his or her productive time due to ill health; this has policy implications. But regardless of its general acceptance, the human capital model is inadequate for analyzing the demand for curative medical care such as hospital inpatient care because it is conceptualized and generalized as if all health care were preventive.

On the other hand, curative health care is distinguished from preventive care in that, by definition, a person gets sick first and then seeks curative care (Arrow, 1963). Just like food, curative care is better suited to the standard consumer theory. Even in inter-temporal (or dynamic) consumer models, contemporaneous consumption is often assumed to enter only contemporaneous utility functions (Deaton and Muellbauer, 1980) based on the usual separability and additivity assumptions.

Arrow (1963) suggests that the demand for medical care is a response to the risks of ill-health. Grossman's (1972) model can be thought of as one early attempt to formalize Arrow's notions, but based on the human capital theory (Becker, 1965) rather than the standard consumer theory, with the underpinning idea that health is a depreciable stock of capital goods; and medical care is utilized as a human-capital investment in order to counteract depreciations in health (capital). In Grossman's original model, the consumer is a health producer and the optimal amount of health care (input) demanded is determined in the capital market at an equilibrium where the marginal efficiency of health investment (that is, of health care) is equal to the cost of capital (or interest rate) plus the rate of depreciation. According to Cropper (20??), Grossman's analysis is necessarily incomplete and his results are, at best, a lower limit because the model neglects additional benefits from the existence of a marginal rate of substitution (between health care and non-medical consumption) equal to the ratio of the marginal utilities of medical care to non-medical consumption which equals their respective price ratios.

In the human capital model, the person's health stock depreciates increasingly with age until, eventually, the person's budget is overextended and his health stock falls to the death level and the person dies at age T. With medical care demanded in order to counteract (aging or natural) health depreciations, Grossman's model yields the result that the demand for hospital inpatient care is income elastic, contrary to intuition and contrary to existing empirical evidence (Feldstein 1974). Starting with Phelps and Newhouse (1974), there exist other models of demand for medical care that attempt to utilize standard consumer theory. Muurinen (1982) refers to such models as 'pure consumption' models. But such consumer models follow Grossman in ignoring distinctions between curative care and other types of care (see Folland et al., 1990, Phelps 1992). These models have not been as popular possibly because they give no intuitive justification for the choice. An alternative conceptual framework is what has been termed the cost-of-illness framework (Goddeeris 1983, Harrington and Portney 1987) the essence is the delineation of the economic consequences of illness and death, with the demand for medical care zero if illness and death did not exist or were not possible. The cost-of-illness framework has intuitive appeal but the framework is not well-developed analytically; an appropriate model would be in the context of lifetime optimization.

According to Arrow (1963), a person wants to be healthy because the healthier he is (the less severely ill) relatively speaking, the higher his survival probability, in addition to the reduction of material, psychic and time...
costs imposed by illness (Cropper 1981, Muurinen 1982) which are higher the more severe the illness. This suggests that models of demand for medical care take illness and disease processes into account. Intuitively, curative medical care is demanded in order to restore the sick person’s health to its prior level with no effort to exceed that level (Phelps 2013). But even with a general understanding that a person’s health outcome depends on the severity and nature of the illness as well as on the effectiveness of medical care (Dardanoni and Wagstaff, 1987) these concepts yet to be formalized in analytical models of demand for medical care.

3. Relation between Severity of Illness and Medical Care

Consider a single time period. Let \( s_0 \) represent a measure of how ill a person is at the beginning of the period. The purpose of curative medical care is to cure this existing illness; that is, to reduce \( s_0 \) possibly to zero. Let \( M \) be the amount of medical care the person utilizes within the period. Let \( h(M) \) represent an index of the effectiveness of medical care, where \( h'(.) > 0, h''(.) < 0 \). A person’s medical care need is the amount of medical care required to cure the illness, to reduce \( s_0 \) to zero. That is, the person’s medical care need is the value of \( M \) that solves the following identity:

\[
3.1 \quad s_0 - h(M) = 0.
\]

Stated differently, let \( m^+ \) be the person’s medical care need (Boulding 1966) defined implicitly as the amount of medical care required to cure the illness; that is, to reduce \( s_0 \) to zero. Note that the person’s medical care need is determined independent of the person’s preferences or income. The following results can be deduced from Equation 3.1: (a) curative care need, \( m^+ \), is zero for a healthy person, \( m^+ = 0 \leftrightarrow s_0 = 0 \); (b) the sicker a person is, the larger his or her medical care need, all else equal: \( s_0 > s_0 \leftrightarrow m^+_1 > m^+_2 \) for person 1 and person 2. If medical care were free, the sick person would choose to utilize medical care at the need level, all else equal (Aaron 1981). If, on the other hand, medical care is not free then, possibly because of the consumer’s budget constraint, the optimal amount of medical care (actually) utilized may be less than his medical care need. In which case his or her illness may or may not be fully cured within the given period. It is possible that at the end of the time period under consideration, the person is not fully cured, meaning that some level of the illness remains. Let \( s \) measure the amount (or severity) of illness ‘remaining’ in the period. Then the technical relationship between the severity of the person’s illness and the person’s medical care usage in a time period is represented generally as follows:

\[
3.2 \quad s = s_0 - h(M) \geq 0.
\]

Equation 3.2 states that how ill a person is in a time period (for example, at the middle or end of the period) depends, all else equal, on the initial health conditions and on the effectiveness of medical care utilized in that period. In a static, non-stochastic model, it is assumed that \( s = 0 \) if \( s_0 = 0 \). In the health economics literature, \( h(.) \) is termed a health production function (Grossman 1972, Pauly 1980). Note that in the relationship between health and medical care as presented in Equation 3.2, medical care \( M \) is utilized in order to cure existing illness. Compare Equation 3.2 to Equation 3.3 or Equation 3.4 where medical care \( M \) counteracts the depreciation in health (Muurinen 1982):

\[
3.3 \quad H = H_0(1 - \delta) + h(M), \text{ for a static model;}
\]

\[
3.4 \quad H(t) = H_0e^{-\delta t} + h(M(t)), \text{ for a multi-period model,}
\]

where \( 0 \leq \delta \leq 1 \) is the rate of health depreciation and \( H_0 \) is the person’s initial health stock.

4. The Demand for Curative Care

This section derives the consumer’s demand for curative medical care. Only sick persons utilize curative medical care, implying satiation in the utilization of curative care when the person is healthy. Satiation, in turn, implies a state-dependent utility function of the following form (see, for example, de Meza 1983, Goodeeris 1983):

\[
4.1 \quad U(C) = \begin{cases} 
U(C) & \text{if } s = 0 \\
U(C, s) & \text{if } 0 < s < s_d \\
0 & \text{if } s = s_d
\end{cases}
\]

Where \( U(.) \) is a (real-valued) utility function. In Equation 4.1, illness (measured by its severity) \( s \) enters the utility of only the sick person, for whom \( 0 < s < s_d \). \( C \) is a composite non-medical consumption. The following assumptions are made regarding the differentials: \( U_c > 0, U_{cc} < 0, U_s < 0, U_{ss} < 0, U_{cs} < 0 \). For example, \( U_{cs} < 0 \)
can represent the fall in a sick person’s appetite for food due to the illness; \(U_s < 0\) would imply that a more severely illness induces relatively higher disutility of the illness.

Consider the consumer’s budget constraint. Let \(p(M)\) represent a cost function, for example, the consumer’s out-of-pocket medical expenses which may depend on the variety and depth of insurance coverage (determined outside the present model). The price of the non-medical consumption is \(p_c\). Let \(Y(s)\) be the person’s potential income function, the maximum income he can earn if he allocated all his available time to income generating activities. Available time can be affected adversely by illness. For example, more severe illnesses tend to cause more absenteeism or lower productivity. Both bequest and unearned income can be ignored (assumed to be zero).

In a more simplified case, \(Y = \omega T\), where \(T\) is 24 hours or 365 days and \(\omega\) is the given wage rate. If the person is not sick, then he can potentially work the whole 24 hours or 365 days and earn potential income of \(24\omega\) or \(365\omega\). If the person is sick in the period, then his available time is only \((T - L)\) and the person presumably can earn only \(Y = \omega(T - L)\), where \(L\) is sick time. It is the available time \((T - L)\) rather than the potential time \((T)\) that a person shares into leisure time, labour time, etc.

The consumer’s budget constraint is:

\[
4.2 \quad p_c C + p(M) = Y(s), \quad p'_c > 0, p''_c \geq 0, p(0) = 0, Y' \leq 0, Y'' \geq 0.
\]

Note that illness (through its severity \(s\)) can affect an individual’s utility Equation 4.1 as well as his budget constraint in Equation 4.2 specifically the income function. Medical care consumption \(M\) does not appear in the utility function. Instead, \(M\) appears in both the budget Equation 4.2 and health Equation 3.2 constraints. This is what is meant by the saying that the demand for medical care is a derived demand, derived from the demand for health.

Viewed as an optimization problem, the consumer chooses a medical bundle \(M\) and nonmedical consumption bundle \(C\) to maximize his single period utility function in Equation 4.1 subject to his budget constraint from Equation 4.2:

\[
4.3 \quad \max_{C,M} U(C, s), \quad \text{such that} \quad 4.4 \quad p_c C + p_m M = Y(s)
\]

Note that the budget constraint in Equation 4.4 is a simplification of Equation 4.2 with \(p_m\) as the per unit price of medical care. \(Y(s)\) represents the consumer’s real income (or available resources). If illness affects the person’s income (or earning power, etc.), then income is modeled as endogenous, \(Y = Y(s)\).

The relation of illness to medical care usage represented in Equation 3.2 can be simultaneous because the severity of a person’s illness can itself be endogenous. For example, measures of severity of illness found in data sets can be ex-ante (before medical care is utilized), contemporaneous, or ex-post. For predicting medical care usage, the present model presupposes that ex ante measures of SOI are the appropriate choice. This allows the severity of a person’s illness to be considered exogenous as needs be.

\(p_c\) can be normalized to a value of one, then the nonmedical consumption \(C\) becomes a numeraire good, interpreted as the total real expenditure on everything else other than medical care. Then the generalized consumer problem is:

\[
4.6 \quad \max_{C,M} U(C, s), \quad \text{subject to} \quad 4.7 \quad s = s_0 - h(M) \geq 0, \quad 4.8 \quad C + p(M) = Y(s), \quad 4.9 \quad s(0) = s_0 \text{ given} \quad 4.10 \quad 0 \leq s < s_0, 0 \leq s_0 < s_d, \quad 4.11 \quad s = 0 \text{ if } s_0 = 0.
\]

The optimal choice of medical care \(M\) is obtained by solving the system of Equations 4.6 – Equations 4.11.

For a healthy person, for whom \(s_0 = 0, s = 0; Y = Y(0) = Y_0\), the optimal non-medical consumption is \(C^* = Y\), optimal medical consumption is \(M^* = 0\) and medical expenditure \(p(M^*) = 0\); curative medical care is not demanded by a healthy person. On the other hand, if a person is sick initially, \(s_0 > 0\), he can get cured such that \(s = 0\). It is possible also for such a person to remain sick in the period, \(s > 0\), even after utilizing some medical care, if the illness was treated but not completely cured. After substituting Equations 4.7 and Equations 4.8 into Equation 4.6, the sick person’s problem is the following:

\[
4.12 \quad \max_{C,M} \{Y[s_0 - h(M)] - p(M), s_0 - h(M)\}
\]

The equilibrium equation is:

\[
4.13 \quad h'Y U_e + p'U_e + h'U_s = 0
\]

Equation 4.13 can be rearranged to get Equation 4.14. For presentational clarity, the arguments are dropped from Equations 4.13 and Equations 4.14.
\[ 4.14 \quad -[\gamma' + \frac{U}{U_c} h'] = p' \]

\( U_c \) is the marginal utility of the non-medical consumption \( C \). \( U_s \) is the direct marginal effect of illness on utility. The equilibrium Equation 4.14 states that the marginal benefits of curative care equal its marginal cost to the individual, \( p'(M) \). The marginal benefits are the two benefits of reducing illness, made up of a direct marginal utility (the marginal rate of substitution) plus an indirect marginal effect through income \( Y \). Each term is multiplied by the marginal efficiency of medical treatment (or marginal health product), \( h'(M) \). The person’s demand for curative medical care \( M \) is obtained by solving Equation 4.13. The above model can be solved by parameterization. A possible example of the utility function in Equation 4.1 is the following Cobb-Douglas function (Deaton et al., 1980):

\[ 4.15 \quad U(C,s) = C^\alpha (s_1 - s)^\beta, \quad 0 \leq \alpha, \beta \leq 1, \quad 0 \leq s \leq s_1. \]

5. Comparative Static Analyses and Interpretation of the Result

Comparative static analysis can be performed on Equation 4.13. For example, the effects of changes in severity of illness or effects of changes in income on the utilization of curative medical care can be derived from Equations 4.7 and Equations 4.8 for \( s \) and \( C \), respectively into Equation 4.13 gives the following:

\begin{align*}
5.1 \quad & h'(M)Y[s_0-h(M)]U(M+s) - p(M), [s_0-h(M)] + \\
& + p'(M)U(M+s) - (h')^2 Y' + (h')^2 Y'' \quad \text{d}Y + \\
& + h'(M)U(M+s) - h''(M) - p'(M)Y' + (h')^2 Y'' \quad \text{d}M
\end{align*}

Totally differentiating Equation 5.1 with respect to \( M, Y, p \) and \( s_0 \) gives:

\begin{align*}
5.2 \quad & -\{(p' + h'Y)^2 \frac{U_{cc}}{U_c} + p'(p' + h'Y)^2 \frac{U_{cs}}{U_c} + (1 + Y) \frac{h'}{U_c} \frac{U_{ss}}{U_c} - h'' \frac{U_s}{U_c} - p'' - h'' Y' + (h')^2 Y'' \} \quad \text{d}M + \\
& + \{(p' + h'Y) \frac{U_{cc}}{U_c} + h' \frac{U_{cs}}{U_c} \} \quad \text{d}Y + \\
& - \{(p' + h'Y) \frac{U_{cc}}{U_c} + h' \frac{U_{cs}}{U_c} \} \quad \text{d}p + \\
& + \{(p' + h'Y) \frac{U_{cc}}{U_c} + (p' + 2h'Y) \frac{U_{cs}}{U_c} + h' \frac{U_{ss}}{U_c} + Y'' \} \quad \text{d}s_0 = 0
\end{align*}

Note the symmetry between the effect on medical care utilization of changes in income (\( \text{d}Y \)) and changes in price (\( \text{d}p \)). In order to avoid strange results, it is reasonable to assume that the terms inside the brackets on the first line of Equation 5.2, taken together, differ from zero. The relation between medical care usage \( M \) and the initial (and exogenous) severity of illness \( s_0 \) can be ascertained from Equation 5.2 by holding income and price constant (\( \text{d}Y = 0, \text{d}p = 0 \)) and solving the resulting equation. That entails dividing the terms inside the brackets on the fourth line of Equation 5.2 by the terms inside the brackets on the first line of Equation 5.2. In the same way, the relation between medical care usage and income (endogenous in this case) can be ascertained by holding severity of illness and price constant (\( \text{d}s_0 = 0, \text{d}p = 0 \)) and solving the resulting equation. That way, \( \frac{\partial M}{\partial s_0} \) and \( \frac{\partial M}{\partial Y} \) can be obtained. Bearing in mind the following conditions:

\[ 5.3 \quad U_c > 0, U_{cc} < 0, U_{ss} < 0, U_{cs} < 0, U_{ss} = U_{cs}, h' > 0, h'' < 0, p' > 0, p'' > 0; \]

it can be shown, from Equation 5.2, that \( \frac{\partial M}{\partial s_0} \) as well as \( \frac{\partial M}{\partial Y} \) can be either positive or negative. As such, the comparative static analyses are inconclusive and suggest: (a) the demand for curative medical care is not a monotonically increasing function of the severity of a person’s illness, (b) curative medical care is not necessarily a normal good. These inconclusive results are to be expected and are reasonable due to the possible interaction of income and substitution effects arising from the endogeneity of income. That is, a person’s illness interacts with
socioeconomic variables to influence a person's demand for medical care. In an effort to obtain definite results, consider the special case of a person on fixed income, \( Y' = 0 \) and \( Y'' = 0 \) in Equation 5.2. Equation 5.2 becomes:

\[
5.4 \quad -\left\{ (p')^2 \frac{U_{cc}}{U_c} + 2h'p \frac{U_{cs}}{U_c} + (h')^2 \frac{U_{ss}}{U_c} - h'' \frac{U_s}{U_c} - p'' \right\} dM \\
+ \left\{ p' \frac{U_{cc}}{U_c} + h' \frac{U_{cs}}{U_c} \right\} dY \\
- \left\{ p' \frac{U_{cc}}{U_c} + h' \frac{U_{cs}}{U_c} \right\} dp \\
+ \left\{ p' \frac{U_{ss}}{U_c} + h' \frac{U_{ss}}{U_c} \right\} ds_0 = 0 
\]

On each line of Equation 5.4, it can be show that the terms inside the brackets are negative except for the price change on the third line. Therefore, in this simplified model, curative medical care usage is a positive function of income, \( \frac{\partial M}{\partial Y} \geq 0 \), and also a positive function of the severity of illness, \( \frac{\partial M}{\partial s_0} \geq 0 \), but a negative function of price. It follows that it is the effect of illness on endogenous income confounds the results in the more general model above.

6. Conclusion

The status of a person's health usually depends on personal characteristics as well as on the environment and lifestyle, and only in the event of illness or injury does health care become important. In general, a person gets sick first and then seeks care, a sequential action that characterizes curative care. A sick person's medical expenditures are a positive function of the severity of his or her existing illness, all else equal. It is difficult to imagine why a rational person would utilize any medical care if he is not sick and does not expect to be sick. Many wealthy persons utilize low (indeed, zero) curative care because they are not sick. As a result, curative medical care is not necessarily a normal good. This explains the empirically observed low income elasticity of demand for hospital inpatient care. Some studies have even found that income has a negative effect on hospital inpatient care. In order to accentuate a conceptual point, this paper analyzes only curative medical care which is demanded by only the sick. But the dichotomy between curative and preventive care is often difficult to establish empirically because 'preventive' and 'curative' represent motivations for medical care use. These ideas can be formalized by recognizing the intuitive distinction between illness and aging. Yet a sick person's demand for health care would be expected to differ from an aged person's demand. Finally, the paper provides a formal justification for the application of consumer theory to the demand for health care.

References


