A DYNAMIC DEMAND FOR MEDICAL CARE

By Jennifer Leigh Kohn

A dissertation submitted to the

Graduate School – Newark

Rutgers, The State University of New Jersey

in partial fulfillment of requirements

for the degree of

Doctor of Philosophy

Graduate Program in Management

Written under the direction of

Professor Robert H. Patrick

and approved by

Robert H. Patrick

Mahmud Hassan

Sharon Gifford

Howard P. Tuckman

Newark, New Jersey

May 2009

ABSTRACT OF THE THESIS

A Dynamic Demand for Medical Care

By Jennifer L. Kohn

Thesis director: Professor Robert H. Patrick

I develop a theoretical model to explain observed patterns of medical care demand and test the hypothesis that demand is greater the greater the decline in health at any level of health. Medical care demand is highly skewed: the top 5% of individuals consume nearly 50% of expenditures, and nearly half of lifetime expenditures occur after age 65. Extant economic models don't explain this behavior. For example, Murphy & Topel (2006) suggest the willingness to pay for health decreases with age and illness. Grossman (1972) concludes that we demand less health over time, and maintained assumptions about health transition make observed spikes in medical spending unlikely. Tomas Philipson (2007 iHEA plenary) suggested either consumers act irrationally or economists have not adequately modeled behavior. I explore the latter explanation.

I specify an optimal control model that extends the seminal Grossman (1972) model in three ways. I include the change in health in utility; I model depreciation as an amount rather than a rate; and I allow the health state to increase health production. Contrary to the Grossman model, the resulting demand for health suggests an inevitable disequilibrium as health declines between increasing benefits and *declining* costs of health capital that individuals can only balance by increasing medical care. The time path for medical care demand suggests the change in health rather than the state of health drives increasing demand and that price sensitivity declines over time.

I test the central hypothesis that the change in health is significant using the first 14 waves of the British Household Panel Survey (BHPS). I specify a non-linear seemingly unrelated system of demands for consumption and medical care and impose symmetry restrictions on the cross-price parameters so that inferences are consistent with utility maximization theory. I identify instruments for unobservable health and price using a multiple correspondence analysis. I find support for the theory and the assumption that health and wealth are not separable.

Results suggest single period, single equation models of medical care demand omit relevant variables that capture dynamic decision making and the relationship between health and wealth.

Acknowledgements

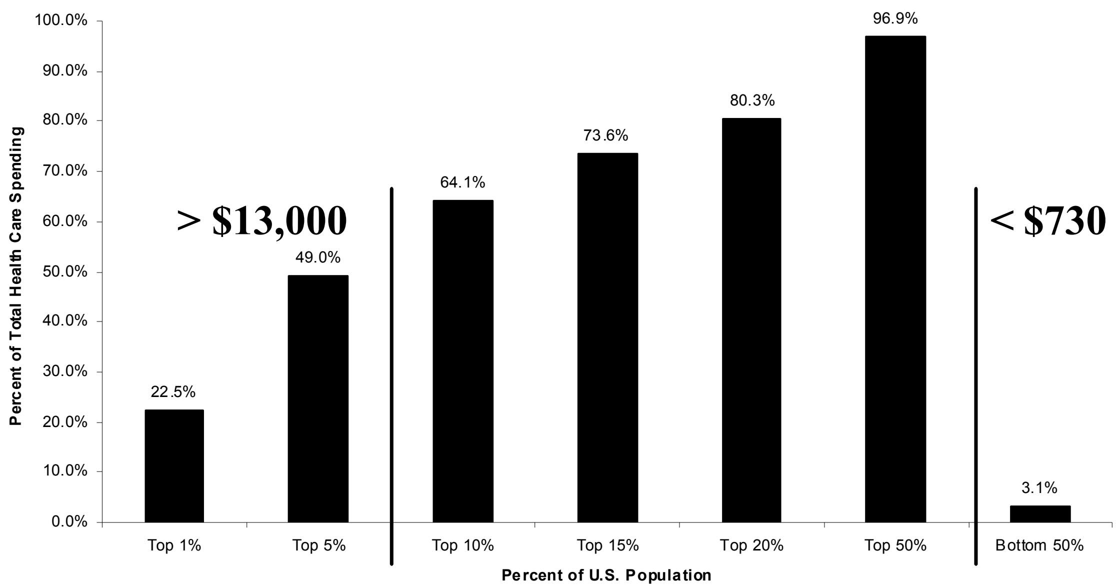
I am indebted to Robert H. Patrick for his immeasurable patience in guiding me through this research and opening up new ways of thinking about interesting problems. I also gratefully acknowledge the rest of my committee, Mahmud Hassan, Sharon Gifford and Howard P. Tuckman, for many valuable conversations and penetrating questions. I would not have been able to complete this work without the support of my husband, David Lagasse, and my children Max and Sam who spent many weekends without me. Finally, I thank my parents Marilyn and Irving Kohn for instilling in me the persistence and expectation of the highest standards that is reflected in these pages.

Table of Contents

List of Figures

Introduction

Medical care spending is nearly 20% of the US GDP[1] and the largest component of aggregate consumer spending.[2] But spending on medical care is highly skewed: the top 5% of spenders incur nearly 50% of expenses.[3] Figure 1: Distribution of U.S. Medical Care Spending (2004)



Reflective of this skewness, medical care expenses are associated with nearly 50% of personal bankruptcies.[4] These snapshot statistics paint a picture of a large market with high risk. Arrow (1968) suggested that "…the special characteristics of the medical care market are largely attempts to overcome the lack of optimality due to the non-marketability of the bearing of suitable risks…"[5] In markets from commodities to consumer credit, even natural disasters, markets manage risk with financial tools such as

[1] Data from U.S. Department of Health and Human Services as cited in Kolata (2006).
[2] As cited in Reinhardt et. al., (2004).
[3] Graph from the Kaiser Family Foundation (2007) with data from the U.S. Department of Health and Human Services, Agency for Healthcare Research and Quality, Medical Expenditure Panel Survey (MEPS) 2004 data. Population includes individuals with no medical care expenditures, and expenditures include payments from all sources (private, insurance, government) but do not include insurance premiums.
[4] Himmelstein et. al. (2005).
[5] Arrow (1968) p. 857.

derivatives and securitization. By contrast, the financial risk in the market for medical care is managed predominantly by large diversified risk pools run either by insurers or governments.

Unlike demands for real estate, consumer products or financial instruments, the parameters underlying the demand for medical care are not well defined either theoretically or empirically. The finiteness of human life and the direct link between a consumer's choice of medical care and his longevity make health care an acute subject on which to test the propositions of economic theory. A *New York Times* article framed the issue poignantly: "do you want to give up steak and martinis to live to 99, and forego dream vacations to finance such longevity?"[6] In this context, medical care is like any other commodity which consumers should demand to the point where the marginal benefit from the resulting health is equal to its marginal cost including the cost of medical care. Moreover, the statistical issues associated with the demand for medical care including latent variables, unobservable heterogeneity and non-linear models make this a particularly interesting area for econometric analysis.[7]

Not surprisingly, high spenders tend to be older and sicker. Almost half of the top 5% of spenders are over 65 years old with limited longevity prospects.[8] Many of the top 5% report poor health and may have low survival prospects even if younger. From a biological perspective it makes sense that the old and sick would spend more on medical care than the young and healthy. However, from an economic perspective, large investments in health when there may be little chance to realize the benefits seems counterintuitive. Indeed, Becker, Murphy and Philipson (2007) argue that "existing

[6] Bruni,(2006) p. 4.
[7] Jones and O'Donnel (2002) p. 1, make this point eloquently.
[8] Conwell and Cohen (2005) p. 1.

estimates of the value of a life year do not apply to the valuation at the end of life."[9] Illustrative of this point, Murphy & Topel (2006) suggest that the higher the probability for survival the greater the willingness to pay for additional longevity,[10] which is inconsistent with high spending among those with low survival prospects. Grossman (1972), the seminal model of individual health demand, concludes that: "biological factors associated with aging raise the price of health capital and cause individuals to substitute away from future health *until death is chosen.*"[11] Grossman suggests that individuals can spend more for medical care while still demanding less health, [12] but the functional form he assumes for health transition makes increasing investment unlikely, particularly to the degree we observe. Ehrlich and Chuma (1990) suggest that an increasing value of life extension can explain increasing medical care expenditures at the end of life.[13] However, this conclusion is also dependent on modeling health decline based on the traditional accounting identity for depreciating assets as is predominant in the literature.

Moreover, the need-based explanation for medical care demand cannot explain the significant differences in demand among individuals with similar medical needs.[14] While the proportion of old and sick is higher among high spenders, over half of these high spenders are still under 65 and two-thirds report better than poor health.[15] There are many competing theories. The demand may be induced by insurance which can reduce

---

[9] Becker, Murphy and Philipson (2007) p. 2.
[10] Murphy and Topel (2006) see in particular equation (3) p. 876 and equation (11) p. 880.
[11] Grossman (1972) p. 240, emphasis added.
[12] Grossman (1972) p. 238.
[13] Ehrlich and Chuma (1990) p. 776.
[14] The Dartmouth Atlas Project has documented extensively the variation in medical care usage across the United States. See http://www.dartmouthatlas.org/.
[15] Conwell and Cohen (2005) p. 1.

the consumer's price of care to zero.[16] Or, it may be driven by physicians who have an economic incentive to induce demand.[17] Still another explanation may be that health is not an individual good, but a societal good whereby family and society may contribute to end-of-life medical care demand.[18] It is likely that all of these play some role in driving high medical care spending. In order to identify the effects of these outside determinants we need a reasonable model of individual demand. A model that can explain the distribution of demand can be used to identify "excess" demand suggested by these other theories.[19]

The contribution of this dissertation is to extend extant models of medical care demand and test the central implied hypothesis. I offer an alternative explanation in a neo-classic economic framework for why high spenders spend so much. Specifically, I show that the change in health and the way depreciation is modeled drive inferences about the demand for medical care. A critical implication is that individuals will demand more medical care the greater their decline in health *at any level of health*. A key assumption is that health and consumption are not separable. Empirical tests support both the hypothesis that the change in health is relevant to the demand for medical care and the non-separability assumption between health and consumption. The theoretical and empirical findings suggests that single equation, single period models of the demand

---

[16] See in particular Keeler, Newhouse and Phelps (1977) and other studies emanating from the Rand Health Insurance Experiment.
[17] See Reinhardt (1985) for an overview.
[18] The 2005 case of Terri Schivo is an extreme example of demand by social institutions rather than by individual choice. Also see Bolin et. al., (2002) for an extension of the Grossman model to incorporate strategic interactions among family members. Becker et.al. (2007) refer to such socially induced demand as "altruism."
[19] I draw a parallel here with the use of the Capital Asset Pricing Model (CAPM) to identify "abnormal" risk-adjusted market returns.

for medical care omit relevant variables that capture dynamic decision making and the relationship between health and wealth.

The dissertation has two parts. Part I specifies the theoretical model and draws out the testable hypotheses. Section I places the dissertation within the extensive literature on the demand for health and medical care. Section II specifies the model, and Section III draws out the testable hypotheses from the equilibrium demand for health and time paths of medical care and consumption. Part II empirically tests the central hypothesis that the greater the decline in health the greater the demand for medical care at any level of health. Section I briefly reviews the empirical literature with a focus on demand for medical care models and econometric issues. This section ends with a review of the literature to proxy for unobservable health and price of medical care and my justification for using a multiple correspondence analysis. Section II specifies the estimating equations and hypotheses. Section III describes the data including empirical results for the multiple correspondence analyses for health and price. Section IV discusses the estimation strategy and Section V presents the empirical results for the systems of equations and robustness checks with single-equation models. Part II concludes with a discussion of these results. The dissertation concludes with a summary of contributions and areas for future research. The appendices contain a list of symbols, sufficiency conditions, derivations of the equilibrium condition and time paths, a detailed list of BHPS variables, and tables of the empirical results.

Part I: A Dynamic Demand for Medical Care

I. Theoretical Literature Review

The value of medical care is the value of the health that is produced. Therefore, the key theoretical question is: what is the value of health? Viewed as any other asset, the value of health is the discounted present value of future benefits. While some argue that you cannot put a price on health, limited resources force both individuals and policy makers to confront this uncomfortable valuation task in order to make necessary tradeoffs between investment in health and investment in competing projects. This further frames the question: what are the benefits from health? The literature shows that how the benefits of health are defined and how these benefits are reflected in the consumer's optimization problem changes the determination of value.

There are two major branches of this literature: the human capital models based on Grossman's seminal 1972 work "On the Concept of Health Capital and the Demand for Health" and the willingness to pay models which focus on the trade-off between assets and the risk of death.[20] Both branches of the literature generally conclude that the value of health declines with age because individuals have less time to reap the benefits from their investment. Yet the old and the sick make up a significant proportion of the top 5% of spenders; and extant models offer no explanation for high spending among the relatively younger and healthier. This review first takes a historical look at health

---

[20] There is another branch of this literature that focuses almost exclusively on the price-elasticity of medical care demand. Two key theoretical works are Keeler et al (1977) and an extension by Ellis (1986). These models generally follow Grossman's derived demand specification for medical care and are not separately reviewed here.

valuation models and then more specifically compares and contrasts the human capital and willingness to pay approaches.

A. A Brief History of Health Valuation Models.

Some date health valuation models to 17th century England,[21] but this review will begin in early 20th century America. The early models were commonly referred to as "human capital" models, but they were little more than actuarially projected discounted earnings. Earnings were considered quite literally: "early estimates incorporated a zero value for persons without labor income."[22] Imputing a value for women's housework was considered a major innovation. The primary difficulty with these models was to determine an appropriate discount rate, and estimates of the value of life ranged widely and were often applied opportunistically.[23]

Frustration with the lack of theoretical grounding for the early human capital models inspired the "willingness-to-pay" (WTP) models. WTP is the marginal rate of substitution between a decrease in mortality and an increase in consumption derived from a representative consumer's optimization problem. Mishan (1971) provided one of the first conceptualizations of the WTP approach:

> Consistency with the criterion of a potential Paeto improvement and, therefore, consistency with the principle of evaluation in cost-benefit analyses would require that the loss of a person's life be valued by reference to his CV (compensating value); by reference, that is, to the minimum sum he is prepared to accept in exchange for its surrender.[24]

---

[21] See Petty (1699).
[22] Landefeld & Seskin, (1982) p. 556
[23] Mishan (1971) p. 693.
[24] Mishan (1971) p. 693.

It was instantly clear to Mishan that "in ordinary circumstances, no sum of money is large enough to compensate a man for the loss of his life."[25] Rather, the WTP models take a position of a general rather than a specific life. Tolley et. al., (1994) in their comprehensive review of WTP suggest that "anonymity is a way of rendering life-and-death decisions a matter of consumer choice."[26]

In 1972, Michael Grossman changed the definition of human capital as it pertains to health and medical care. His seminal work, "On the Concept of Human Capital and the Demand for Health" has been cited by nearly 1,400 subsequent works.[27] Like the early WTP models, Grossman grounded his valuation of health in the consumer's utility maximization problem. To address those in and out of the labor market, Grossman applied the wage rate to total time, not just labor time. In addition, he uniquely defined a "health production function" that extended the seminal work in human capital by Becker (1964) and Ben-Porath (1967). Notably, Grossman's production function included education, which was a primary focus of Becker's work.

It is interesting to consider other historical movements that influenced the two major branches of health valuation literature. While Grossman was influenced by societal and academic interest in the value of education, the WTP scholars drew inspiration from the environmental movement. Notably the first Earth Day occurred in 1969, barely two years prior to the Mishan quote above. Much of the environmental debate at the time was about societies' "willingness to pay" for pollution remediation and resource conservation. Similarly, many examples in the WTP health valuation literature reference environmental factors. Berger et. al. (1987) offer two examples of the health

---

[25] In this observation Mishan quoted the other early WTP scholar, T.C. Schelling (1968).
[26] Tolley et. al. (1994) p. 5.
[27] Google Scholar search as of 1/8/09

outcomes they are interested in: "the occurrence of a specified type of cancer as affected by environmental irritants" and "traffic accidents due to poor visibility brought on by air pollution…"[28]

Finally, the late 1960's saw the first post World War II spike in medical spending and sparked a wave of research into medical care demand. The most ambitious was the Rand Health Insurance Experiment (RHIE) which tracked the health and medical care usage of 2,000 families who were randomly assigned to different insurance plans for three or five years. The Rand study investigators published dozens of articles both during and after the experiment, and the data from the experiment has been used in hundreds of subsequent studies.[29] However, the vast majority of this research has been empirical and has not tested explicitly hypotheses derived from a theoretical model of medical care demand.[30] To the extent that the studies using the RAND data use underlying theoretical models, they generally reflect Grossman's assumptions.[31] This empirical literature will be reviewed in Part II of the dissertation.

There have been many extensions of the Grossman model. Dardanoni (1986) and Wagstaff (1986) offer simplifications; Koc (2004a) restrictions; and Cropper (1977), Chang (1996), Ehrlich (2000) and Liljas (2005) offer various stochastic extensions, just to name a few. All of these extensions preserve the key elements of Grossman's utility

[28] Quoted from Tolley et. al (1994) p. 29.

[29] The first book from the Rand study was Manning et. al. (1988) *Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment.* A later book: Newhouse et. al. (1993) *Free for All?: Lessons from the Rand Health Insurance Experiment* has been cited by 390 subsequent works, and the main journal article, Manning et. al. (1987) has been cited by nearly 700 according to Google Scholar, 1/8/09.

[30] A Federal Reserve Bank of San Francisco Economic Letter "More Life vs. More Goods: Explaining Rising Health Expenditures" May 27, 2005 commented on a 1992 Newhouse paper that "by itself, …Newhouse's story is incomplete. People do not have to purchase the new medical technologies if they don't want to…" (p. 2) In other words, Newhouse's framework does not incorporate consumer preferences in a coherent way.

[31] See in particular Keeler, Newhouse and Phelps (1977) which was one of the theoretical works that provided a foundation for the RAND study.

and health transition specifications that I generalize. The next section will detail the Grossman and Ehrlich and Chuma (1990) models on the one hand and the Murphy & Topel (2006) model on the other (hereafter referred to as E&C and M&T). I will adopt E&C's optimal control framework and draw support for my specification of health from the M&T model.

B. Human Capital and Willingness to Pay Models

The two branches of the literature pose distinctly different research questions and make different assumptions about the benefits from health. Each section will discuss these differences with particular emphasis on how the different assumptions about health benefits are incorporated into each optimization model. I will conclude each section with a more detailed analysis of the resulting health valuation from each model.

1. Grossman (1972): The Demand for Health

In the abstract to "On the Concept of Health Capital," Grossman states: "The aim of this study is to construct a model of the demand for the commodity "good health."[32] Grossman's primary contribution to the field was to characterize the demand for medical care as a derived demand. In other words, the object of consumer choice is not a doctor's visit, or drug, or heart by-pass procedure, but the commodity "good health." In marketing terms, consumers demand the sizzle, not the steak.

[32] Grossman (1972) p. 223.

Grossman defines the benefits from health narrowly in terms of "healthy days." He defines a production function for healthy days: $h_t = \phi(H_t)$ [33] which is increasing and concave in health. However, it is not always clear whether the maximum number of healthy days is the number of days in a year, 365, or the number of days in a lifetime. In other words, in Grossman's specification it is difficult to identify the effect of health on incremental longevity. E&C address this as discussed in detail below.

Grossman defines an intertemporal utility over healthy days, h, and consumption, Z. His model is in discrete time employing the calculus of variations where n is an endogenously determined length of life.[34]

$$U = U\left(\phi_0 H_0, ..., \phi_n H_n, Z_0, ..., Z_n\right)$$

In empirical work Grossman and others following the Grossman model assume strict separability between health and consumption and estimate either a pure consumption or pure investment model of health. Grossman does not maintain separability in his theoretical work; yet the cross-partials of utility play no role in his theoretical implications.

Grossman defines the health transition and health production functions as follows.

$$H_{t-1} - H_t = I_t - \delta_t H_t$$
$$I_t = I\left(M_t, TH_t; E_t\right)$$

The change in health is a function of investment, I, minus a rate of depreciation, δ, multiplied by the state of health. Health investment is a constant-returns-to-scale function of the market input of medical care, M, and the time spent on producing health,

---

[33] I am replacing Grossman's index for time, i, with t for clarity in this exposition.
[34] The issue of whether or not Grossman endogenously determines the end of life will be further discussed in the context of the Ehrlich and Chuma critique on this issue.

TH, given the level of education, E. As mentioned in the history of health value models, This functional form for health transition is consistent with the accounting identity for depreciating assets at the time; but it is no longer consistent with accounting for intangible assets such as health. This functional form has significant implications for the user cost of health capital and the equilibrium demand for health and will be a central element of the dissertation.

Grossman defines the full wealth constraint as follows:

$$R_t = \sum \frac{P_t M_t + V_t X_t + W_t \left(TL_t + TH_t + T_t \right)}{\left(1+r\right)^t} = \sum \frac{W_t \Omega}{\left(1+r\right)^t} + A_0$$

Where P is the price of medical care, V is the price of market goods, X, used in the production of commodities, Z, W is the exogenous wage rate per unit of time, TL is time lost to sickness, TH is time spent on health and T is time spent on producing market goods. Ω is the total time available and A is discounted property income. Thus, Grossman's wealth constraint states that "full wealth equals initial assets plus the present value of the earnings an individual would obtain if he spent all of his time at work."[35] Again, this was an innovation over other so called human capital models in that it included a value for non-work time.

In the 1972 exposition Grossman does not explicitly define the initial conditions for health and wealth, but implies that H_o is greater than H_{min}, which is the terminal health capital (or else the problem would be trivial) and A_o is greater or equal to zero. Grossman does not explicitly provide a terminal condition for wealth, and he does not discuss the issue of debt in either the 1972 or the 1999 papers. In the 1972 paper

[35] Grossman (1972) p. 228.

Grossman states the terminal condition for health as $H_t \le H_{\min}$; however in the 1999 paper he clarifies that $H_t > H_{\min}$ and $H_{t+1} \le H_{\min}$.[36] This distinction is relevant for the endogenous determination of terminal time in the discrete time formulation; but is not relevant in continuous time.

The central theoretical implication of the Grossman model is seen in the equilibrium demand for health which sets the marginal benefit of holding a unit of health capital (on the left-hand side) equal to the marginal user cost of health capital (right-hand side.) G is the marginal change in healthy days: $G_t = \partial h_t / \partial H_t$, π is the marginal cost of health investment, and with the tilde it is the percent change over time.

$$G_t \left[W_t + \left(\frac{Uh_t}{\lambda} \right) (1+r)^t \right] = \pi_{t-1} \left(r - \tilde{\pi}_{t-1} + \delta \right)$$

From this equation Grossman concludes that as the rate of depreciation increases individuals reduce their demand for health to maintain equilibrium with a rising cost of health capital. He maintains that "even though health capital falls over the life cycle, gross investment might increase, remain constant, or decrease" because the increase in depreciation changes both the right-hand-side marginal cost as well as the left-hand-side marginal benefits G_t. If an increase in the rate of depreciation decreases health such that the marginal benefits from health are higher than the marginal cost, then individuals will invest more in medical care to increase health and decrease the marginal benefits to regain equilibrium. This result will be explored in detail when contrasted with the equilibrium demand for health from the dynamic lifecycle model. Finally, from this

[36] Grossman (1999) fn 9 p. 81.

equilibrium demand for health Grossman argues that individuals ultimately chose death by choosing a level of health equal to H_{min}.

2. Ehrlich and Chuma (1990): The Demand for Health and Longevity

Ehrlich and Chuma (1990) extend Grossman's model by explicitly specifying "a demand function for longevity, or 'quantity of life,' along with corresponding demand functions for indicators of 'quality of life'…"[37] Their focus on the terminal time of death leads them to specify a continuous-time optimal control model which explicitly defines the terminal time as a choice variable and derives transversailtiy conditions to define the parameters of this choice. Beyond this choice of time perspective and methodology, the only change E&C make to the Grossman model is to assume a diminishing rather than constant returns-to-scale health production function. They do this to ensure an interior optimum for medical care demand.[38] E&C's assumption of diminishing returns to health production is reflected in the marginal cost of investment included in the wealth transition function with α as the exogenous productivity parameter. All other notation is the same as Grossman's except A is the state of wealth, ψ denotes the marginal cost of investment in commodities, ρ is the individual's time preference, λ_A, λ_H are the costate variables for wealth and health respectively, and m rather than TH denotes the time spent on health production.

$$U = \int_0^T e^{-\rho t} U(Z(t), h(t))dt$$

[37] E&C (1990) p. 761.
[38] See E&C (1990) p. 768 and Grossman's response to the E&C critique on this point in Grossman (1990) p. 17.

$$\dot{H}(t) = I(t) - \delta(t)H(t)$$
$$I(t) = I(M(t), m(t); E(t))$$

$$\dot{A}(t) = rA(t) + w\phi(H(t)) - \pi I(t)^{\alpha} - \psi Z(t)$$

Finally, E&C explicitly state the initial and terminal conditions for the state variables:

$$H(0) > H_{\min} > 0$$
$$H(T) = H_{\min}$$
$$A(T) \geq 0 \text{ if T is finite}$$

E&C do not state an initial condition for A(0); however, if A(T) is non-negative than A(0) cannot be negative without a lower bound. The key point is that E&C maintain the same utility function, health transition and health production functions as Grossman and these drive the major theoretical implications of their model.

E&C's equilibrium demand for health is similar to Grossmans except the marginal cost of medical care, π, is replaced by the marginal value of health capital, g(t). In E&C this is defined as the ratio of the shadow price of health to the shadow price of wealth: $g(t) = \lambda_H(t)/\lambda_R(t)$:

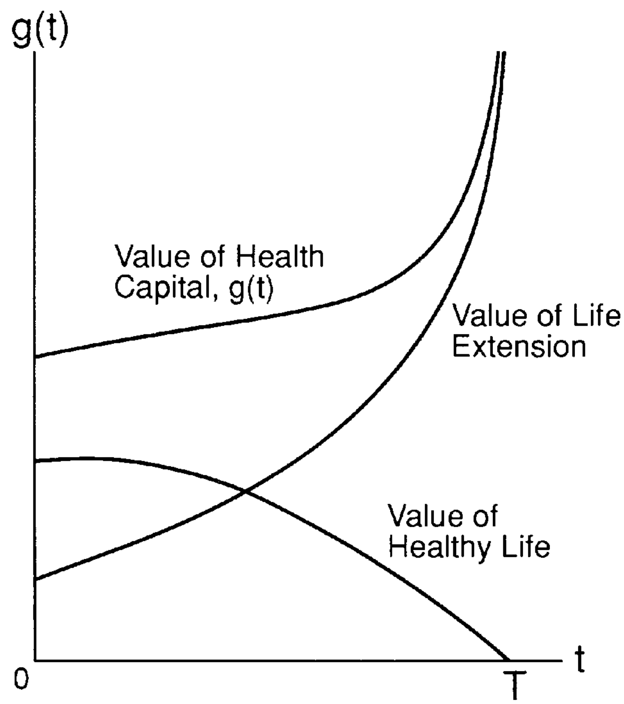
$$\left[\frac{U_h}{\lambda_R(0)} e^{(r-\rho)t} + w\right]\phi^{'}(H(t)) = g(t)\left[\delta(t) + r - \tilde{g}(t)\right]$$

The implication is the same as Grossman's: as depreciation increases the marginal cost of health capital increases requiring individuals to decrease their demand for health capital to increase the marginal benefits from health and regain equilibrium. The difference is that this equilibrium condition does not offer a "myopic rule" for investment in health because the shift in depreciation (and other parameters) may change the value of health capital over the planning horizon.[39] Instead, E&C turn to the expression for the value of health capital to draw out the key theoretical implications of their model.

$$\frac{\lambda_H}{\lambda_A} = g = g(T)e^{-\int_t^T [\delta(u)+r(u)]du} + \int_t^T \left\{ \left[\frac{1}{\lambda_A(0)} U_h(u) e^{[(r-\rho)u} + w \right] \phi'(H(u)) e^{-\int_t^u [\delta(s)+r]ds} \right\} du$$

In words, the value of health capital is equal to the ratio of the shadow price of health to the shadow price of wealth. This expression is similar in principal to the willingness to pay concept of a change in mortality given a change in wealth. The first term on the right hand side represents the discounted value of terminal health, g(T).[40] Because the integral declines as t approaches T, this value increases with time, as depicted in the graph below.[41]

Figure 2: Ehrlich & Chuma (1990) Value of Health Capital Over Time



[39] See E&C (1990) p. 769. This results from the simultaneous determination of the state, control and co-state variables in optimal control.
[40] E&C suggest that Grossman does not derive this component of marginal benefit because he does not endogenously determine the terminal time.
[41] E&C (1990) p. 777.

E&C suggest that this increasing component of the value of health "provides a partial explanation for the empirically observed tendency of health care expenditures to rise even over the more advanced phases of life".[42] However, it is critical to note that the terminal value of health capital increases only because the addition of the rate of depreciation makes the integral in the exponent unambiguously positive. This follows directly from the specific functional form of a rate of depreciation multiplied by the stock of health. This can also be seen in E&C's expression for the change in the value of health capital.

$$\dot{g}(t) = g(t)[\delta(t) + r] - \left\{ \left[\frac{1}{\lambda_A(0)} \right] U_h(t) e^{[(r-\rho)u]} + w \right\} \phi'(H(t))$$

Again, the first term represents the value of life extension, and it is always positive since all of its elements are assumed positive. E&C's depiction suggests that the terminal value increases at an increasing rate, but this depends on the relative changes of g and δ.

3. Murphy and Topel (2006): The Value of Health and Longevity

Murphy and Topel's main contribution to the WTP literature was to explicitly distinguish two benefits from health: H(t), which "raises the quality of life" and G(t) which affects mortality.[43] This helped to bridge the historical WTP focus on mortality risk with the human capital focus on health in a similar way as E&C added the focus on longevity to the Grossman model.

Murphy & Topel's utility function is as follows:

---

[42] E&C (1990) p. 776. E&C also make an interesting point that, in addition to a myriad of statistical problems with health expenditure data, "aggregate statistics concerning health care outlays by age are subject to a 'natural selection' bias since those surviving to older ages should be those with higher initial health and wealth endowments, which are expected (according to their model) to raise the value individuals ascribe to health capital and, thus, their demand for voluntary health investments" fn. 9 p. 776.

[43] M&T (2006) p. 876.

$$U = \int_a^\infty H(t)u(c(t), l(t))\tilde{S}(t,a)e^{-\rho(t-a)}dt$$

Utility is the sum of the discounted value of consumption and leisure *multiplied* by the state of health, H(t). In other words, if the individual has perfect health with value 1, then she enjoys 100% of her utility. However, if her health is impaired, say .5, then she will only enjoy 50% of her utility. The upper bound on the integral is infinity, implying an infinite lifespan is possible, but of course no one lives forever. The terminal time of death is determined when the "survivorship function," $\tilde{S}(t,a)$, the probability that the consumer is alive to enjoy the utility, goes to zero. The second component of health, G(t), is contained within this survivorship function:

$$\tilde{S}(t,a) = \exp[-\int_a^t \lambda(\tau, G(\tau))d\tau]$$

M&T are interested in the change in the value of health for an *exogenous* change in the risk of mortality. M&T acknowledge that health is affected by "individuals' choices", but they relegate these choices "to the background."[44] Therefore, they do not need to specify a health transition function or health production function or specify in detail how the survivorship function trends to zero. They do need a wealth constraint which is specified as follows:

$$A(a) + \int_a^\infty [y(t) - c(t)]\tilde{S}(t,a)e^{-r(t-a)}dt = 0$$
$$y(t) = w(t)[1 - l(t)] + b(t)$$

The first equation states that individuals must consume all assets plus the sum of lifetime earnings, y(t), minus consumption, c(t). The second equation specifies that lifetime earnings are a function of the wage rate, w, which is exogenous, multiplied by non-

[44] M&T (2006) p. 876.

leisure time. Therefore, consumers choose earnings based on their choice of leisure. The b(t) represents non-wage income such as pensions. It is meant to reflect income in retirement, not transfer payments associated with the state of health. The only role for health in the budget constraint is embedded in the survivorship function, which, like in the utility function, determines the terminal point of the integral.

The key implication of the M&T model is reflected in the expression for the value of health expressed as the value of remaining life:[45]

$$V_{\lambda}(a) = \int_{a}^{\infty} \left[\frac{H(t)u}{H(t)u_c^{'}} + y(t) - c(t) \right] e^{-r(t-a)} \tilde{S}(t,a)dt$$

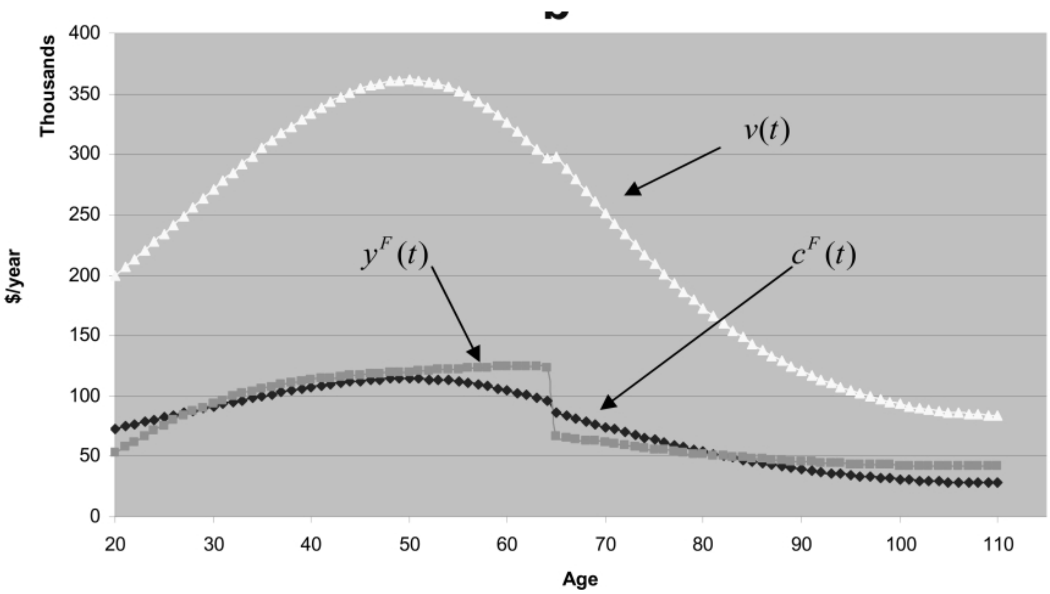
The bracketed term is the value of a life year, and the sum of the discounted value of life years equals the value of the remaining life at age a. The bracketed term begins with the ratio of utility to marginal utility of consumption, which represents the monetary value of instantaneous utility. Because the quality of life, H(t) is incorporated multiplicatively, it affects both the utility and marginal utility proportionally. Therefore, the terms cancel and the quality of life plays no role in the calculation of the value of remaining life.[46] As t increases, the survivorship probability decreases and as a result the total value decreases over time as depicted in the graph below:

---

[45] I have substituted M&T's equation (8) into (7) to clearly show how H(t) cancels out of the ratio in this calculation. See M&T (2006) p. 11.

[46] M&T argue "H is valuable…yet willingness to pay for additional life-years does not depend on H." They cite an Environmental Protection Agency study: "There are no published studies that show that persons with physical limitations or chronic illnesses are willing to pay less to increase their longevity than persons without those limitations." (Both quotes from p. 878.) I would argue to the contrary, that the often exorbitant prices people with chronic illnesses pay for pharmaceuticals just to live suggests that people with lower health are willing to pay *more* to increase their longevity because not taking the drugs would clearly increase their mortality risk. (See in particular, Anand, G., "Through Charities, Drug Makers Help People – And Themselves," December 1, 2005; "A Biotech Drug Extends a Life, but at What Price?" November 16, 2005; and "How Drugs for Rare Diseases Became Lifeline for Companies," November 15, 2005, all in *The Wall Street Journal* p. A1.

Figure 3: Murphy and Topel (2006) Value of a Statistical Life Year



The critical implication is that while the quality of life, H(t), plays no role in the value of remaining life, it does affects the value of a life year. The result is a value of a life year that changes with age and the level of health. Prior to M&T all people of any age and health status were assumed to have the same value of a life year.[47] Their expression for the change in the value of a life year is:

$$\dot{v}(t) = \frac{y}{v}\left[s_w\dot{w} + (1-s_w)\dot{b}\right] + \left[1 - \frac{y-c}{v}\right]\left[\dot{H} + r - \rho\right]$$

The first term represents the lifecycle of earnings, and the second term represents the lifecycle of consumption. Health is assumed to decrease over time and therefore the value of the second term also decreases over time.

M&T's specification for utility imposes the assumption that health and consumption are complements: $U_{cH} = U_{Hc} = U_c > 0$. Many of their conclusions are driven by this assumption. In the time path $\dot{v}(t)$ the complementarity between health and consumption implies that: "persons with declining health are, in effect, more

---

[47] A common synonym is value of a statistical life.

impatient."[48] M&T calibrate the time path of health in their model by linking it to the observed time path of consumption to get the shape of the value of a life year illustrated in the figure above. While I include a quality of life component to health consistent with M&T, I do not assume the relationship between health and consumption: it is left as an empirical issue and further explored in Part II of the dissertation.

II The Dynamic Demand for Medical Care Model

A. Model Specification

This model is firmly within the human capital branch of the literature but reflects the more general benefits from health assumed in the willingness to pay approach. : The model extends the Grossman (1972) model by generalizing the specification of the utility and health transition functions which allows for expanded benefits from Health and medical care. Since my focus is on the demand for medical care, I follow E&C's optimal control framework with medical care and consumption as controls, health and wealth as states and the terminal time of death determined endogenously. The assumptions associated with the utility function, health transition, wealth transition and endpoint conditions of the model are discussed in turn.

1. The Utility Function

An individual's lifetime utility (*LU*) is defined to be the discounted utility from consumption, $Z(t) \geq 0$, the state of health, $H(t) \geq H_{\min}$, and the change in the state of

[48] M&T (2006) p. 879.

health, $\alpha(t)$, over all $t \in [0,T]$, T being the individual's endogenously determined lifetime:[49]

$$LU \equiv \int_0^T e^{-rt} U\left[Z(t), H(t), \alpha(t)\right] dt. \tag{1}$$

This specification generalizes Grossman's by including the state of health rather than a specific flow of healthy days in the utility function. The willingness-to-pay literature provides support for this more general specification reflecting a quality of life component to utility:

> When investigating how workers and consumers make choices regarding risks to health, it is important to recognize that *the utility individuals derive from consumption depends on their state of health.*[50]

In other words, health is relevant not only because it provides time for productive activity, but also because it enhances the utility from consumption. This is consistent with M&T's treatment of health in utility. I generalize their multiplicative form to preserve the ability to empirically test the role of quality of life in the value of health and the relationship between health and consumption.

I extend Grossman's model by including the change in health, $\alpha(t)$ in the utility function. The intuition is that the utility associated with health is relative rather than absolute. For example, an individual with say health = 100 will get a different amount of utility from the same level of health if her change in health is positive (negative); if she is relatively healthier (sicker) than before. In other words, the choice of medical care and

[49] The stated model does not include a bequest function. Adding a bequest function increases the shadow value of wealth along the planning horizon and thereby decreases the marginal value of health capital resulting in a lower level of demand for health. However, none of the model's implications drawn from the equilibrium demand for health or time paths for medical care or consumption are changed by adding a bequest function.

[50] Berger et. al. (2004) p. 26, emphasis added.

consumption is conditional not only on the state of health, but also on the change in the state of health.

Including the change in health in the utility function follows the significant literature in economics on habit formation and social status.[51] More specifically, in the health economics literature this follows Groot (2000) and Gjerde et. al. (2005) who model the frequently observed phenomenon that "chronically ill patients generally report levels of quality of life that are much higher than one would expect given their condition."[52] Gjerde et al. suggest that "the change in health rather than the absolute health matters" and specify a particular functional form for "adaptation." [53] My specification preserves a role for the level of health independent of the change in health and does not impose a specific functional form on an individual's adaptive process. The specification is general in that it allows asymmetry in the utility from improving and declining health.

The more interesting case is when health declines since a substantial proportion of high spenders are in advanced ages and/or poor health. Therefore, I consider $\alpha(t) < 0$, that the change in health is negative, in most of the discussion. In this context, it is not the direction of the health change (positive or negative), but the magnitude (small or

---

[51] Constantinides (1990) offers a short review of the literature on habit formation.

[52] Groot (2000) p. 403. Sacket & Torrance (1978) and Cassilleth et al. (1984) among others in the medical literature offer empirical evidence of individual's adaptation to declining health. Cassilleth et. al. conclude that, "psychological adaptation among patients with chronic illness is remarkably effective and fundamentally independent of specific diagnosis," p. 506.

[53] Gjerde et al (2005) p. 1284. Gjerde et al put a specific adaptation function, K(t), into the utility function. K(t) is a function of H(t); however, they do not reflect the marginal utility of health in their necessary conditions, see Appendix A equation A(2) and A(3). As a result, they conclude that the co-state on health can be negative, in violation of the sufficiency conditions for a maximum (see Caputo (2005) Corollary 3.1 p. 55). This leads to a counterintuitive conclusion that if individuals can adapt to changing health then they will demand less health to avoid future declines in health: "A person who is able to adapt to a lower health level*, would presumably try to avoid a high H* (level of health) in an early stage of life that would give high cost in terms of a large fall in H in later periods." p. 1289, emphasis added.

large) that is relevant to utility. When health declines, the assumption of a positive partial derivative, $U_\alpha > 0$, means that individuals prefer a smaller negative change to a larger one.[54] For example, an investment in health may not completely re-coupe health that is lost to depreciation, so the change in health will be negative but *less negative* than it would have been without any medical care. The implication is that individuals have a greater ability to adapt to declining health if the decline is gradual rather than sudden. The utility function is considered concave in its other elements as well: $U_Z > 0; U_H > 0; U_{ZZ} \leq 0; U_{HH} \leq 0; U_{\alpha\alpha} \leq 0$.[55]

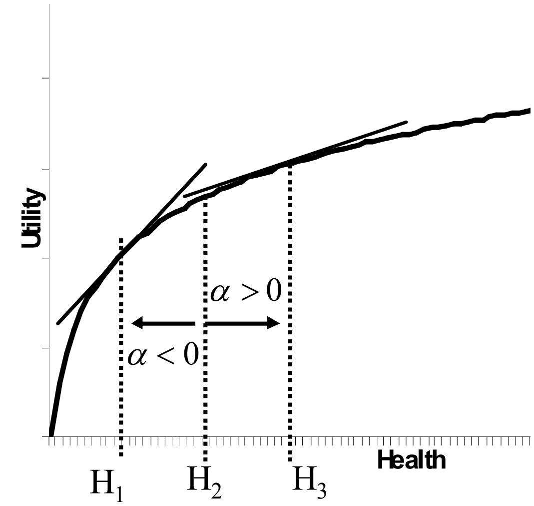
While the sign of $U_{\alpha Z}$ must be determined empirically, I assume the sign of $U_{\alpha H}$ is negative. Since α is defined as the change in the health state, the second cross-partial derivative of utility with respect to health and the change in health is a movement along the concave utility curve with respect to health as depicted in Figure 4 below. An increase (decrease) in health decreases (increases) the marginal utility of health. The same holds for a change in health that is negative but small compared to a change that is more negative. We can also think of the change in health changing the shape of the utility curve with respect to health. $U_{H\alpha} < 0$ implies that the marginal utility from health is hypothesized to be greater (lower) when health declines (increases). Intuitively, this suggests that individuals are better able to withstand a decline in health at higher levels of health than at lower levels of health.

---

[54] The partial derivative would (of necessity) be of the same sign if $\alpha(t) > 0$: a larger positive change in health is preferred to a smaller positive change or a negative change.

[55] Subscripts represent derivatives, e.g., $U_Z = \partial U / \partial Z$, $U_{ZZ} = \partial^2 U / \partial Z^2$, $U_{ZH} = \partial^2 U / \partial Z \partial H$, etc. A full derivation of the model's sufficiency conditions is presented in the Appendix.

Figure 4: Marginal Utility of Health With Respect to the Change in Health



The alternative assumption that $U_{H\alpha} > 0$ would require that the utility of health become flatter with a decline in health. This would suggest that individuals become more indifferent to different levels of health as health declines, which is not an intuitively reasonable assumption.

2. Health Transition Function

The change in health over time is a function, α, of time, t, medical care, $m(t) \geq 0$, the state of health, $H(t) > H_{\min}$, and an exogenous amount of depreciation, $\delta(t) \geq 0$ [56] which is assumed to increase over time $\dot{\delta}(t) \geq 0$.[57]

$$\dot{H} = \frac{\partial H}{\partial t} = \alpha\left[t, m(t), H(t), \delta(t)\right] \tag{2}$$

[56] Depreciation can also be considered a negative value, in which case it would be more negative over time $\dot{\delta} \leq 0$. This would make no difference to the sign of the time path of medical care or any other results.
[57] Prior versions of the model specified a depreciation function that is endogenous with the choice of medical care and/or the state of health. This merely added component(s) to the marginal benefit from medical care, but did not fundamentally alter the conclusions.

Modeling depreciation as an amount rather than a rate is a critical departure from the literature.[58] Recall that in the Grossman model the change in the stock of health is the difference between investment, I(t) and the *rate* of depreciation multiplied by the existing stock: $\dot{H}(t) = I(t) - \delta(t)H(t)$. This follows traditional accounting for depreciating assets. However, in 2001 the U.S. Financial Accounting Standards Board (FASB) issued statement #142 changing the accounting for *intangible* assets. Intangible assets are no longer amortized (at constant or varying rates); rather, they are tested periodically for impairment and written down accordingly. In justifying the change, the Statement noted:

> …financial statement users…indicated that they did not regard goodwill amortization expense as being useful information in analyzing investments…(and) the change…will better reflect the underlying economics of those assets.[59]

Similarly, I will show that modeling depreciation as an amount better reflects the way health declines and is more useful in analyzing investments in health.

First, advances in medical technology over the 20th century appear to have changed the aggregate experience of health decline over the lifecycle. Figure 5 represents the percent of people still living at each age in 1900 and in 2003. In effect it reflects the aggregate lifecycle of health for a course measure of H_{min} (dead) or $H>H_{min}$ (alive.) In 1900 (after the initial drop reflecting high infant mortality) the time path is more linear; the 2003 path is more concave. The 2003 time path suggests that more people remain healthy for a longer period, but the death rate – the aggregate decline in health – is much higher in later years than a century earlier. This reflects the fact that

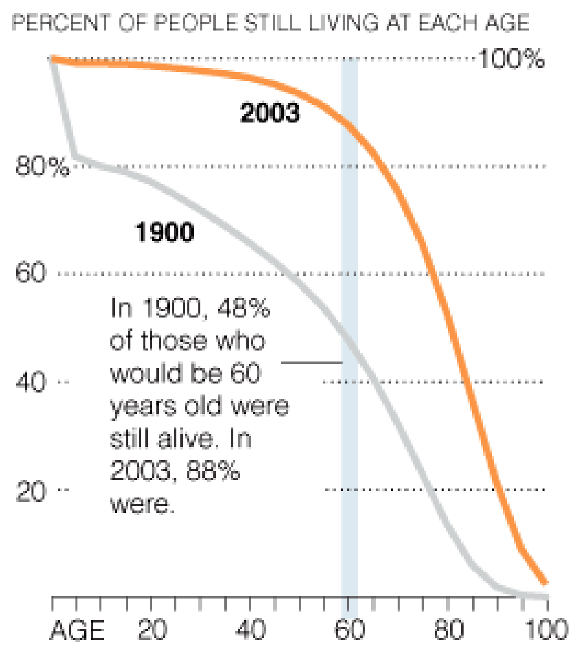
---

[58] Keeler et. al (1977) model health shocks *l* rather than a rate of health depreciation; however, their analysis is in terms of a monetary loss *l** and they do not specify a function mapping *l* to *l**. Ellis (1986) specifies both a health shock ε as well as a "decay rate of poor health" ρ (see p. 164.) Both papers focus on a marginal rate of substitution between medical care and other consumption with a strictly separable utility function between health and consumption and with prior period health only relevant to the accumulated use of medical care affecting the current period price.

[59] http://www.fasb.org/st/summary/stsum142.shtml

while we have been able to extend the duration of good health for many, we have not been able to extend the total length of life. [60]

Figure 5: Time Path of Aggregate Health Capital 1900 and 2003[61]

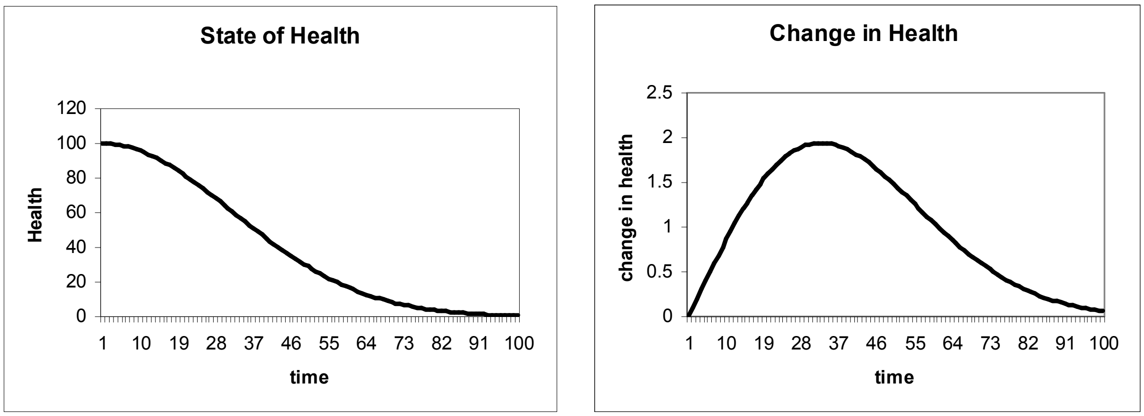


Assuming that health declines over one's lifetime, this graph suggests that a lower (higher) level of health is associated with a more (less) negative rate of change in health, or $\dot{H}_H \geq 0$. By imposing a multiplicative form of depreciation, δH, the Grossman model assumes the opposite: $\dot{H}_H = -\delta \leq 0$. In other words, the Grossman model suggests that the higher the state of health, the larger the negative change in health at any given rate of depreciation. As depicted in figure 6 below, the multiplicative functional form for depreciation implies an asymptotic decline of health at the end of life and not the steep

[60] See Nuland (1993): "Though biomedical science has vastly increased mankind's average life expectancy, the maximum has not changed in verifiable recorded history. In developed countries only one in ten thousand people lives beyond the age of one hundred" p. 85.
[61] This graph is from http://www.nytimes.com/imagepages/2006/07/30/health/20060730_AGE_GRAPHIC2.html as part of the *New York Times* article from Kolata (2006).

decline suggested by the longevity data in figure 5. The amount of health decline decreases at the end of the lifecycle even if the rate of depreciation increases and the largest amount of health decline will occur prior to the end of life. Changing the rate of depreciation will change the slope of this curve, but not its basic shape.

Figure 6: State and Change in Health with Multiplicative Depreciation



Ultimately, whether $\dot{H}_H$ is positive or negative is an empirical question. The general specification for health transition allows for two hypotheses which would be consistent with a positive relationship between health and the change in health. First, an individual in a higher state of health may be better able to recover from an exogenous health shock without any medical care. In other words, a healthy body is better able to "heal thyself."[62] Second, the general specification allows health to be an element in the health production function. The implication is that the state of health is relevant to the productivity of medical care, or in medical terms, co-morbidities hurt the prognosis for recovery. For example, if an individual needs open heart surgery, she is more likely to recover fully and improve her state of health if she does not also suffer from diabetes,

[62] This is also consistent with the amount of depreciation modeled as a declining function of health as in Liljas (1998.) His result is consistent with mine in that a negative term associated with the marginal change in health associated with the state of health is subtracted from the user cost of health capital (see p. 158.) However, the rate of depreciation is still added to the cost of health capital in Liljas' model.

does not smoke and has a history of good diet and exercise. Mathematically this is reflected in the sign of the cross partial of health transition: $\alpha_{mH} \geq 0$.[63]

I assume weak diminishing returns to the change in health from both medical care and the state of health: $\alpha_m > 0, \alpha_{mm} \leq 0, \alpha_H \geq 0, \alpha_{HH} \leq 0$ and diminishing returns to scale for the joint production process.[64] This assumption will be critical to ensure a finite length of life, and will be discussed further in the context of the equilibrium demand for health.

Finally, I assume $\alpha_{mt} \geq 0$, reflecting the advancement of medical technology over time. For example, heart by-pass surgery is more effective in restoring health today than it was 20 years ago because of the smaller incisions necessary and advances in anesthesia, by-pass blood filtering and recovery procedures just to name a few improvements. Moreover, advances in medical technology have made more aggressive surgery viable for older patients.[65] However, in the future drug-resistant bacteria or other pandemics may call this assumption into question. In any case, the more general specification allows for empirical testing of this critical relationship of the benefits of medical care over time to medical care demand.

The proposed changes to the health transition function should be considered critically. The WTP literature critiques the human capital literature in part because of the sensitivity of the results to the specification of the health production function. Berger

---

[63] This assumption of a positive cross-partial does not further assume that this relationship is constant for all health states. At very low and very high levels of health the impact of medical care may be lower because there is little medical care can do to improve health.
[64] Grossman assumes diminishing returns to scale for medical care with time held constant, but constant returns to scale for the joint production process with medical care and time. Ehrlich & Chuma assuming a diminishing returns to scale for medical care both independently and jointly with time. See Ehrlich and Chuma's critique of Grossman's assumption in 1990, p. 768.
[65] McGinley (2006) p. R7.

(1987) argues that, "the fundamental problem with the health production function approach is that it is hard to identify and measure all of the inputs that affect health."[66] Harrington and Portney (1987), Atkinson and Crocker (1992), and Mullahy and Portney (1990) document problems with omitted variables, measurement error, and endogeneity bias respectively. The different implications I derive from the resulting equilibrium demand for health and time path of medical care demand support the WTP critique. The changes are made with the argument that they are more intuitively reasonable than the assumptions underlying the existing human capital models, but their validity must certainly be checked empirically with care taken to address econometric estimation issues.

3. Wealth Transition Function

The changes to the wealth transition function are minor generalizations consistent with extant human capital literature. The wealth transition function is a constant rate of interest, r, times accumulated wealth, R, plus income, w(H(t)), as a function of the state of health, minus the cost of medical care, $P(t)m(t)$, and the cost of consumption, Z, with price normalized to one:

$$\dot{R} = \frac{\partial R}{\partial t} = rR(t) + w(H(t)) - P(t)m(t) - Z(t) \tag{3}$$

The wage function is more generally a function of the state of health rather than healthy days and can include non-labor income from disability and/or social insurance payments. This assumption allows for the state of health to affect the wage rate as well as the amount of time available for work. It also reflects the modern blurring of boundaries

[66] Berger et. al., (1987) p. 33.

between work and non-work time: individuals often work when they are sick, and spend time procuring medical care when they are at work. I assume that $w_H \geq 0; w_{HH} \leq 0$ reflecting diminishing returns to health with respect to income.

Price of medical care is assumed to vary with time but not with the quantity of medical care. This assumption of diminishing returns to medical care is reflected in the health transition function, $\alpha_{mm} < 0$, following Grossman rather than in the cost function following E&C..

4. Endpoint conditions

The endpoint conditions are as follows:

$$\begin{aligned}
&H(0) = H_0 > H_{\min} \\
&H(T) = H_{\min} \\
&R(0) = R_0 \geq 0 \\
&R(T) \geq 0 \\
&T \leq T_{\max}
\end{aligned} \tag{4}$$

The model assumes a positive initial endowment of health and wealth. The terminal health state, H_{min}, is the minimum health necessary to sustain life which is exogenously fixed at the beginning of the planning period. In other words, the model does not allow medical technology to lower the level of H_{min} over the course of the planning horizon. Similarly, the maximum biological lifespan, T_{max}, is exogenously fixed at time 0. The actual terminal time, T, is determined endogenously by the transversality condition discussed in the next section. Finally, the terminal wealth R(T) must be nonnegative, eliminating the possibility of debt at the time of death. However, debt is not precluded for $t \in (t, T)$.

B. Optimization Results

1. Necessary Conditions

Based on the specification of the model in part A, the problem is to find the piecewise continuous control functions $Z(t)$ and $m(t)$, the terminal time *T*, and associated piecewise differentiable state functions $H(t)$ and $R(t)$, defined on $\left[0, T \le T^{\max}\right]$, to maximize (1) subject to (2), (3), and (4). The Hamiltonian function is defined by: [67]

$$V = e^{-rt}U\left[Z(t), H(t), \alpha(t)\right] + \lambda^{H}\alpha\left[t, m(t), H(t), \delta(t)\right] + \lambda^{R}\left[rR(t) + w(H(t)) - P(t)m(t) - Z(t)\right] \tag{5}$$

Let Z^{*} and m^{*} be the optimal controls defined on the interval $\left[0, T^{*}\right]$ that solve the problem. Then there exist continuous adjoinnt functions $\lambda^{H}(t)$ and $\lambda^{R}(t)$ such that for all $t \in \left[0, T^{*}\right]$, Z^{*} and m^{*} maximize (5). Necessary conditions include the following.

$$\dot{\lambda}^{H} = -\frac{\partial V}{\partial H} = -\left[e^{-rt}\left(U_{H} + U_{\alpha}\alpha_{H}\right) + \lambda^{H}\alpha_{H} + \lambda^{R}w_{H}\right] \tag{6}$$

$$\dot{\lambda}^{R} = -\frac{\partial V}{\partial R} = -\lambda^{R}r\,, \tag{7}$$

except at points of discontinuity of Z^{*} and m^{*}. Given $Z^{*} > 0$ and $m^{*} > 0$,

$$\frac{\partial V}{\partial m} = e^{-rt}U_{\alpha}\alpha_{m} + \lambda^{H}\alpha_{m} - \lambda^{R}P(t) = 0\,, \tag{8}$$

and

$$\frac{\partial V}{\partial Z} = e^{-rt}U_{Z} - \lambda^{R} = 0\,. \tag{9}$$

[67] The superscripts on λ denote the multiplier for health, H, and wealth, R, respectively. There is also a multiplier on the objective function, but it is equal to one for this problem. See Kamien and Schwartz (1991) Part II.

Given $T \le T^{\max}$,

$$V(T^*) \ge 0;\ T_{\max} - T^* \ge 0;\ \text{and } V(T^*)(T_{\max} - T^*) = 0. \tag{10}$$

The terminal condition $H(T) = H_{\min}$ implies no transversality condition on $\lambda^H(T^*)$, and the terminal condition $R(T) \ge 0$ implies the transversality condition $\lambda^R(T^*) \ge 0$ $(= 0 \text{ if } R(T^*) > 0)$. For notational ease, I drop the * notation below. All that follows is based on the simultaneous solution of these necessary conditions for the optimal values for the controls, states and terminal time as indicated above.

2. Equilibrium Demand for Health

The equilibrium demand for health offers a different explanation for why individuals may demand more medical care when health declines than that suggested by the equilibrium condition from the Grossman model. The equilibrium condition is:[68]

$$\frac{\left(U_H + U_\alpha \alpha_H\right)}{\lambda^R(0)} + w_H = g(t)\left[r - \alpha_H - \tilde{g}(t)\right] \tag{11}$$

Where g(t) is the ratio of the shadow prices of health and wealth $\frac{\lambda^H}{\lambda^R}$ following E&C. The corresponding continuous time equilibrium condition from Grossman is:[69]

$$G_i\left[\frac{U_{h_i}}{\lambda} e^{ri} + W_i\right] = \pi_i\left[r + \delta_i - \tilde{\pi}_i\right]$$

where π represents the marginal cost of health investment. The tilde over g(t) and π represents the percent change in these values. In both specifications the left-hand side represents the marginal benefits of health capital and the right-hand side the marginal

[68] Appendix B provides a detailed derivation.
[69] Grossman (1972) Appendix A equation (A13) p. 250.

costs. The model more generally reflects the marginal benefits from health and includes an additional term representing the marginal utility from the change in health. In Grossman's model the only marginal benefit from health is in terms of marginal healthy days, denoted $G = \frac{\partial h}{\partial H}$, the marginal product of health capital or the change in the number of healthy days associated with a change in health.

The critical difference between the two specifications appears in the marginal cost of health capital. In the Grossman model the rate of depreciation is added back to the user cost of capital consistent with traditional accounting. Thus, as the rate of depreciation increases, the user cost of health capital increases. To maintain equilibrium, individuals decrease their demand for health in order to increase the marginal benefits from health. This is the mathematical foundation for Grossman's conclusion that as health depreciates an individual continues to reduce his demand for health until "death is chosen."

In the Grossman model an individual will simultaneously demand less health and more medical care only if the *amount* of depreciation increases the marginal benefits from health more than the *rate* of depreciation increases the marginal costs. In this case, an individual would demand more medical care to increase health and bring down the marginal benefits to be in line with the marginal costs. However, the multiplicative form for health transition makes any significant increase in medical spending unlikely. If a high rate of depreciation occurs when health is already at a low level (as would occur at advanced ages) the resulting amount of health decline will be low, but the change in marginal benefits will still be high because the individual will already be on the steeper part of the concave benefit curve. If a high rate of depreciation strikes when health is still at a high level (at younger ages) then the amount of depreciation would be high and also push the individual to the steeper part of the benefit curve. Either way, in the Grossman model the change in marginal benefits is likely to be commensurate with the change in marginal costs requiring little investment in additional health to maintain equilibrium. Moreover, in the presence of a maximum biological lifespan the potential marginal benefit of additional years of longevity decreases as we age. In other words, as we approach the maximum longevity the upper bound on marginal benefits declines. Thus, including any longevity benefit, it is even less likely that we would see significant spending on medical care at advanced ages.

By contrast, the more general specification for health transition results in the marginal productivity of health, α_H, *decreasing* the marginal cost of health capital. As depreciation reduces the stock of health, the marginal productivity of health increases which decreases the user cost of health capital. At the same as time, the decline in health increases the marginal benefits from health. Disequilibrium is inevitable. The only way an individual can regain equilibrium is to invest in medical care. Investing in medical care does two things. First, investment increases the stock of health, which decreases the marginal benefits from health. Second, investing in medical care increases g(t), the marginal cost of health capital, due to the assumption of declining productivity of medical care.[70] Moreover, following the assumption $\alpha_{mH} > 0$, a decline in health makes medical care less efficient in producing more health, so even more spending on medical care is needed to drive down marginal benefits. Thus, as health declines an individual

[70] Grossman assumes a constant returns to scale health production function and investment at constant equilibrium proportions of medical care and time (see equation (11) p. 228) so additional investment in health does not change the marginal cost of gross investment.

may want a full recovery, but increasing marginal costs force him to settle for less health and ultimately a finite life. In this way, the model is consistent with economic theory, the observed increase in medical care spending in old and/or sick states and a finite life.

The impact of the assumption of diminishing returns to medical care is not obvious from inspecting the equilibrium condition because it is embedded in g(t). In E&C's model g(t) exactly equals the marginal cost of health investment:

$g(t) = \frac{\lambda^H(t)}{\lambda^R(t)} = \pi\alpha I(t)^\alpha$. Deriving g(t) from the necessary condition (8) yields:

$$g(t) = \frac{\lambda^H}{\lambda^R} = \frac{P(t)}{\alpha_m} - \frac{e^{-rt}U_\alpha}{\lambda^R} \tag{12}$$

The first term on the right-hand-side in (12) is consistent with E&C's assumption of increasing marginal cost of investment: as the quantity of medical care increases the marginal productivity of medical care decreases increasing the price term. The second term is the discounted dollar value of marginal utility with respect to health change. This value is subtracted from the marginal price decreasing the over-all marginal cost of health capital. The more negative the change in health, the higher the marginal utility and the lower the marginal cost of health capital. This suggests that individuals are forward looking in their valuation of health capital and implies that in equilibrium the demand for health is higher than in E&C.

Taking the derivative of g(t) with respect to medical care yields the following:

$$\frac{\partial g(t)}{\partial m} = -\frac{P(t)\alpha_{mm}}{\alpha_m^2} - \frac{e^{-rt}U_{\alpha\alpha}\alpha_m}{\lambda^R(t)} > 0 \tag{13}$$

The first term is positive because of the assumption of diminishing returns to medical care, $\alpha_{mm} < 0$, and the preceding negative sign. The second term is also positive due to the preceding negative sign, the concavity of the utility function in α, and necessary non-negativity of the wealth costate $\lambda^R \geq 0 \forall t$. Thus, an increase in medical care increases the marginal cost of health investment for two reasons: the marginal price of medical care increases, and the second-order impact on marginal utility decreases.

Finally, $\tilde{g}(t)$, the percent change in the marginal cost of health investment, can contribute either positively or negatively to the marginal cost of health capital. Decomposing $\tilde{g}(t)$:

$$\tilde{g}(t) = \frac{\dot{g}}{g(t)} = \left[\frac{\dot{\lambda}^H}{\lambda^R(t)} - \frac{\lambda^H(t)\dot{\lambda}^R}{\lambda^{R2}(t)} \right] \left(\frac{\lambda^R(t)}{\lambda^H(t)} \right) = \frac{\dot{\lambda}^H}{\lambda^H(t)} - \frac{\dot{\lambda}^R}{\lambda^R(t)} \tag{14}$$

The sign of $\tilde{g}(t)$ depends on the relative magnitudes of the co-states of health and wealth, which in turn depend on the endowed level of health and wealth and the relative rates of change. A higher level of health decreases the rate of change of the health co-state (see necessary condition (6)). Thus, increasing medical care to increase the state of health is more likely to result in a capital loss on health capital and a further increase to the user cost of health capital. This contributes to an individual ultimately settling for less health and a finite life.

3. The Value of Longevity and the Quality of Life

Contrasting the value of g(t) with that derived in E&C shows that an increasing value of longevity supports increasing medical care spending at the end of life only if we assume the specific functional form for health transition that adds back a rate of depreciation to the cost of health capital. The more general functional form suggests both value of longevity and the value of a healthy life fall at the end of life. The relative values of these two components of the value of health capital are indeterminate in both models; but the likelihood that quality of life is more valuable than quantity of life is greater in the more general specification because the quality of life value gets compounded rather than discounted as health declines.

The time path and general solution for g(t) can be derived from the equilibrium demand for health equation (15):

$$\begin{aligned}\dot{g}(t) &= g(t)\left(r-\alpha_H\right)-\left[\frac{U_H+U_\alpha\alpha_H}{\lambda^R(0)}+w_H\right] \\ g(t) &= g(T)e^{-\int_t^T (r-\alpha_H(u))du}+\int_t^T\left\{\left[\frac{U_H(u)+U_\alpha(u)\alpha_H(u)}{\lambda^R(0)}+w_H(u)\right]e^{-\int_t^u (r-\alpha_H(s))ds}\right\}du\end{aligned} \tag{16}$$

The corresponding equations in E&C are:[71]

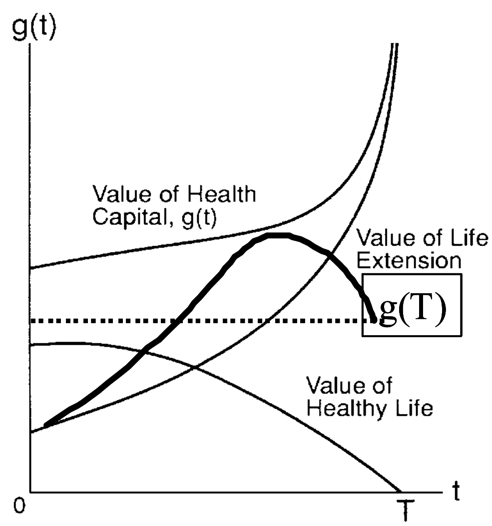
$$\begin{aligned}\dot{g}(t) &= g(t)\left(r+\delta(t)\right)-\left[\frac{U_h(u)}{\lambda^R(0)}e^{(r-\rho)t}+w\right]\phi'\left(H(t)\right) \\ g(t) &= g(T)e^{-\int_t^T [\delta(u)+r(u)]du}+\int_t^T\left\{\left[\frac{U_h(u)}{\lambda^R(0)}+w_h(u)\right]e^{-\int_t^u [\delta(s)+r]ds}\right\}du\end{aligned}$$

Recall E&C suggest that the first term of g(t) represents "the value of life extension" while the second is the "value of healthy life."[72] As with the equilibrium demand for health, the critical difference between the two specifications is that the more general form for health transition subtracts the marginal productivity of health instead of adds back the rate of depreciation. The key implications are that the value of longevity no longer increases unambiguously over the lifecycle and the value of a healthy life is compounded rather than discounted at advanced ages.

[71] These are equations (17) and (18) respectively in Ehrlich and Chuma (1990) p. 775.
[72] Ehrlich and Chuma (1990) p. 776.

Looking at the first term, when the rate of depreciation is replaced by $-\alpha_H$ from the more general specification the sign of the integral is not necessarily positive. At some point as t approaches T α_H is may equal r and exceed it for some time thereafter. If the integral in the exponent turns negative, the terminal value of health capital, g(T) will be compounded rather than discounted due to the preceding negative sign. This suggests that the value of life extension would be discounted at a decreasing rate up to the point where the integral would equal zero and then be *compounded* at an increasing, and then ultimately a decreasing rate as the terminal time approaches and g(t) must inevitably equal g(T). Intuitively, this picture suggests that when we are young we discount the value of extending our life. At some point (perhaps a mid-life crisis) we start to value life extension more; but ultimately we can only extend life so much so the value of life extension falls to just equal the value of health at the terminal time of death. Therefore, we cannot necessarily explain the increase in medical spending at the end of life with an increase in the value of longevity.

Figure 7: New Time Path of the Value of Life Extension



In both specifications the value of a healthy life declines over the lifecycle as $t \to T$ because the bracketed value of the second term is still positive for all $t \in [t, T]$. In the E&C model, the value of a healthy life gets discounted at an increasingly large rate as the second integral $-\int_t^u \left(r + \delta(s)\right)ds < 0$ gets unambiguously larger (in absolute value) over time. Thus, while the bracketed quality of life value gets larger, it is discounted by a larger amount so does not add increasingly to the total integrated value of a healthy life. By contrast, in the more general specification the instantaneous marginal value of a healthy life may be compounded rather than discounted if $-\int_t^u \left(r - \alpha_H(s)\right)ds > 0$. This is more likely at later ages for two reasons. First, health is lower so $\alpha_H > r$ is more plausible. Second, the integral from t to u is larger allowing the times when $\alpha_H > r$ to outweigh the earlier years when it is more likely that $\alpha_H < r$. Thus, the higher quality of life value experienced when old and sick contributes more to the integrated value of a healthy life. This potential for compounding makes the total value of a healthy life greater in the general specification than with the more specific functional form for health transition.

The relative values of longevity and quality of life, particularly at the end of life, are indeterminate in both specifications. The time path $\dot{g}(t) < 0$ declines at a faster rate in the general specification; but g(t) is likely to be higher because the quality of a healthy life is greater and the value of terminal health capital, g(T) reaches its peak prior to the end of life. Whether the value of longevity is greater than the value of a healthy life at any time, particularly towards the end of life, depends on the value of terminal health capital, g(T), and the rate of change in the integrated value of a healthy life, which in turn depend on all of the model parameters as the time paths of the co-states, states and controls are determined simultaneously with the terminal time, T. Nonetheless, if I consider just the single period right before death, (T-1 in discrete time) I can make some predictions about relative values of longevity and quality of life. In my specification with the quality of life on the left and value of terminal capital, or in E&C's terms the value of life extension, on the right:

$$\left[\frac{U_H(u)+U_\alpha(u)\alpha_H(u)}{\lambda^R(0)}+w_H(u)\right]\overset{?}{>}g(T)=\frac{\lambda^H(T)}{\lambda^R(T)}=\frac{P(T)}{\alpha_m(T)}-\frac{e^{-rT}U_\alpha(T)}{\lambda^R(T)} \tag{17}$$

And in E&C:

$$\left[\frac{U_h(u)}{\lambda^R(0)}+w_h(u)\right]\overset{?}{>}g(T)=\frac{\lambda^H(T)}{\lambda^R(T)}=\pi\alpha I(T)^\alpha$$

First, note that my specification requires the denominators $\alpha_m\big|_{m=0}\neq 0$ and $\lambda^R\neq 0$ which in turn requires R(T)=0 for g(T) to be defined. The E&C specification requires I(T)>0 for g(T) to be >0.[73] In both models the relative value of quality of life declines with more spending and increases with the shadow price of wealth.[74] The key observation is that adding the change of health to utility both increases marginal benefits and decreases marginal costs so that the greater the decline in health the greater the relative value of quality of life over longevity.

[73] Eisenring (1999) explores the trade-off between longevity and quality of life for a discrete-time pure investment specification of the Grossman model. To solve the model using dynamic programming he assumes zero investment at time T. This strategy would not be applicable to Ehrlich and Chuma's continuous-time specification.

[74] Since a bequest function would increase $\lambda^R(T)$ it would increase the relative importance of quality of life over longevity.

4. Time path of medical care demand

The time derivative of the necessary condition (8) solved for $\dot{m}$ yields the time path for medical care demand.[75] The signs above each term are either by assumption as discussed in part A, as required by the sufficiency conditions, or discussed below.

$$\dot{m} = \frac{1}{D}\left\{ \begin{array}{l} +e^{-rt}\overset{+}{\alpha}_m\left(\overset{?}{U}_{\alpha Z}\overset{?}{\dot{Z}}+\overset{-}{U}_{\alpha H}\dot{H}+\overset{-}{U}_{\alpha\alpha}\left(\overset{-}{\alpha}_t+\overset{+}{\alpha}_H\dot{H}+\overset{-}{\alpha}_{\delta}\overset{+}{\dot{\delta}}\right)\right) \\ -\overset{+}{\alpha}_m\left[e^{-rt}\left(r\overset{+}{U}_{\alpha}+U_H+U_{\alpha}\overset{+}{\alpha}_H\right)+\overset{+}{\lambda}^H\alpha_H+\overset{+}{\lambda}^R w_H\right] \\ +\left(e^{-rt}\overset{+}{U}_{\alpha}+\overset{+}{\lambda}^H\right)\left(\overset{+}{\alpha}_{mt}+\overset{+}{\alpha}_{mH}\dot{H}+\overset{-}{\alpha}_{m\delta}\overset{+}{\dot{\delta}}\right) \\ +\lambda^R\left(rP(\overset{-}{t})-\dot{P}\right) \end{array} \right\} \qquad (18)$$

The time-path of medical care demand cannot be unambiguously signed. Rather, it depends on the relative magnitudes of the different terms and, critically, the interaction between the time paths of medical care and consumption. The complexity of the time path reinforces Grossman's conclusion that the value of health and thereby the demand for medical care "depends on many other variables besides the price of medical care."[76]

The preceding scalar 1/D aggregates the second derivatives associated with medical care:

$$D = \left[-e^{-rt}\left(U_{\alpha\alpha}\alpha_m^2+U_{\alpha}\alpha_{mm}\right)-\lambda^H\alpha_{mm}\right]>0 \qquad (19)$$

The second derivatives are negative to satisfy the sufficiency conditions. The preceding negative terms $-e^{-rt}$ and $-\lambda^H$ make D unambiguously positive.

[75] Appendix B provides a detailed derivation.
[76] Grossman (1972) p. 223.

The co-state variables λ^H and λ^R are non-negative. The co-state with respect to health is positive from the sufficiency conditions[77] and the co-state with respect to wealth is nonnegative from the solution to the differential equation (7)[78] and the transversality condition associated with $R(T) \geq 0$.

I interpreted the first term of the time path of medical care demand as the "quality of life." Notably, this quality of life term is associated with the *change* in health, α, rather than the state of health. When health declines, $\dot{H} < 0$, the quality of life term will contribute positively to the time path for medical care demand depending on the sign and relative magnitude of the term associated with the time path of consumption, $\dot{Z}$. (This interaction will be discussed in detail in the next section.) The first line of the time path suggests forward looking consumers who demand medical care to forestall increasing declines in health rather than for any immediate utility value from the state of health.

The marginal benefits from the state of health are aggregated in the second line and contribute negatively to the time path of medical care demand. While perhaps counterintuitive,[79] this is a result of the Pontryagin maximum principal that requires $\dot{\lambda}^H = -\dfrac{\partial V}{\partial H}$: the shadow price of health depreciates at the rate of its marginal

---

[77] This is so because the health transition function, α, is concave in both the state and control variables and the optimization problem is for a maximum. See Caputo (2005), Corollary 3.1, p. 55.

[78] The explicit solution for $\lambda^R(t)$ from the necessary condition (7) is $\lambda^R(t) = e^{-rt}\lambda^R(0)$, which declines monotonically with t for any starting value $\lambda^R(0)$.

[79] Counterintuitive in that one might think the higher the marginal benefits from health the higher the demand for medical care. However, the time path for medical care demand looks at the change in demand for medical care over time rather than the level of medical care demand. Higher marginal benefits from health would be associated with higher medical care demand at all times, but not necessarily with an increasing demand over time.

contribution to the value function.[80] When a stock is physically used up in the creation of value (e.g. mining a stock of minerals in the ground and selling them) such depreciation is obvious. However, health is not necessarily "used" in the process of deriving value from it. In this case, health depreciation is independent of value creation. Nonetheless, the *value* of the stock of health still declines in proportion to the benefits derived. In other words, each unit of additional health created by medical care is worth less by the sum of the marginal benefits including both the marginal utility from the state of health, U_H, and the change in health, U_α and the marginal changes in health and wealth with respect to health, α_H and w_H. Thus, the optimal choice of medical care at any point in time needs to account for the change in the value caused by the change in the stock of health associated with the choice. In other words, investing in health decreases the shadow price of the next incremental investment which reasonably is associated with a decline in the *time path* of demand.

The other positive contribution to the time path for medical care demand is the α_{mt} term in the third line. This cross-partial derivative of health production with respect to time and medical care can represent what Becker et. al (2007) refer to as "hope." In their words, it is "the current consumption of future survival" associated with the marginal additional health produced by medical care, as well as the "option value of seeing a new treatment being discovered before one's death."[81] This value is multiplied by the shadow value of health *plus* the discounted marginal utility associated with the

[80] See Dorfman (1969) for this intuitive interpretation. The second line of the time path is from substituting the time path of the health co-state plus an additional term, rU_α, which is associated with the time derivative of the necessary condition. See Appendix B.
[81] Becker et. al. (2007) p. 3.

change in health: $\left(e^{-rt}U_{\alpha}+\lambda^{H}\right)$. While both terms will decrease with time (due to discounting and the negative time path for the co-state) the marginal utility of health change will increase when health declines sharply. If sharp declines occur at younger ages this multiplier will be higher, increasing the positive impact of this "hope" term on the time path of medical care demand.

The remaining terms on the third line reflect the inevitability of death. As depreciation increases over time causing health to decline, the effectiveness of medical care declines which is associated with a negative time path of medical care demand. This is consistent with the demand for a finite life suggested by the equilibrium demand for health. An individual will only demand more medical care to the point where the marginal cost equals the marginal benefit, and as health declines an individual will not spend ever increasing amounts on medical care to recoup all lost health.

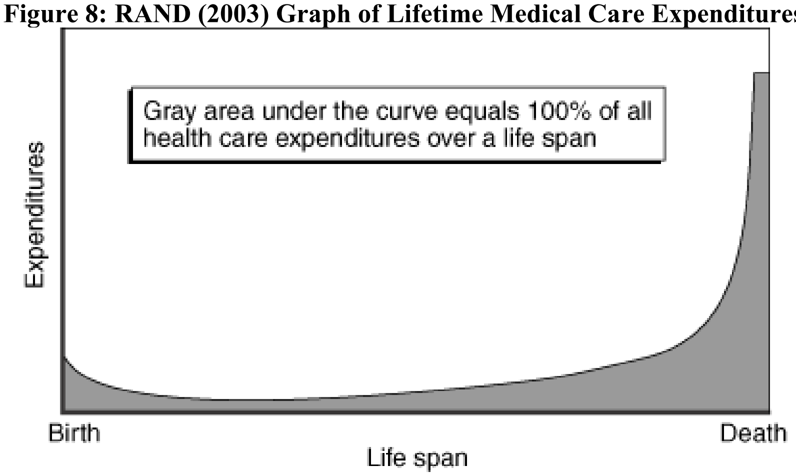
Finally, the last line represents the traditional downward sloping demand curve of medical care relative to price. If the change in price of medical care over time, $\dot{P}$, exceeds the real rate of interest then the demand for medical care will decline. However, the price term is multiplied by the shadow value of wealth which declines over time. This suggests the testable hypothesis that the impact of a change in price on the demand for medical care decreases over an individual's life. In other words, the frequently cited price elasticity of medical care demand that was empirically estimated in the RAND Health Insurance Experiment from observing individuals under 63 years old may not be the same as that for older individuals.[82] Since individuals over 65 are disproportionately represented in the top 5% of spenders, their price elasticity is particularly important to

[82] The price elasticity estimated from the RAND study was approximately -.2. See Manning et. al (1987). The RHIE will be reviewed in the next section.

evaluate policy aimed to reduce total medical care expenditures. Moreover, as the change in health becomes more negative, the positive elements of the time path are more likely to outweigh the negative. Thus, the impact of price on the demand for medical care is likely to be less for those with larger declines in their health at any age.

Taken in sum, the time path of medical care demand has more factors associated with declining than increasing demand over time. However, a declining time path for medical care demand would be inconsistent with the observed increase in medical care spending over an individual's lifecycle. This dramatic increase is illustrated heuristically in the RAND graph below which reflects lifetime medical care expenditures from all sources – private, government and insurance:

Figure 8: RAND (2003) Graph of Lifetime Medical Care Expenditures



The relative magnitude of the positive factors in the time path in (18) may still result in increasing demand; but without including the change of health in the utility function, the only remaining positive impact would be the "hope" factor. This suggests that if the change in health is not relevant for medical care demand, then the hope factor dominates all others at the end of life. The implications of this time path are also consistent with the demand for medical care being a function not of age directly, but of some other factor that is confounded with age. Several empirical studies have found that the effect of age on health care expenditures is reduced and even insignificant when controlled for the time to death.[83] If the change in health is relevant, and it is more negative the shorter the time to death, then this may account for these empirical results.

Nonetheless, the time path for medical care demand is consistent with the ambiguous time path of health investment described by Grossman:

> …even though health capital falls over the life cycle, gross investment might increase, remain constant or decrease. This follows because a rise in the rate of depreciation not only reduces the amount of health capital demanded by consumers but also reduces the amount of capital supplied to them by a given amount of gross investment.[84]

I can demonstrate this by applying Grossman's assumptions to the time path of medical care demand in equation (20). First, Grossman does not include the change in health in utility which eliminates the entire first line as well as the U_α elements of the second and third lines. Second, Grossman's health investment function is independent of the state of health and time, which eliminates the cross-partial terms α_{mH} and α_{mt} from the third line. Finally, Grossman's specification of depreciation as proportional to the state of health changes the remaining α_H in the second line to $-\delta$.[85] What remains is:

$$\dot{m} = \frac{\alpha_m \left(e^{-rt} U_H - \lambda^H \delta + \lambda^R w_H \right)}{\lambda^H \alpha_{mm}}$$

Assuming declining returns to scale for medical care[86] makes the time path negative *except* if the depreciation term outweighs the the marginal benefits to utility and income.

---

[83] See among others Zweifel et. al (1999) and Sterns and Norton (2004.)
[84] Grossman (1972) p. 238.
[85] For simplicity Grossman also assumes that the price for medical care changes at the rate of interest which eliminates the price term in line four.
[86] In the Grossman model this would be declining returns to medical care holding the time spent on health constant as the total production function is homogenous of degree one.

Only in this case would both numerator and denominator be negative resulting in an increasing demand for medical care.

5. The Time Path of Consumption

Following the same process as with the time path of medical care, the time derivative of the necessary condition (9) solved for $\dot{Z}$ yields the following:[87]

$$\dot{Z} = \frac{\dot{H}\left(U_{ZH} + U_{Z\alpha}\alpha_H\right) + U_{Z\alpha}\left(\alpha_t + \alpha_m \dot{m} + \alpha_\delta \dot{\delta}\right)}{-U_{ZZ}} \tag{21}$$

Just as the time path for consumption, $\dot{Z}$, is an element in the time path for medical care, the time path for medical care, $\dot{m}$, is an element in the time path for consumption. Thus, any empirical estimate of the demand for medical care that is not estimated simultaneously with the demand for other consumption risks omitted variable bias. Moreover, the model does not assume a sign for the cross-partial between consumption and health or the change in health (though the two must necessarily be of the same sign.) By contrast, M&T assume health and consumption are compliments in utility. Similarly, E&C assume a positive cross-partial in their comparative dynamics. I consider the sign of the cross partial an empirical question.

The sign of the cross partial has implications for the relationship between health and consumption. Consider three possible scenarios for the relationship between utility from consumption and health. First, suppose that health shifts the utility of consumption curve by a constant amount at all levels of consumption, including Z=0. This would occur if the utility function were additive in health, for example: $U = Z + H + \alpha$. In this

[87] Appendix B contains a detailed derivation.

case, the slope of the utility of consumption curve stays the same no matter the level or change in health. The cross partials are zero, $U_{ZH} = U_{Z\alpha} = 0$. If the cross-partial derivatives are zero, the time paths drop out of each equation leaving them independent of each other. However, this assumes that health provides some utility in the absence of consumption: $U(0, H, \alpha) > 0$ which would be contrary to any pure investment model for health demand.

Second, assume, consistent with M&T that Health contributes to utility proportionally to the utility of consumption. In this case, health provides no utility in the absence of consumption, $U(0, H, \alpha) = 0$, consistent with pure investment, but not with pure consumption models. The slope of the utility of consumption curve would increase (decrease) with an increase (decline) in health: $U_{ZH}, U_{Z\alpha} > 0$.[88] Looking at the first term of the numerator in $\dot{Z}$, this would mean that a decline in health would be associated with a decline in consumption. Moreover, with $U_{Z\alpha} > 0$, an increase in $\dot{m}$ would be associated with an increase in $\dot{Z}$. This is inconsistent with equilibrium given a binding budget constraint: consumers cannot simultaneously spend more on both medical care and other consumption. However, it is consistent with the demand for a wealth transfer in sick states suggested by Nyman (2003). Nyman argues that "consumers demand health insurance because they desire an income transfer from those who remain healthy

[88] Both cross-partials will have the same sign because the change in marginal utility of consumption will shift in the same direction for a change in health, ∂H, and a change in the change in health, $\partial \alpha$. For example, a change in health that is marginally less negative (a positive change) nonetheless reduces health, but results in a marginal utility of consumption that is marginally higher (a positive change) than the marginal utility of consumption associated with a marginally more negative change. It is not possible for health to provide no utility in the absence of consumption and for the cross-partials to be negative rather than positive because this would violate the FOC that $U_H > 0$. One way to show this is if health contributes to utlity inversely, for example U = Z/H then $U_H = -Z/H^2 < 0$ which is not admissible.

in the event that they become sick," not because they want to avoid uncertainty as under conventional expected utility theory.[89]

The cross partial can be positive in a pure consumption model if the marginal utility from consumption increases at an increasing rate at higher levels of health. In other words, the slopes of the utility curves with respect to consumption would be steeper at higher levels of health. While consistent with pure consumption models, the positive cross partial is still inconsistent with a binding budget constraint. However, there is no way to empirically distinguish the pure consumption from the pure investment models if the cross-partial of the derivatives is positive.

However, a negative cross-partial is inconsistent with a pure investment model. If health provides utility in the absence of consumption, but increases the utility of consumption at a declining rate as consumption increases, then the utility of consumption curve becomes less (more) concave as health increases (declines) and $U_{ZH}, U_{Z\alpha} < 0$.[90] The intuition for this case is that at higher levels of consumption consumers are materially better able to buffer the effects of declining health. For example, they can purchase comforts and other material goods to replace some of the utility lost from declining health. Viewed the other way, the higher the level of health, the flatter the utility of consumption curve or the more indifferent individuals are between different levels of consumption. Looking at the time path of consumption, this case implies that a decline in health would be associated with an increase in $\dot{Z}$; however this increase would be off-set at least in part by an increase in $\dot{m}$ which would be associated with a decrease

[89] Nyman (2003) p. 30.

[90] As with the prior scenario, the opposite sign is not plausible. If health provides utility in the absence of consumption then higher levels of health must be associated with higher levels of utility at the y-axis otherwise this would violate the first-order condition that $U_H > 0$.

in $\dot{Z}$. The latter effect would need to outweigh the former to be consistent with a binding budget constraint.

C. Summary of the Model's Implications

1. The Value of the Change in Health

The effect of the change in health appears in the equilibrium demand for health and the time paths for both medical care and consumption. The marginal utility from the change in health increases the marginal benefits from health when health declines, and the increase is greater the greater the decline in health. This increases the disequilibrium in the demand for health as the cost of health capital declines. It is this disequilibrium that can occur at any age or state of health that underlies the model's explanation for high medical care spending. In addition, the cross-partial $U_{\alpha Z}$ links the time paths of for medical care and consumption which suggests that the demands for medical care and consumption should be estimated jointly. Finally, the positive impact of the quality of life component of the time path of medical care demand would not exist but for the value of the change in health in utility. Therefore, the theoretical model strongly suggests that the change in health is relevant to the demand for medical care demand and may help to explain the diversity of demand among individuals with the same state of health.

2. Quality vs. Quantity of Life

Ehrlich and Chuma (1990) suggest that an increasing value of longevity drives an increase in medical care spending over the lifecycle. However, this conclusion is entirely dependent on a multiplicative functional form for health transition. Generalizing the functional form for health transition and including the change of health in the utility function suggests that both the value of longevity and the value of a health life fall towards the end of life. However, the model suggests that quality of life has a higher value than longevity the greater the change in health.

3. Health and Wealth

The time path for consumption is an element in the time path for medical care, and the time path for medical care is an element of the time path for consumption. If consumption and the change in health are not independent, then the sign of this relationship suggests further hypotheses about the relationship between health and wealth over time. If health and consumption are complements, then according to the time path of medical care demand an increase in consumption is associated with an increase in medical care demand, which would violate a binding budget constraint in the absence of savings. However, according to the time path of consumption, an increase in medical care is associated with a decrease in consumption consistent with a binding budget constraint. Whether or not both medical care and consumption increase depends on the magnitude of the change in health which affects the magnitude of the cross partial and the magnitude of the first term in the time path for consumption (see equation (21)). This suggests a testable hypothesis for Nyman's (2003) theory that the demand for insurance is a demand for a wealth transfer rather than a demand for certainty. Specifically, the model suggests that individuals demand both more medical care and more consumption the greater the change in health and the lower the accumulated savings. Furthermore, if health and wealth are not independent, then pure consumption and pure investment models of the demand for medical care would suffer from omitting the relevant relationship between health and consumption.

4. The Advance of Medical Technology Over Time

The time path for medical care demand specifically includes the element α_{mt} : the cross partial of the change in health with respect to medical care and time. Notably, this is the only element associated with an increasing time path of medical care demand other than those associated with the change in health. If the change in health is irrelevant to utility, then "hope" must dominate all other elements in order to explain the observed significant increase in the time path of medical care demand over the lifecycle.

5. Decline in Price Sensitivity Over Time

The time path for medical care demand suggests that the change in demand becomes less sensitive to price over time. This is because the magnitude of the shadow price of wealth decreases over time (see the necessary condition(7)). This decline in price sensitivity holds for any health state and change in health. This implies that the positive contribution to medical care demand from the change in health may dominate the negative price effect more easily for large negative changes and at advanced ages even for smaller negative changes. Thus, this may help to explain the increase in demand for medical care with increased age without reference to declining states of health. It also has significant implications for policy proposals that aim to reduce the demand for medical care by increasing the price. The impact among the older high spenders will be less than expected as will the impact on younger high spenders who experience significant declines in their health even if their state of health is relatively high.

6. The Relationship Between the Health and the Change in Health

The model generalizes the specification for health transition to be consistent with current accounting for intangible assets and reflect the medical observation that those in better health have a better prognosis for recovery. Mathematically, this is reflected by a positive marginal change in health with respect to health: $\alpha_H > 0$. This is a critical assumption of the model that along with the utility value of the change in health induces disequilibrium in the demand for health as health declines. This assumption and the assumption of non-separability between health and consumption will be tested along with the central hypothesis that the change in health matters in Part II of the dissertation.

Part II: Empirical Tests of the Dynamic Demand for Medical Care

The dynamic demand for medical care specified in Part I suggests several testable hypotheses summarized in the prior section. I focus on the central hypothesis that the change in health is a significant factor in the demand for medical care. To test the hypothesis I specify estimating equations that are consistent with the theory. A key assumption is that the utility of health and consumption are not separable, and thereby the demands for medical care and consumption are not independent (see the time paths of medical care and consumption demand equations (18) and (21)). In addition to testing the theoretical hypothesis I test for the significance and estimate the sign of the relationship between health and wealth and the demands for medical care and consumption. I leave testing of the other hypotheses to future research which I describe briefly in the conclusion of the dissertation.

I. Empirical Literature Review

The literature on the demand for medical care is vast. This review will focus on empirical tests of the Grossman model, health-state dependent structural estimates of utility of consumption, and literature associated with the econometric issues of skewed discrete demands, particularly in the presence of unobservable heterogeneity. This review concludes with an extensive discussion on the methods used to proxy for unobservable health and price of medical care and my justification for using a multiple correspondence analysis in the present study. The theoretical model abstracts from the insurance issue which makes modeling this joint demand is outside the scope of the present work. Therefore, I do not review the extensive literature modeling this joint demand and estimating the moral-hazard effect of insurance on demand.[91] I do use insurance as an instrument for price and address the literature associated with insurance briefly in this context.

A. Estimates of the Grossman Model[92]

Recall the central contribution of the Grossman model that the demand for medical care is a derived demand from the demand for good health. The Grossman model assumes instant adjustment such that the desired health is equal to the observed health. Thus, the coefficient on health in the demand for medical care equation is expected to be positive: the higher the state of health the more medical care demanded. Unfortunately, much of the literature finds a negative relationship: the higher the state of health the less medical care demanded. Muurinen (1982), Wagstaff (1986, 1993) and Erbsland et. al. (1995) all estimate variants of the Grossman model and find a negative rather than a positive relationship. Based largely on these studies, Zweifel and Breyer (1997) conclude that:

> …the notion that expenditure on medical care constitutes a demand derived from an underlying demand for health cannot be upheld because health status and demand for medical care are negatively rather than positively related.[93]

Grossman (1999) argues that such a dismissal of the theory is premature given the likely biases in the estimates due to various aspects of model misspecification. Specifically, Grossman argues that Wagstaff's estimates are biased because he considers health status

---

[91] See the classic papers by Akerlof (1970) and Pauly (1968), a critique of the standard theoretical argument by Nyman (2003), and empirical studies by Cameron et. al. (1988) and Zweifel and Manning (2000), Koc (2005) and references therein.)

[92] There is a large literature testing Grossman's hypotheses regarding the effect of education and the wage rate on health production and the demands for health and medical care. Since the role of education is not a focus of the dissertation and my theoretical model abstracts the time element in the Grossman model I will not discuss this literature.

[93] Sweifel and Breyer (1997) p. 62.

as exogenous, which would induce bias if depreciation (which is unobserved and so subsumed in the error in both Grossman's and Wagstaff's models) is correlated with health, which it clearly is in Grossman's theoretical model.[94]

Part of the problem may be that the reverse causality may be difficult to tease out in a single period model. According to Grossman, the desired end-of-period health induces the demand for medical care over the period. However, it is clearly the case that the beginning of period health influences the end of period demand for health. In other words, the end of period demand for health is relative to prior health, not absolute. Van Doorslayer (1987) and Wagstaff (1993) estimate dynamic models with longitudinal data to incorporate this lagged relationship. Both also model a cost of adjustment to desired health. Grossman (1999) argues that these adjustment models are "ad hoc" and are biased because the cost-of-adjustment cannot be captured with only lagged health, but also needs the next period health (H_t requires both H_{t-1} and H_{t+1}.) However, this would not be the case for a feedback or closed loop form of an optimal control solution.[95] None of the extant tests of the Grossman model are of this form or use more than two periods of observations on health.

The dynamic demand model offers an explanation for the negative relationship between current health and the demand for medical care that is consistent with the derived demand hypothesis. First, the demand for medical care is a function of the change in health, $m = f(H_t - H_{t-1})$, explicitly incorporating reference to the prior-period's level of health. If the demand for medical care is greater the greater the decline in health, then the coefficient on the change in health would be negative. Distributing

[94] Grossman (1999) p. 50.
[95] See Caputo (2005) p. 511.

this negative coefficient to the two terms of the change in health suggests a negative, rather than a positive coefficient on contemporaneous health. Second, following from the more general specification for health transition, the higher the state of health the more positive the change in health, $\alpha_H > 0$ and the more productive any given input of medical care $\alpha_{mH} > 0$. Thus, the higher the state of health, the less medical care is needed to achieve the demanded state of health. However, the state of health in this case is lagged health, which is not included in single-period models. As a result, inferences on the derived demand hypothesis from single period models may be confounded by the several different health effects: demand for health, the utility effect of the change in health, and the productivity effect of lagged health.

B. Estimates of the Relationship Between Health and Consumption

A critical empirical question is: what is the sign of the cross-partial of utility with respect to health and consumption? The sign can go either way: a decline in health can increase the marginal utility of consumption if consumption can substitute for health as a source of utility. On the other hand, a decline in health can reduce the marginal utility from consumption of health and consumption are compliments in utility. Or, as is assumed by the vast majority of demand estimates, the states of health and wealth may be independent.[96] There is no consensus in the literature. Viscusi and Evans (1990), Gilleskie (1998), Sloan et. al. (1998) and Finkelstein et. al. (2008) all find a positive relationship: better (worse) health is associated with higher (lower) utility from consumption. However, Rust and Phelen (1997) and Lillard and Weiss (1997) find the

[96] See a recent working paper by Finklestein et. al (2008) and references therein for a brief list of literature that assumes state independence and other estimates of the effect of health on the utility from consumption.

opposite: a decline in health is associated with higher utility from consumption. Evans and Viscusi (1991) find no state dependence, and Blau and Gilleskie (2006) who extend Rust and Phelen (1997), find mixed signs: sometimes good health is associated with higher utility from consumption, but other times poor health is associated with higher utility from consumption and.[97]

All of this literature estimates the utility function directly. This approach raises numerous econometric issues from the appropriate dependent variable for utility to reasonable measures of risk aversion and modeling of bequest functions (see Finklestein et. al. (2008) for a brief discussion of these issues.) I will take a different approach. Rather than estimating a utility function, I will estimate a system of Marshallian demand functions and impose symmetry restrictions on the cross-price parameters such that the inferences can be interpreted as consistent with an admissible yet unspecified utility function. The specification of a system of demands relaxes the assumption of independence between the demands for medical care and consumption, and including the state of health in the demand for consumption allows for state-dependent demand. This avoids the difficulty of specifying the utility function and estimating risk aversion parameters, but is still subject to other econometric issues further discussed in the next section.

C. Empirical Issues

1. Discrete, non-negative skewed counts of medical care demands

[97] Blau and Gilleskie (2006) estimate numerous combinations of health status and retirement status of married couples, and the coefficient associated with the utility of consumption varies among different scenarios.

Empirical work that uses counts of medical care services as a dependent variable faces the problem of a high percentage of zero counts: many people do not report using any medical services in a given study period. In addition, many of the observed count distributions have a long-right tail: a small number of observations have vary high medical care usage. Many studies facing this issue use some variant of a negative binomial distribution to model such skewed count data. Negative binomial models have been extended to hurdle models consistent with the two-part decision models underlying the RHIE (Pohlmeier and Ulrich (1995)). These two-part models estimate parameters conditional on any use of medical care services. Deb and Trivedi (1997) extend the finite mixture model of Heckman and Singer (1984) to the demand for medical care. They argue that a finite mixture negative binomial model offers more flexibility than the hurdle models and results in a better fit to the data.

Beyond the problem of zero counts, models that estimate the usage of medical services face the problem of capturing the substitutability among services as well as the omission of significant classes of services from most datasets.[98] Studies that estimate counts tend to estimate the demands for different types of services independently (e.g. they estimate different models for hospital use, general practitioner visits, etc.) which ignores the interrelationship among medical services. Wagstaff (1986) is a notable exception. He estimates a recursive system of structural equations for GP Visits, Hospital Stay and Medicine built with the assumption that GP visits effect the other two utilizations. However, his specification requires the additional assumption of the general practitioner as the gatekeeper which is otherwise not modeled in the theory. I have

[98] Few datasets have detailed consumption data on prescription pharmaceuticals let along over-the-counter drugs and alternative therapies such as acupuncture or chiropractic.

found no analysis as to whether such additional assumptions produce more consistent and/or efficient parameter estimates or any literature that explicitly tests the principal-agent assumption associated with the gatekeeper model.

I take the approach of modeling the demands for different types of medical care jointly in a seemingly unrelated system of demands. This allows for the likely correlation in the error terms as well as imposition of the economic restrictions on the cross-partials price parameters. However, it makes modeling the overdispersion of the data more difficult. All of the literature regarding the estimation of medical care counts cited above uses single-equation models. While the theory of SUR models does not preclude each equation having a different specification,[99] I have found no literature that specifies a SUR using a negative binomial with different dispersion parameters for each equation. Similarly, I have found no literature that extends either a two-step latent variable or hurdle model to a system of equations in the cross section or over a panel. Beyond the implementation challenges, the properties of the resulting variance-covariance matrix for a two-step estimator of a SUR are unclear and undeveloped in the literature.

2. Discrete demands and unobservable heterogeneity: the incidental parameters problem

The problem of accounting for unobservable heterogeneity in models of count data lies at the frontier of econometric methodology.[100] Known estimators all entail uncomfortable trade-offs. One is between a fixed or random effect.[101] The random

[99] See Greene (2003) p. 370 equation 14-41.
[100] Many models of count data take the form of discrete choice models, and I found Train (2003) to be an exceptional treatment and review of the recent literature that uses new simulation methods to implement these models.
[101] There are many well-regarded textbook treatments of panel data models. The two I used most extensively were Greene (2003) and Cameron and Trivedi (2005).

effects model requires the often untenable assumption that the unobservable heterogeneity is independent of the regressors. This assumption is particularly difficult to maintain in demand for medical care models with the health state as a regressor because of its likely correlation with unobservable depreciation, the presence of insurance, and errors associated with the price of medical care.[102] Fixed effects models relax this independence assumption, but raise two other problems. The first is that fixed effects models cannot identify parameters associated with time invariant regressors. Perhaps the more critical barrier is the incidental parameters problem.

The incidental parameters problem was first identified by Neyman and Scott (1948), popularized in the econometrics literature by Lancaster (2000) and perhaps most clearly stated by Arellano and Hahn (2005):

> Because only a finite number T of observations are available to estimate each (fixed effect) α_i, the estimation error of $\hat{\alpha}_i(\theta)$ does not vanish as the sample size n grows, and this error contaminates the estimates of parameters of interest.[103]

The bias is trivially corrected for normally distributed linear models with a degree-of-freedom correction. However, there are only a few known non-linear models with a sufficient statistic upon which to condition the fixed effect, specifically the logit and the poisson. While Hausman Hall and Griliches (1984) specify a popular fixed effects negative binomial model, Greene (2007) illustrates that this model is not immune to bias.[104] Greene (2004) illustrates that "the finite sample behavior of the fixed effects estimator is much more varied than the received literature would suggest" for both

[102] I am referring specifically to the parameters that effect the non-linear price function in the presence of insurance deductibles examined by Ellis (1986).
[103] Arellano and Hahn (2005) p. 3.
[104] Greene (2008) demonstrates that the bias associated with the HHG fixed effects model comes from an "omitted variable bias" rather than an incidental parameters problem because the HHG model builds the fixed effect into the variance rather than the mean. Green specifies a "true" NB FE model but concedes that "the specification may also suffer from the incidental parameters problem" p. 38.

discrete and truncated continuous distributions such as the tobit.[105] Greene has done much recent work in this area, but has not yet extended his analysis to multiple equation models.

3. Unobservable Health

Modeling the unobservable health state has received much attention; but there is no universally accepted measure of health in the literature. There are four commonly used approaches: a single or several observed measures of health as independent variables, multiple-indicators-multiple-causes (MIMIC), latent variable models, and principal components analysis (PCA).[106] I will focus on extant empirical models of medical care demand and discuss the pros and cons of each method in terms of the present analysis.

A prevalent and straightforward approach is to use a single measure of health or several observed health indicators all as independent variables (Grossman (1972), Manning (1982), Cameron et. al. (1988), Wagstaff (1993), Hunt-McCool et. al. (1994), Pohlmeier and Ulrich (1995), Gould and Jones (1996), Deb and Trivedi (1997), Gilleskie (1999) among many others[107]). The most common single indicator of health is self-assessed health (SAH). In the BHPS this is the answer to the question:

> Please think back over the last 12 months about how your health has been. Compared to people of your own age, would you say that your health has on the whole been… with catgories from 1 = excellent to 5 = very poor.[108]

---

[105] Greene (2004) p. 98.

[106] There is another method using anthropometric measures such as height, weight, blood pressure, etc., (see Steckel (2008) for a recent review of literature associated with height and facial skeletal measures). These are often used to measure health in developing countries and/or child health. Since such measures are not available in the BHPS, I do not review this literature.

[107] While this review is focused on models of medical care demand, Contoyannis, Jones and Rice (2004) is notable for their use self-assessed health in their analysis of the dynamics of health in the BHPS.

[108] http://www.iser.essex.ac.uk/survey/bhps/documentation/volb/wave1/aindresp3.html#AHLSTAT

There is an extensive literature on the potential measurement errors including framing issues associated with this question (see Contoyannis et. al. (2004) and references therein.) Nonetheless, SAH is frequently used because it is available in most applicable surveys and has been shown to correlate with mortality but not with socioeconomic status more generally (van Doorslaer and Gerdtham (2003)). However, correlation with mortality is not the same as correlation with morbidity which is more relevant to the demand for medical care since you do not demand care when you are dead. While a substantial proportion of lifetime medical care expenses occur in the months leading up to death (see Sterns and Norton (2004) and references therein) a measure of mortality cannot reflect the demand for medical care at earlier points of the lifecycle. Ider and Kasl (1995) show that SAH predicts functional limitations, but like mortality, functional limitations is at best a noisy indicator of the need for medical care. Contoyannis et. al (2004) cite van Doorslaer et. al (2000 and 2002) to suggest that "categorical measures of SAH have been shown to be good predictors of subsequent use of medical care."[109] However, these papers indicate that SAH increases the value of horizontal inequity measured with aggregate medical care expenditures, not of individual demand for medical care.

Another health indicator is limitations in daily activities such as dressing, bathing and shopping (see in particular Cohen et. al (1995) and Gould and Jones (1996)). On the one hand, questions about limits to functioning appear more objective than categorical questions about health. However the lists of activities are often inconsistent across studies, and what constitutes a "limitation" is not universal among respondents. In

---

[109] Contoyannis et. al. (2004) p. 476.

addition, the issues of self-reporting relative to peers and the degree of adaptation are still present.

There are also problems with physician-reported health indicators such as ICD-9 indication codes or "objective" measures of health such as Body-Mass Index (BMI) and lung function. Blaxter (1985) found only 80% correspondence between self-reports of chronic illnesses and GP reports; others have found the errors to be systematically related to socioeconomic status and other demographic variables (O'Donnell and Propper (1991), Senior (1998)). The assumption in these papers is that the individual commits the error rather than the physician. I find no justification for this assumption.

Using multiple health indicators risks collinearity that can impair inferences. However, studies that use multiple health indicators typically use them for controls rather than as primary objects of inference.[110] This is not the case in the present study where the primary hypothesis is the effect the change in the state of health on the demand for medical care.

Bound (1991) proposed a latent variable method to create a single health measure. His primary purpose was to explain labor force participation decisions, and the concern with SAH was a likely endogeneity between the bias in the self-reported measure and the object of interest: the decision to leave the labor force.[111] Specifically, ill health is perceived as a legitimate reason for not working, so those not working make be more

[110] An exception is Koc (2005) who constructs an interesting indicator of health based on "paths" from good to ill health associated with indicators of illness, disability and disease. However, while Koc talks about "paths" in a seemingly dynamic sense his estimates are based on a single observation of health. His interpretation of the moral hazard effects of the different paths are difficult to interpret in the cross-section. Moreover, his inferences based on these health measures are curious: he finds that poor health decreases the probability of insurance (perhaps confounded by the probability of employment). I have not found Koc's methodology to identify the state of health replicated in any other studies.

[111] See Bound (1991) for a detailed discussion of the issues associated with using SAH in retirement decision models. Bound et. al. (1999) p. 180 notes that the measurement errors associated with SAH are exacerbated in models that use lagged as well as contemporaneous health as is the case in the present study.

likely to report ill health to "rationalize their behavior." The latent variable model uses the demographic and health indicators in the data to instrument the endogenous and error-ridden self-reported health variable.

This latent variable method is not appropriate for the present study for two reasons. First, the latent variable method creates a single health measure along a single dimension of health. This dimension is identified by the choice of the dependent variable in the health equation. Bound et. al (1999) use self-reported health as the dependent variable. The resulting index linearly combines the coefficients of the other indicators that correlate with the chosen dimension. Those aspects of functional limitations and other indicators that do not correlate with self-reported health are stripped out of the resulting index. Therefore, while the latent variable model may produce an index with less bias and measurement error, it does not produce a multi-dimensional index despite using multiple indicators of health. In the present study it is these uncorrelated dimensions of health beyond self-reported health that may have significant explanatory power for the demand for medical care. Second, unlike the retirement decision context, in demand for medical care models the latent variable strategy does not improve the consistency of parameter estimates because the instruments (ADL limitations, disability) will still be correlated with the behavior being modeled -- the demand for medical care. Moreover, to the extent there is a correlation between the health indicators and the demand for medical care (e.g. hypochondria) it should not be "stripped away" since it may be associated with (potentially excessive) demand for medical care.

Multiple Indicators Multiple Causes (MIMIC) models (van de Ven and van der Gaag (1982), van der Gaag and Wolfe (1982), Wagstaff (1986), Ersblad et. al (1995)) explicitly model the health state in a system of equations. While addressing the likely endogeneity between health and the demand for medical care, specifying a separate equation for the health state raises other difficult modeling issues. Extant MIMIC specifications are notably unclear as to whether the health equation represents a demand or a production function.[112] Specification of a health demand function requires difficult assumptions about unobservable depreciation (Grossman (1972, 1999)), Wagstaff (1986, 1993), Muurinen (1982)). Grossman assumes that the net rate of disinvestment in health is small relative to the rate of depreciation. However, this assumption is unlikely to hold at precisely those times when medical care demand is high: when individuals experience high rates of depreciation or later in the lifecycle when the rate of depreciation is increasing at a high rate.[113] Grossman otherwise subsumes unobservable depreciation in the error term despite the likely correlation between depreciation and the independent variable age. Wagstaff (1993) notes that Grossman's assumption is "at odds with the theoretical model and eliminates entirely the dynamic character of the net investment identity." Instead, Wagstaff assumes that depreciation varies by age and splits his sample into two age groups (over and under 41). Unfortunately, he finds nonsensical results that depreciation is not only lower, but negative for the older group.[114] Moreover, Wagstaff's method would not account for younger individuals who are in the top 5% of medical care users because they experience health shocks. Specification of a health production

---

[112] Van der Gaag and Wolfe (1982) p. 31: "The (Health) equation can be interpreted as either a production function or a demand function of health."

[113] See Grossman (1999) fn. 32 where he uses the identity: $\frac{-\tilde{H}}{\delta} = \varepsilon\delta^2 \frac{d\delta_t}{dt}$ to argue that net disinvestment in health is "small" relative to the rate of depreciation "at modest rates of depreciation" and because the marginal elasticity of health capital, ε , is less than one. However, at low levels of health this marginal elasticity rises; and later in life the rate of change in depreciation rises weakening the assumption that is the linchpin of Grossman's empirical specification for the demand for medical care.

[114] Wagstaff (1993) in Jones & O'Donnel (2002) p. 17.

function has its own perils, in particular omitting relevant correlated variables such as diet and exercise. Berger et. al (1987) present a compelling critique that "the fundamental problem with the health production function approach is that it is hard to identify and measure all of the inputs that affect health."[115]

Since the purpose of the present analysis is to test the underlying theory, the nature of any health equation as a production or a demand function must be clear. Again, since the primary hypothesis is regarding the state of health and the change in health, consequences of mis-specification of the health function are acute. The effect of model mis-specification on the inferences in other MIMIC papers with different goals is potentially less severe.

Finally, a few studies of the demand for medical care have used a principal components analysis (PCA) to create a single index of health (van de Ven and Wolfe (1982), Wagstaff (1986) and many examples in the biomedical literature such as Zhang et. al (2006)). However, PCA critically assumes that the mean and variance are sufficient statistics to describe the joint distribution of the data.[116] This assumption is more tenable in the biomedical literature where inputs are typically continuous and normally distributed (e.g. blood pressure, peak heart rate and other biomedical measures) rather than in models that use discrete indicators from survey data. It is difficult to maintian these assumptions with ordered categorical data that exhibit significant non-linearity between the categories, which is the case with SAH as well as other available health state instruments in the BHPS.

[115] Berger et. al (1987) and references therein. Quote from p. 33.

[116] See Shlens (2005) for a particularly accessible introduction to PCA.

Like PCA, Multiple Correspondence Analysis (MCA) aims to reduce the dimensionality of a matrix of values. MCA is a weighted form of PCA that takes as its weights the distance between the categories, which need not be defined by the variance as in PCA.[117] MCA produces scale values for each category that maximize the average squared correlation between each category and the resulting index or "score vector."[118] In other words, the scale values are the weights assigned to each component of the index with the objective to maximize the variation in the resulting index by maximizing the correlation between each component and the index.

Thus, the justification for using a multiple correspondence analysis for the present study are: 1) MCA incorporates multiple health indicators in a single multi-dimensional health index that does not risk multicollinearity and thereby supports inferences on the health state; and 2) MCA is a statistical method that does not require specifying either a health production or health demand function. In addition, MCA is particularly well suited for the present analysis focused on the change in health because it maximizes the variation in the resulting health index among the observations in the sample including among individuals over time. This maximum variation allows for more observations of changing health. By contrast, if I were to use only the single measure of SAH, I would observe relatively few changes (e.g. from poor to very poor) and the direction of the changes would be truncated for two of the five categories (individuals can only go down from excellent health and up from very poor health.) There are several examples of correspondence analysis for health states using two instruments: SAH and activities of daily living (see Erickson et. al. (1995) and references therein and Hadley and Weidman

[117] Greenacre (2001) p. 3.
[118] Greenacre (2002) p.165.

(2006)). However, I know of no literature that extends this to multiple categorical instruments in a model of medical care demand.

4. Unobservable Price

There is very little literature on modeling the unobservable price of medical care, and I know of no examples of MCA used for this purpose. When testing his theory, Grossman (1972, 1999) dismisses the price of medical care from his empirical specification by assuming that this price does not vary across the individuals in the sample or that as an omitted variable it is not correlated with the other regressors.[119] Similarly, Wagstaff (1986, 1993) does not include a price variable when testing Grossman's theory. Unlike the present study, neither Grossman nor Wagstaff imposed economic restrictions on their empirical models when testing the hypotheses associated with the underlying utility maximization theory. In addition, neither were either interested in making inferences on price or cross-price parameters. As a result, neither considered the omission of price a significant issue. Propper (2000) considers price a significant element of her theoretical development, but does not empirically examine the impact of price because "the prices paid by individuals are not observed" in the BHPS.[120]

The seminal work associated with the effect of price on the demand for medical care is from the Rand Health Insurance Experiment (RHIE). However, rather than estimating a price-elasticity, the study in effect estimated an insurance elasticity since the "price" parameters were different levels of insurance co-payments (Manning et. al. (1987)). Keeler et. al. (1977) and Ellis (1996) further explored theoretically and

[119] Grossman (1972b) p. 41. Grossman (1999) p. 43 acknowledges that "neither assumption is likely to be correct in light of the well known moral hazard effect of private health insurance."
[120] Propper (2000) p. 864.

empirically the non-linear price function induced by insurance deductibles. Nonetheless, many studies that include a price variable use an indicator of insurance coverage (or different co-payment rates) rather than any model for unobservable price (Cameron et. al. (1988), Pohlmeier and Ulrich (1995) Deb and Trivedi (1997)).[121] One key reason is lack of specific price information. Eichner (1998) overcame this problem in part by using a proprietary data set of insurance claims from a large employer. His estimates of the price-elasticity of medical care were consistent with the RHIE estimates; however his sample was restricted to workers between 22 and 55 years old, and his specification did not include any health state indicators other than age and gender. Therefore, like the RHIE estimates it is questionable whether these results are generalizable to the population that makes up the top 5% of spenders.

Finally, one aspect of the unobservable price of medical care that is occasionally included in demand for medical care models is the opportunity cost of time to see the doctor. Grossman models time as an input to the household production of health but does not include time as an independent variable in his empirical specifications. The seminal work on the time-price of medical care remains Acton's 1975 study of the travel-costs to New York City clinics. Acton concluded that time measured as travel distance does function as price when care is nominally free, and that the elasticity of distance approaches the money-price elasticity.[122] More recently Janssen (1992) found that time prices measured by wage and employment status were significant in the demand for general practitioner visits and "ignoring time prices could result in the mis-specification

[121] Deb and Trivedi (1997) p. 324 acknowledge that using Medicaid status as an indicator of the price of medical care likely captures income (poverty) effects.
[122] Acton (1975) p. 610.

of demand equations."[123] Nonetheless, most of the work on the time-price of medical care has been in the public health literature due in part to the lack of data on travel and waiting time in the large surveys typically used in the economics literature.[124] The economics literature tends to use more blunt proxies for waiting and travel time. An exception is Wagstaff (1986) who computes time cost for hospitalization and GP visits as distance times wages, but these parameter estimates are insignificant in his regressions. Propper (2000) matches the BHPS to NHS data on regional waiting lists (over and under 12 months) and includes these times as controls but does not report coefficient estimates. Several other economic models include an urban/rural indicator and an indicator of service availability as proxy controls for travel time and service availability (Pohlmeir and Ulrich (1995) and Wagstaff (1986, 1993) use population and physician density, Hunt-McCool et. al. (1994) and Deb and Trivedi (1997) use region). None of these studies make inferences on these parameters with respect to price elasticity of demand because as blunt indicators they likely capture other socioeconomic factors.

While the empirical literature on proxying for unobservable price is limited, the significance of the opportunity cost of time, particularly in situations where the money-price trends towards zero, appears well founded both theoretically and empirically. The literature has identified several instruments that can proxy for unobservable price: insurance status, waiting lists, employment status, wages, geographic region, and provider density (e.g. number of hospital beds.) As with the multiple indicators of the unobservable health state, a MCA analysis can incorporate these different indicators into

[123] Janssen (1992) p. 725.

[124] See Goodman et. al. (1997) and references therein for a brief review of the relationship between travel time and hospitalization.

a single indicator of unobservable price that maximizes the variation in the data and support inferences on the price parameters.

II. Empirical Specification

In order to test the theory the empirical specification must be consistent with it. Two key features of the theory are that the demands for medical care and consumption should be estimated jointly and that significance and sign of the relationship between health and consumption is not assumed a priori. To reflect this, I specify a system of flexible Marshallian demands that relax assumptions of strict separability between consumption and medical care and health and consumption. A negative semi-definite Slutsky matrix and symmetry of the cross-price parameters is necessary and sufficient for these demands to reflect an admissible utility function.

I estimate demands for three categories of medical care: j = 3 = {hospital days (H) tests and services (TS) and general practitioner visits (GP)} along with the demand for consumption, Z. I estimate three separate medical care demands to avoid either arbitrarily combining such different services as inpatient hospital days with outpatient eye exams or eliminating any particular service.

The Marshallian demands are as follows:[125]

$$Z_t = \exp\left\{ b_Z + c_{ZZ} p_t^Z + \sum_{j=1}^{3} c_{Zm_j} p^{m_j} + b_{ZZ}(NI_t - S_t) + b_{Z\alpha}\left(H_t^* - H_{t-1}^*\right) + b_{ZH} H_{t-1}^* + b_{Zc} Couple_t + b_{ZE} Edu_t \right\} + \varepsilon_Z$$

$$m_{jt} = \exp\left\{ b_{m_j} + c_{m_j Z} p_t^Z + \sum_{k=1}^{3} c_{m_j m_k} p^{m_j} + b_{m_j m_j}(NI_t - S_t) + b_{m_j \alpha}\left(H_t^* - H_{t-1}^*\right) + b_{m_j H} H_{t-1}^* + b_{m_j c} Couple_t + b_{m_j E} Edu_t \right\} + \varepsilon_{m_j}$$

(22)

[125] See Appendix A for nomenclature for the empirical model. Superscripts denote demands and not exponents.

The errors are assumed to be additive, correlated across equations and clustered by individual.

The c's and b's are the parameters to be estimated, and the hypotheses to be tested are:

$H1: b_{m_j a} < 0$: The greater the decline in health the greater the demand for medical care

$H2: b_{m_j H} < 0$: The higher lagged health, the lower the demand for medical care. This would be consistent with the assumption from the theory that $\alpha_H > 0$, or the higher the state of health the less negative the change in health and thereby the less demand for medical care. It is also consistent with $\alpha_{mH} > 0$ which is the co-morbidity hypothesis that the higher state of health the higher the productivity (the better the prognosis) from any medical care. The more productive the care, the less care is needed to produce the optimum level of health demanded.

$H3: c_{zm_j} \neq 0$: The demands for consumption and medical care are not separable.

$H4: b_{za}, b_{zH} \neq 0$: Health and consumption are not separable. If significant, the sign of these coefficients will suggest the sign of the cross-partials $U_{Z\alpha}$ and U_{ZH}.

The theory suggests that both the change in health, $DH_t = H_t - H_{t-1}$, and the state of health, $LH_t = H_{t-1}$, are inputs into the individual's demand for both medical care and consumption (see the time paths for these demands equations (18) and (21) respectively). Critically, the state of health is assumed to be weakly exogenous. The state of health is only weakly exogenous as future states are assumed to be determined by the past demands for medical care. This is a restriction on the theory that determines health endogenously within the optimal control model. However, the economic restrictions associated with symmetry of the cross-price parameters and a negative semi-definite Slutsky matrix assume that the unspecified utility function is maximized subject to all applicable constraints. One of these constraints is the health transition equation that determines the relationship between the demand for medical care in one period and next period's health state and thereby determines the change in health. In other words, the demand for medical care is conditional on the state of health that individuals optimally demand. Given the optimal demand for health, individuals demand medical care to achieve that level of health consistent with lifetime utility maximization; they do not demand medical care and merely accept the state of health that results. This reflects Grossman's derived demand hypothesis and reasonable in a deterministic model.

Similarly, contemporaneous net income is assumed to be weakly exogenous to the demands.[126] Following the reasoning regarding the weak exogeneity of the health state, current health is assumed to affect future income. Still, the wage function is an element of the wealth transition function which is assumed to be satisfied by the negative semi-definiteness of the Slutsky matrix. In other words, the demands take the wealth constraint as given. In this case, the wealth constraint consists of income (including any non-wage transfer payments that may be associated with the state of health) as well as savings or debt, which is completely determined by the wealth transition equation.

The wealth constraint is modeled as net income minus (plus) savings (debt): (NNI-S). The specification of the budget constraint to include savings and/or debt allows the demands to be complimentary at any point in time. This is consistent with the theory

[126] Using lagged net income would be more clearly exogenous and impose the tenable assumption that income is unavailable for use until the end of the period. However, given the structure of the BHPS it is problematic to link households over time. As a result, using lagged income would merely trade an exogeneity problem for substantial measurement error for individuals who change households over the sample period.

that allows for debt at all times up until death and extends empirical models that include only contemporaneous labor income. Not allowing for savings and/or debt in the construction of the budget constraint would impose the condition that the demands for consumption and medical care must be substitutes in each time period since there is no mechanism in the model for intertemporal wealth transfers.[127]

The p^{m_j} price variables represent the individual's point-of-purchase price of medical care including opportunity cost and net of insurance. The model and instrument weights for the unobservable price of medical care is discussed in more detail in the following section. Finally, following prior literature I also control for whether the individual is married or living as a couple and the highest level of attained education.[128]

III. Data Description: The British Household Panel Survey

The theory is tested on the first 14 waves of the British Household Panel Survey (BHPS): 1991 - 2005.[129] The BHPS is an annual survey of adult (16+) members of households nationally representative of the UK.[130] The BHPS began with roughly 5,000 households (over 9,000 full adult individual interviews) in 1991 and added several subsamples over the study period: the United Kingdom European Community Household

---

[127] This is of little importance in extant single-equation models that use contemporaneous net income because of the inherent assumption that demands are independent.

[128] There are extensive literatures regarding the relationships among education, marital status, health and the demand for medical care. A central hypothesis in Grossman (1972) is that education increases household productivity and thereby decreases the demand for medical care. Andersen and Newman (1973) is the seminal work that posited the theoretical relationship between marital status and the demand for medical care. See Wilson and Oswald (2005) for a review of the empirical literature, Kiecolt-Glaser and Newton (2001) and Williams and Umberson (2005) and references therein for a more recent discussion of the different theories of the effect of marital status on health and empirical estimates using panel data. Age and sex are used as instruments for the unobserved health state and not as independent fixed inputs to utility. See the subsequent section on the MCA results for the health state for a further discussion.

[129] See Taylor et. al (2007) for details on the BHPS and Levy et. al (2006) for details on the net income data.

[130] Appendix D includes a complete list of BHPS variables used in the analysis.

Panel (EHCP) from 1997 to 2001 (waves G through K); the Scotland and Wales Extension from 1999 onward (waves I through N); and the Northern Ireland Household Panel Survey (NIHPS) from 2001 onward (waves K through N). The addition of these samples to the BHPS is evident in the increase in sample size in waves 7, 10 and 12 in the table below. Since the focus is the change in health, only individuals with at least two consecutive waves of complete individual response data are in the sample. Given the measurement error inherent in self-reported health, no proxy responses are included. In addition, observations with non-positive net income minus savings (279 observations are negative and 244 are zero) are deleted resulting in an unbalanced panel of 119,750 observations. I do a sub-analysis on a balanced panel of 3,148 individuals (40,896 observations) with full responses for all waves. This is referred to as the "OSM" sample for "original sample members." All monetary figures for consumption, net income and savings are converted from nominal to real using the UK cpi with the base year of 2005 and expressed in thousands of pounds (000's).[131] Figure 9 shows the number of observations per wave and the number of waves each unique individual appears in.[132]

Figure 9: Sample Size by Wave and Number of Waves

Observations by Wave, Full Sample

wave	Freq.	Percent	Cum.
2	7,812	6.52	6.52
3	7,167	5.98	12.51
4	7,165	5.98	18.49
5	7,082	5.91	24.41
6	7,009	5.85	30.26
7	7,306	6.10	36.36
8	8,685	7.25	43.61
9	8,502	7.10	50.71
10	11,633	9.71	60.43
11	11,752	9.81	70.24
12	12,188	10.18	80.42
13	11,862	9.91	90.32
14	11,587	9.68	100.00
Total	119,750	100.00	

Number of Waves for Individual Responses

nwaves	Freq.	Percent	Cum.
1	3,187	2.66	2.66
2	4,599	3.84	6.50
3	9,853	8.23	14.73
4	9,249	7.72	22.45
5	12,751	10.65	33.10
6	3,731	3.12	36.22
7	4,173	3.48	39.70
8	4,707	3.93	43.63
9	5,418	4.52	48.16
10	6,127	5.12	53.27
11	10,730	8.96	62.23
12	4,329	3.62	65.85
13	40,896	34.15	100.00
Total	119,750	100.00	

[131] CPI data from http://www.statistics.gov.uk/statbase.

[132] The Number of Waves table shows that 40,896 individuals appear in all 13 waves. This comprises the OSM sample. There are 13 waves instead of 14 because one wave of data is lost in order to compute the change in health.

Figure 10 shows summary statistics for all variables for both the full and OSM samples. These variables, particularly those that are derived, are discussed more thoroughly below.

Figure 10: Sample Statistics, FULL and OSM Samples

Summary Statistics: Pooled Full Sample

Variable	Obs	Mean	Std. Dev.	Min	Max
Zrk	119750	7.328642	4.389378	0	206.9547
Hdays	119750	.9758664	6.278687	0	365
TS	119750	3.092985	2.331129	0	16
GP	119750	2.415816	1.207569	0	5
pz	119750	1.10851	.0761688	1.020408	1.24844
ph	119750	2.976964	1.252225	0	5.801143
pts	119750	2.24767	1.277249	0	5.283835
pgp	119750	2.845605	1.227373	0	4.748104
NNIk	119750	20.97904	14.33852	.0006122	750.9901
DHt	119750	-1.355056	8.83061	-74.83315	68.49057
LHt	119750	83.79338	14.70637	-6.72e-06	100
couple	119750	.6586305	.4741711	0	1
edu	119750	.6018121	.4895266	0	1

Consumption (Zrk) and Net Income minus Savings (NNIk) are in '000 real pounds

Summary Statistics: Pooled OSM Sample

Variable	Obs	Mean	Std. Dev.	Min	Max
Zrk	40896	7.219372	4.438979	0	206.9547
Hdays	40896	.741564	5.127983	0	330
TS	40896	3.092747	2.257932	0	14
GP	40896	2.371454	1.171202	0	5
pz	40896	1.1235	.0781863	1.020408	1.24844
ph	40896	3.079934	1.267137	.1865218	5.801143
pts	40896	2.399353	1.300301	0	5.283835
pgp	40896	2.914839	1.248871	0	4.735017
NNIk	40896	21.06166	13.44174	.0006213	344.963
DHt	40896	-1.272998	8.085671	-73.38447	68.49057
LHt	40896	84.51561	13.08316	3.990636	99.99999
couple	40896	.746919	.4347823	0	1
edu	40896	.607859	.4882338	0	1

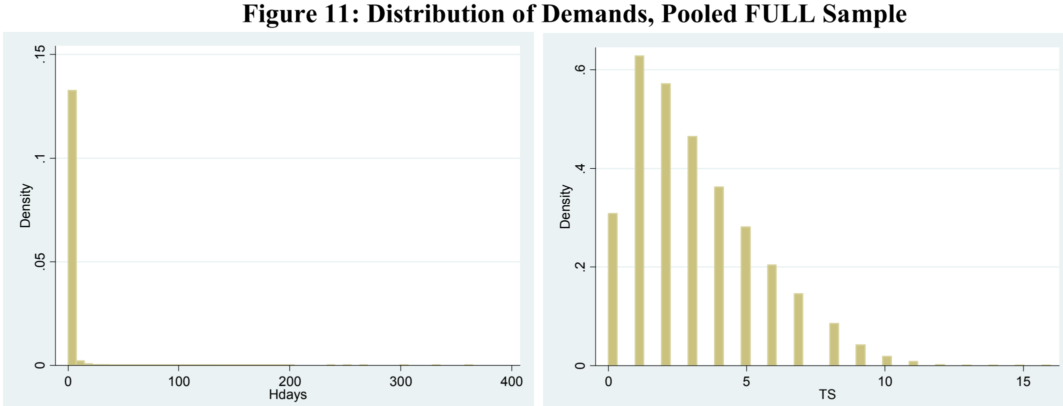
Consumption (Zrk) and Net Income minus Savings (NNIk) are in '000 real pounds

The summary statistics for the FULL and OSM samples are substantially similar. The OSM sample has slightly better average health (84.5 vs. 83.8) and a smaller decline in health (-1.27 vs. -1.36) and is more likely to be living as a couple (.75 vs. .66.) This is consistent with those in better health and more stable relationships being more likely to persist with the survey over time. The OSM sample also has slightly lower values for all of the dependent variables including consumption despite having a higher average budget constraint with a lower standard deviation (NNIk). The biggest difference between the samples is in average hospital days (.74 vs. .98) which may again be reflective of individuals dropping out of the survey due to hospitalizations.[133]

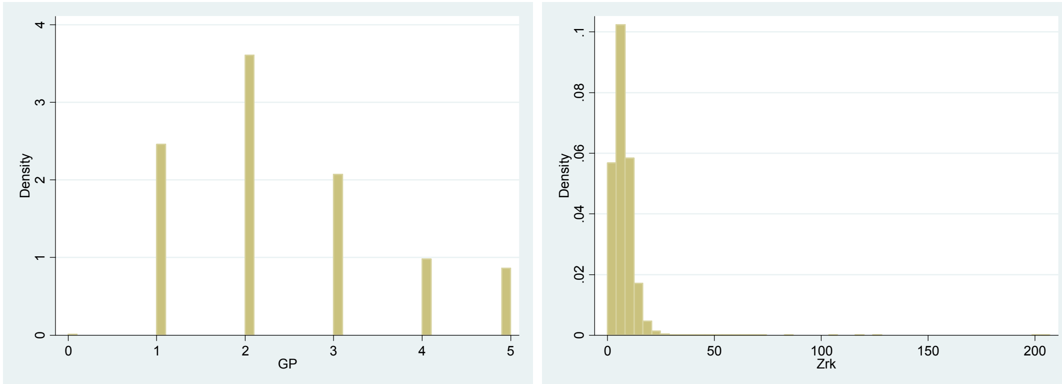
A. Demands

I estimate three different categories of medical care demand: hospital days, tests and services, and general practitioner visits. Hospital days ('w'hospd) is a continuous variable from 0 to 365 with the expected high frequency of zeros and long right tail characteristic of medical usage data. I construct an index for tests and services by aggregating the variables for different health services listed in the BHPS (see Appendix B) with additional tests and services added as they become available in the data. The range for TS is from 0 to 15, with a high frequency of zeros and long right tail as illustrated in the histograms in the figure below. General practitioner visits ('w"hl2gp) is a categorical variable ranging from zero to 4 where 4 represents six or more contacts with a general practitioner (either in person or by phone) per year.[134]

Figure 11: Distribution of Demands, Pooled FULL Sample



[133] There are several individuals in the BHPS who would be in the OSM sample but for having been hospitalized at the time of the survey. These individuals have either proxy or missing responses, but I do not include proxy respondents in my sample as noted above.
[134] Unfortunately the BHPS does not have data for pharmaceutical use which may result in omitted variable bias.



While the BHPS has more consumption data than most household surveys, the data available consistently in all waves is primarily for non-discretionary items such as food, fuel and housing. There is no data for spending on transportation and no consistent data across the waves for important discretionary items such as clothing and entertainment. To the extent that individuals make a trade-off between health and consumption, that trade-off is likely to be more apparent in discretionary consumption than non-discretionary consumption. Moreover, all of the consumption variables are by household. I apply household figures to each member of the household even though the individual may not have participated in each consumption decision.

The aggregate consumption variable is the annualized amount spent on food (‘w’xpfood) fuel (‘w’xpfuel), housing (‘w’xphsn) and consumer durables. The amount spent on consumer durables is bcdnuxp for wave B and an aggregate of the price paid for the specific items for subsequent waves. The food variable was transformed from discrete to continuous by taking the top of each range and arbitrarily using 179 as the maximum weekly spending. There is no fuel variable for wave F: it is linearly interpolated from waves E and G at the individual level. [135] Like the demands for medical care, consumption is non-negative and highly skewed.

B. Independent Variables

The derived annual household net income ('w'hhyneti) is from Levy et. al (2006.) This includes labor and non-labor income from investments, pensions and benefits and is net of taxes including national insurance contributions. Like the consumption variables, net income is only available by household and is distributed to each individual within the household. Conversely, savings is an individual variable. Individual savings ('w'saved) is made consistent with household net income by aggregating savings within the household and then distributing the aggregate figure back to each individual. Applying household financial data to individuals is consistent with the perception that an entire household's resources are available to address the health needs of any individual member of the household. Finally, while the theoretical model and empirical specification allow for debt, there is no specific dollar value of debt available in the BHPS. While the BHPS has categorical variables that indicate the presence of debt, I do not attempt to incorporate these in the analysis at this time.[136] The only reflection of debt in the data is to the extent that the computed household net income includes investment income that would be reduced by interest payments and transfer income that would include loans and to the extent that tax credits include interest deductions.

---

[135] The risk is that the interpolated figure is from two different households if the individual PID changed households from wave E to wave G.

[136] The BHPS has several categorical questions on whether or not the individual has credit card debt, student loans, and other personal debt. However, these questions are only asked in waves E,J and O, are not numerical variables, and do not indicate the time period over which the money was borrowed for consumption.

The other directly observable independent variables are marital status and education. The variable ‘w’mastat is transformed to 1 = married or living as a couple in the variable “couple.” Not living as a couple, whether because never married, divorced or widowed is the excluded category. Similarly, the highest five education categories from ‘w’qfedhi (higher degree, first degree, teaching QF, Other Higher QF and Nursing QF) = 1 and all lower categories are omitted. Thus, the coefficients on “couple” and “edu” measure the incremental effects on consumption of being in a couple and being “highly” educated relative to being single and relatively less educated.

C. Unobservable Health

1. Methodology for MCA for unobservable Health

I use instruments for the MCA analysis for the health state that have been suggested by the literature: sex, age (at date of interview), self-assessed health, subjective well-being, reported health problems, reported limitations in daily activities, whether the individual is disabled, the number of accidents, marital status and whether the individual smokes.[137] See Appendix D for a detailed description of BHPS variables.

The BHPS changed the wording of some of the health related questions over time. The most significant issue for the present study is that in waves I and N the questions regarding limitations in daily activities (ADLs) are changed substantially.[138] For all other waves the ADL questions are: problems with housework, climbing stairs, dressing,

---

[137] There is a long literature on the relationship between smoking and health. For a recent theoretical work with a life-cycle perspective on smoking and the demand for medical care see Carbone et. al (2005). There is a large literature in sociology and psychology indicating that marital status has an affect on health, and may have differing effects on men and women. See Ross et. al. (1990) for a review and Contoyannis et. al (2004) document the difference in self-assessed health between men and women in the BHPS for waves 1–9.

[138] The code but not the wording of the SAH question was changed for wave I but not for wave N. The code for wave I is ihlsf1.

walking and other. These questions have two response options: yes and no. In waves I and N individuals are asked a longer, more detailed list of questions and are given three response options: a lot, a little or none. I experimented with several different specifications for these two waves with the goal of obtaining weights broadly consistent with the other waves. The following table lists the questions used for waves I and N. Comparable categories are detailed in the results table, and results for the full set of instruments for waves I and N are reported separately.

Figure 12: List of ADL Limitations for Waves I and N

name	code
Vigorous activity	w'hlsf3a
Moderate activity	w'hlsf3b
Grocery shopping	w'hlsf3c
Several flights of stairs	w'hlsf3d
One flight of stairs	w'hlsf3e
Bending and kneeling	w'hlsf3f
Walking 1 mile	w'hlsf3g
Walking 1/2 mile	w'hlsf3h
Walking 100 yards	w'hlsf3i
Bathing & dressing	w'hlsf3j
Limited time spent on work	w'hlsf4a
Accomplish less at work	w'hlsf4b
Limited kind of work	w'hlsf4c
Difficulty performing work	w'hlsf4d
Limit social activity	w'hlsf6
Amount of pain	w'hlsf7
Pain interfered with work	w'hlsf8

There are a few additional changes in the BHPS over time. The question associated with current smoking was changed from “do you smoke cigarettes” to “do you smoke cigarettes at all nowadays” and the code changed from ‘w’smoker to ismnow for wave I only. Two reported health problem problems, cancer and stroke, were added to the survey in wave K; however I do not included them in the present study to preserve as much consistency as possible in the health index across the waves. Finally, the question regarding disability was changed from “Can I check, are you registered as a disabled person, either with Social Services or with a green card?” in waves up through J to “Can I check, are you registered as a disabled person" in wave K to "considers self to be disabled" in wave L and then "Is (NAME) registered as a disabled person" in wave M. The code is the same, 'w'hldsbl, except for wave L where it is lhldsbl1.

Weights for the health index are computed separately for each wave of the BHPS for three reasons. First, since the goal is to test the theory, the priority is to measure variables consistently within the sample rather than within the population for out-of-sample forecasting. However, weights are computed using all available observations with individual responses for the required variables in each wave to maximize the sample size and thereby the precision of the estimated weights. This assumes that the estimation sample is similar to the excluded sample (e.g. individuals who do not have two consecutive waves of complete individual responses) in terms of the instruments used to index for unobservable health. Second, a wave-by-wave analysis can address changes in question wording without contaminating the entire sample. Finally, a wave-by-wave analysis can incorporate unobservable time-varying societal preferences that may alter the instrument weights. For example, poor economic times may increase the number of individuals who report limitations with work because health limitations may be a more socially acceptable reason for being unemployed. Assuming that this incentive is felt the same way by all members of the sample, time-varying changes in macroeconomic and other factors would affect the health index consistently within each wave.

2. MCA results for the Health State

The weights applied to the different instruments that comprise the health index are intuitively reasonable and broadly consistent with the literature. However, the changes in waves I and N are apparent in both the weights and the resulting values for health. The weights for each instrument for each wave are reported in Appendix E.

The weights for men and women are of opposite sign but of almost equal absolute magnitude across all waves. This is consistent with the perception in the literature that men report higher health than women (Ross et. al. (1990) and references therein). The health index reflects this without separately estimating weights for men and women thereby maintaining a larger sample size necessary for more precise measurements for some of the less frequently observed ages and health instruments. The absolute magnitude of both weights is lower for wave I but not for wave N.

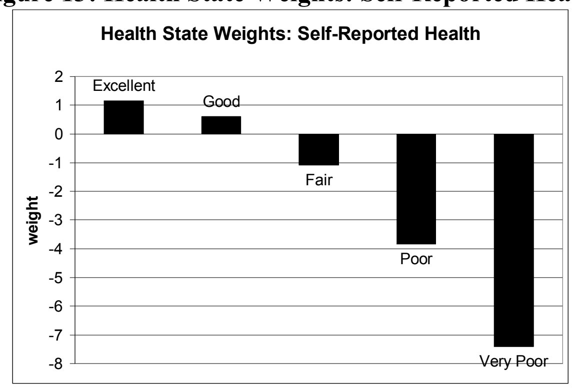
As expected, the weights decrease consistently (though not monotonically) with age.[139] The inflection point between positive weights and negative weights is between ages 46 and 49. In no wave is there a negative weight younger than age 43 or a positive weight older than age 53. Notably, the oldest positive weight occurs in wave I when the reported health problems are expanded (though the oldest weight in wave N with the same list of reported problems is 49 consistent with several other waves.) This is consistent with the literature that suggests the association between age and health is at best a noisy indicator. The weights do not decrease monotonically in part due to the significant variation in sample sizes in the different age groups. The variance in the mean weight for each age across the waves is U-shaped from 16 to 56 (with a minimum at age 36) and then increases significantly at older ages. For example, the weight for the oldest age ranges from a low of -1.7 in wave D to a high of -20.175 in wave B. This reflects a

[139] While all reported ages are used in the analysis, the summary table of weights lists ages 16 (the youngest age in the sample), 36, 56, 76 and the oldest age in the sample, which ranges from 93 in wave C to 99 in wave J.

significant survivorship bias across waves which are statistically apparent in the small sample sizes in many of the highest age categories. .

The weights associated with SAH decrease monotonically and non-linearly as illustrated below. The distance between the weights increases at an increasing rate as the category of health declines.[140]

Figure 13: Health State Weights: Self-Reported Health

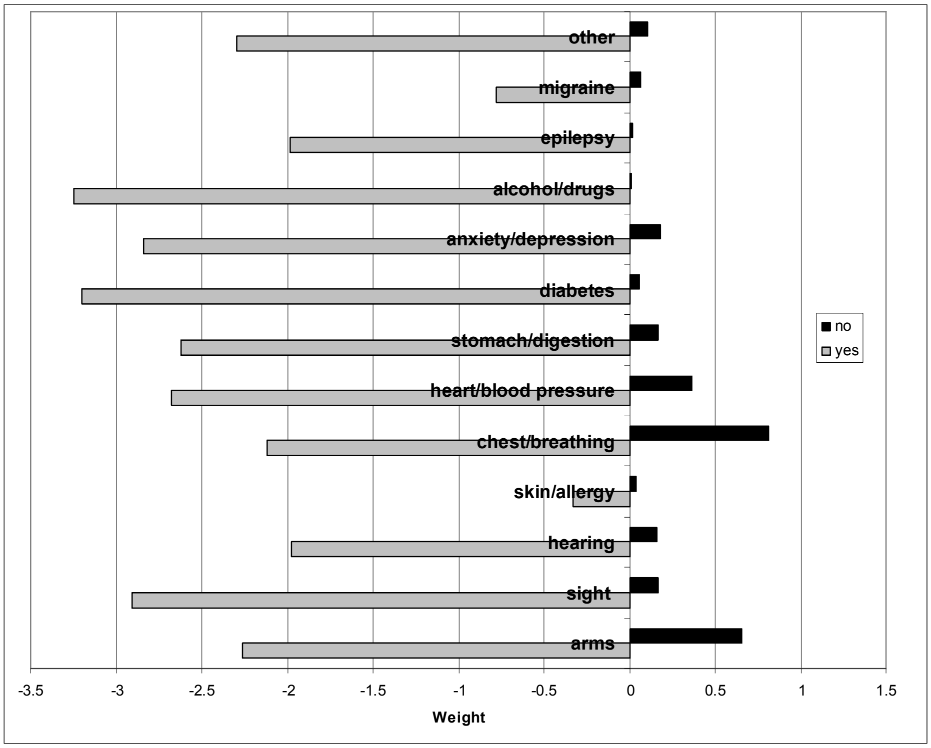


Like SAH, the weights for the 36-point mental and emotional health index decrease with higher values (indicating poorer health) of the index. However, like the weights associated with age, the decrease is not monotonic and the variation across waves increases with increasing values owning to the small sample sizes in many of the higher categories. The range of instrument weights for “0” indicating the highest mental and emotional health is from .604 in wave B to 1.299 in wave A while the range for “36” the highest category is from -2.532 in wave B to -11.535 in wave A.

There are interesting differences in the weights associated with different reported health problems. As illustrated in the figure below, alcohol/drugs, diabetes and problems with sight detract most from health while the absence of breathing problems and problems with arms are the most beneficial.

[140] The figure illustrates weights for wave B but is consistent with the weights for all other waves. Weight estimates for all waves are in the Appendix.

Figure 14: Health State Weights: Reported Health Problems



It is interesting and significant that women tend to report more health problems, but problems that carry lower weights than men. The Appendix report the percentage of each gender reporting each health problem at each wave, and the figure below illustrates the average percent reporting and the difference between the genders over all waves.

Figure 15: Percent Reporting Health Problems by Gender

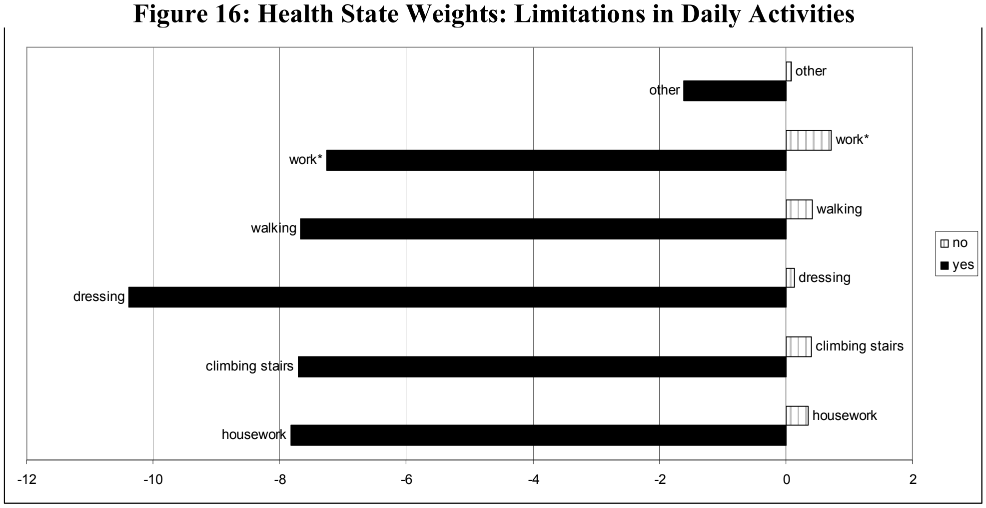
	Average % Reporting Over All Waves			Weight	
	Women	Men	Difference	yes	no
arms	**28.99%**	**23.32%**	**5.67%**	**-2.262**	**0.657**
sight	5.73%	4.49%	1.24%	-2.911	0.164
hearing	6.97%	9.48%	-2.51%	-1.975	0.162
skin/allergy	**13.23%**	**9.17%**	**4.05%**	**-0.332**	**0.037**
chest/breathing	13.16%	12.35%	0.81%	-2.119	0.816
heart/blood pressure	15.42%	13.06%	2.36%	-2.679	0.367
stomach/digestion	7.74%	6.11%	1.63%	-2.624	0.165
diabetes	2.48%	3.22%	-0.73%	-3.199	0.061
anxiety/depression	**9.34%**	**4.61%**	**4.73%**	**-2.839**	**0.179**
alcohol/drugs	*0.26%*	*0.70%*	-0.44%	*-3.250*	*0.010*
epilepsy	0.82%	1.06%	-0.24%	-1.980	0.015
migraine	**11.89%**	**4.38%**	**7.51%**	**-0.782**	**0.066**
other	5.79%	3.32%	2.46%	-2.294	0.104
Average weight for all problems				-2.250	0.216
Average weight when women reort more frequently				-2.094	0.284
Average weight when men reort more frequently				-2.601	0.062

Nine of the 13 categories have a greater proportion of women reporting problems. Of these, arms, skin, anxiety and migraine have at least 4% more women than men reporting problems. However, two of these, skin and migraine, which 7.51% more women than men report, have the lowest of all the weights. By contrast, the problem with the highest weight, alcohol/drugs, is more frequently reported by men. As a result, the average weight for the problems more frequently reported by women is lower than that of men and lower than the average of all weights. At the same time, the weight associated with not reporting the problem is higher for those problems women report more frequently. So, even though women report more health problems, the impact on the resulting health state is muted because the different weights.

In an initial MCA specification I aggregated the reported health problems and used the number rather than the specific type of problem as the instrument. In this analysis the resulting index for health had two distinct modes: one for men and a lower one for women. These two modes occurred even with sex used as an instrument because women tend to report more health problems than men. However, this double mode disappeared when I listed health problems individually for the MCA because while women report more health problems, the problems they report tend to carry less weight in the index. This suggests that models that use indicators of health problems may exaggerate the gender differences associated with this measure of health.

The weights associated with different limitations in daily activities are less varied than the weights associated with different health problems. The figure below shows that the most debilitating is an inability to dress oneself, while the most beneficial is predictably no limitations in work.[141] There are some sex-differences in reported ADLs; but because the weights among the problems are more comparable than the weights associated with reported health problems the impact on the resulting health index is not dramatic. Less than 2% of men and women (approximately 1.5 and 1.9% respectively for each wave) report difficulty dressing. More individuals of both sexes report difficulty with "other" but in this case, slightly more men than women: (approximately 23% vs. 22% respectively in each wave.)

Figure 16: Health State Weights: Limitations in Daily Activities



[141] Data is for wave B but illustrative of all the other waves. Work limitations have five categories; only the highest and lowest, "a lot" and none, are illustrated in the figure.

The weights for ADL limitations in waves I and N, when the categories were disaggregated, are comparable with the corresponding categories in other waves indicated in Appendix E. However, a direct comparison is difficult because there are more ADLs and one more response option in these two waves. Moreover, the effect of the question changes for ADLs is apparent in weights for other questions that do not change. Most notably, the absolute magnitude of the weights for the top and bottom categories of SAH (but not the middle categories) for being disabled (but not for no disability), for being widowed or never married (but not the other marital status categories) and for most of the reported health problems are lower in waves I and N.

The weights for work limitations increase monotonically and non-linearly like the weights for SAH. There are five response categories from none to a lot (consistently through all waves). The weight for "no problems" is positive .75 on average with a low .034 standard deviation across waves. By contrast, the weight for "a lot" is -6.08 on average with a much higher 1.088 standard deviation. Unlike most of the instruments, there appears to be a declining trend in the absolute magnitude of all the work impairment categories with the high and low for the "a lot" category in waves A and N respectively.

As expected, the weight associated with being disabled is negative and much larger in absolute magnitude than that associated with not being disabled. Like the weights on work limitations, the absolute magnitude of the weight associated with disability appears to decline over time. Both trends may reflect improvements in private and social insurance for disability and workplace policies associated with disability discrimination and accommodations.

The weights associated with the number of accidents are negative for any accidents and generally increase in absolute magnitude with the number of accidents. However, there is more variation in these weights across the waves than in many other categories. In wave A the sign of 0 accidents is negative while the sign of 1 and 2 accidents is positive; this is the opposite sign for 0 for all other waves and for 1 and 2 for all waves other than C which has positive weights for 1 and 3 accidents. In most cases the weight for 4+ accidents is higher than that for 3 accidents except in waves I, J and L.

Somewhat surprisingly, the absolute magnitude of the weights associated with smoking are relatively small, averaging -.098 and .041 for smoker and non-smoker respectively. The only smaller weights on average across all waves are those for no accidents and married.

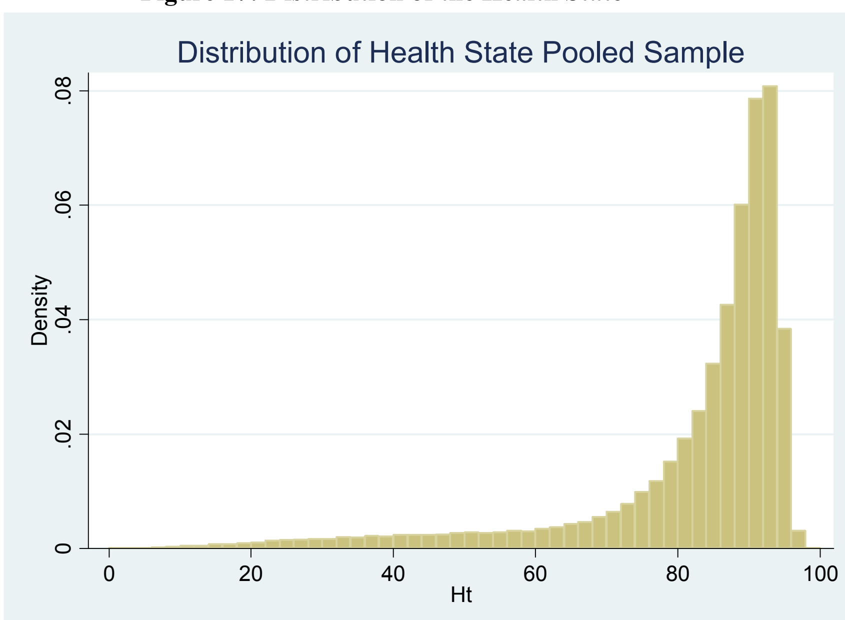
While the weight associated with being married is negative (after wave C) and very small, the weight associated with living as a couple is positive and larger: .749 on average. Interestingly, the weight associated with "never married" is nearly identical in all waves to that associated with "living as a couple." The weight associated with widowed indicates the greatest health impairment averaging -2.398. Women are much more likely to report being widowed: 13% in the pooled full sample vs. 4% for men. Given the magnitude of the "widowed" weight, the higher frequency of women being widowed may have a greater effect on the lower average health state than the negative coefficient on sex.

The health index is computed from the weights for each instrument using row scores based on the indicator matrix of the MCA.[142] I standardize the index to lie

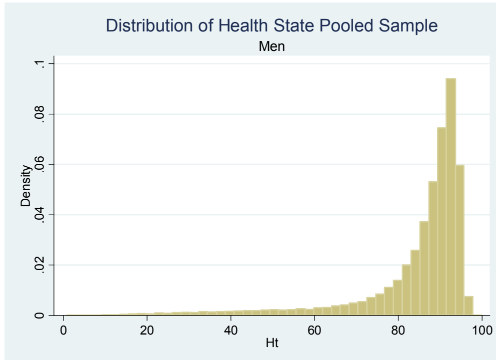
[142] See "Stata Help for MCA Post-estimation" for details on the prediction from the MCA analysis: http://www.stata.com/help.cgi?mca+postestimation.

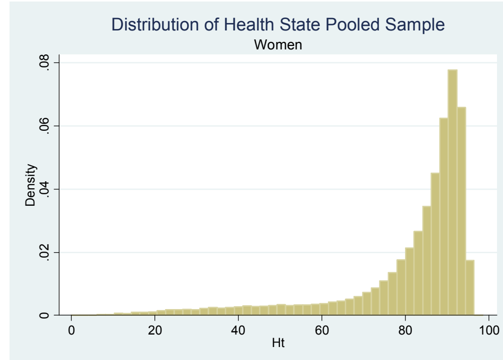
between zero and 100 to aide interpretation and reflect the theoretical assumptions of a minimum level of health to sustain life and a biologically fixed maximum longevity. The average health is 83.91 and it declines almost monotonically over time. Notably, the average health for waves I and N when the ADL questions changed is lower: 81.01 and 77.32 respectively. The standard deviation for these waves is higher: 18.81 and 20.17 respectively compared to an average standard deviation across the waves of 14.66. All of the multiple equation models were run without waves I, J and N and the results were similar and in no case did inferences change.[143] The distribution of health pooled for all waves has a single mode and is highly skewed as expected given the survivorship bias in the sample.

Figure 17: Distribution of the Health State



[143] Wave J is affected by the change in wording because wave I health is used to compute the change in health for wave J.



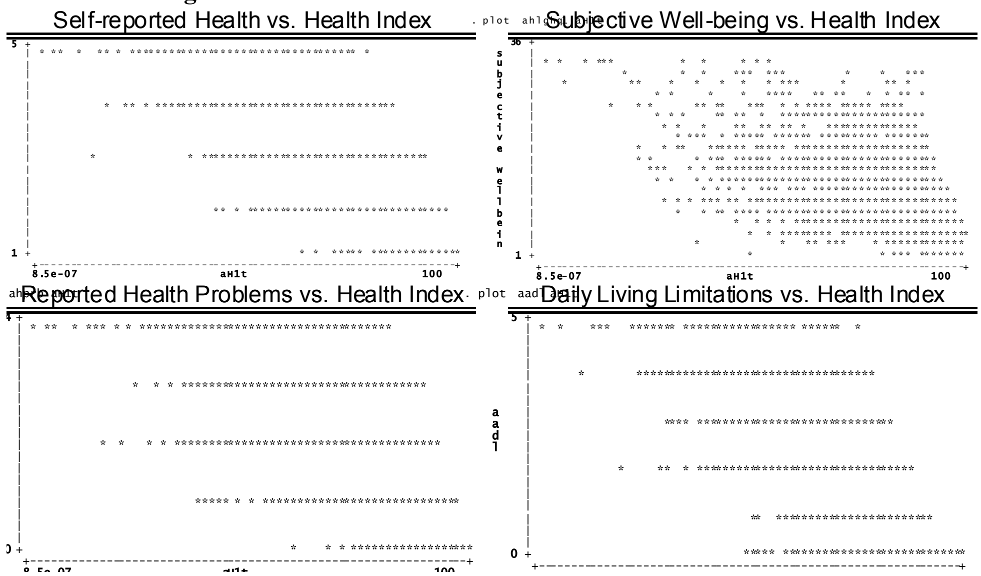


The distributions of the health state (pooled across all waves) are substantially similar for men and women. The mode is the same though at a lower frequency for women with greater mass both higher and lower. While the maximum for both men and women is 100 (99.72 for men) the minimum for men is much higher: 16.22 vs. 0 for women.

One goal of using MCA to compute a health index is to have one number that captured multiple dimensions of health. The computed health index appears to accomplish this goal: it is only weakly correlated with various commonly used single measures of health. The figure below shows scatter plots of the computed health state vs: a) self-assessed health, b) subjective well being, c) the number of reported health problems and d) the number of reported limitations in daily activities. These plots illustrate that individuals who report low levels of health (in each case reflected in higher numerical values) on the single dimension of health nonetheless have a full range of health states.[144]

[144] These plots are for wave A, but are consistent with plots from all other waves.

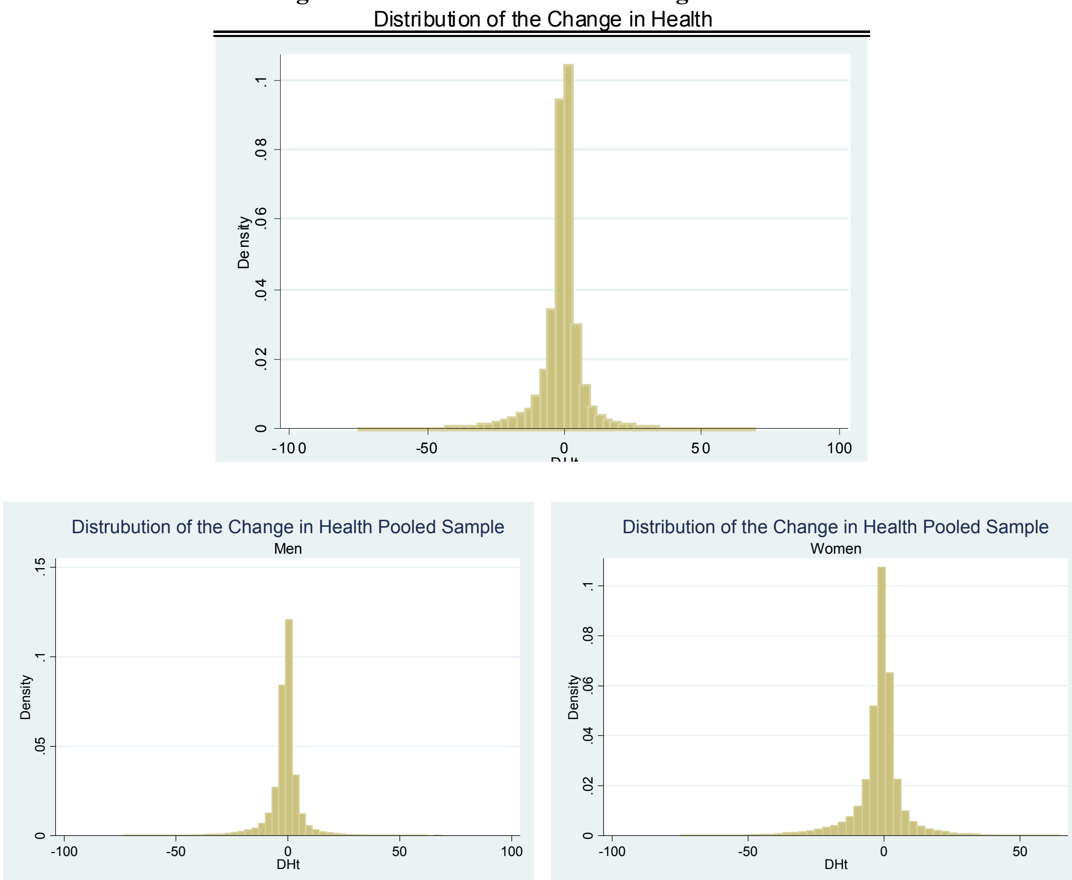
Figure 18: Health Index vs. Individual Measures of Health



A critical point for the present study is that there is a large variation in the change in health. This confirms the benefit of the BHPS for testing the hypothesis that the change in health is significant in the demand for medical care. The change in health is remarkably normally distributed. Figure 19 shows the distribution for the pooled full sample, which is consistent with the change in health for each wave and for the OSM samples.[145] The distribution is skewed towards declining health, which is to be expected given the construction of the BHPS and the aging of the sample.

[145] Observing significant improvements in health is not unique to the BHPS or this construction of a health index. For example, Blau and Gilleskie (2008) p. 490 observe that 20 – 25% of men in their sample experience improving health.

Figure 19: Distribution of the Change in Health



Women have a slightly greater average decline in health (-1.54 vs. 1.14 for men) and a much higher standard deviation (9.59 vs. 7.89 for men) despite a slightly smaller range ([-74.83, 64.52] for women and [-73.40, 68.49] for men.

D. Unobservable Price

<u>1. MCA Methodology for the Price of Medical Care</u>

The price instruments are: whether the service was provided by the NHS, private providers, or both, whether the individual has private insurance coverage (‘w’hlcvr), job status (‘w’jbstat) the wage rate (categorized in quartiles from ‘w’fimnl, the last regular net monthly income) and region (‘w’region).

The BHPS includes several questions about whether individuals pay for medical care services. The hospital days variable and the health check variables (‘w’…ckn) includes “both” as a response option, but the health services questions (‘w’…svcn) only include NHS and private options. The health services questions include a question about whether the service was free, paid for or both (‘w’…svcf); however, I use the NHS/private question to be consistent with the wording of the hospital days and health checks questions. A spot check of responses for the NHS/private vs. the free/paid questions suggests a high degree of consistency between NHS and free responses and private and paid responses for these services. There are very few (less than 1% conditional on any use) “both” responses for the health services payment questions.

I create an aggregate variable for NHS/private/both, ‘w’TSNHS, to correspond to the aggregate count variable for tests and services. This aggregate variable is a weighted average of the service providers for all the tests and services reported by the individual. In other words, if an individual uses all NHS services except for one private use, then he is coded as NHS, not both. This methodology under-counts the “both” responses for tests and services; however, counting any private service as “both” would over-count. I chose to under-count to reflect the NHS anti “top-up” rule that prohibits individuals from accessing both NHS and private services for the treatment of a single condition on a single visit. For example, individuals cannot be in a NHS hospital but get a drug that is not NHS approved.[146] The percentage of individuals who access both NHS and private services (conditional on using any service) is less than 2% for hospitalizations and 4% for aggregate tests and services in all waves.

[146] In 2008 there has been new debate about the NHS modifying the “top-up” rules and/or requiring co-payments under various circumstances. See the BBC for a contemporaneous review: http://news.bbc.co.uk/1/hi/health/7458908.stm

There is no question about NHS/private usage for general practitioner visits because the vast majority of GP visits in the UK are provided by the NHS. I include the categorical usage variable ‘w’gp derived from ‘w’hl2gp as an instrument for the price of general practitioner visits. This usage variable may reflect a few things. First, in principal-agent models of medical care demand individuals choose the first encounter with a physician, but then the physician determines the amount of future encounters. If this is the case, then the affect of “cost” on the idividual’s decision to contact a general practitioner should be higher for the first visit than subsequent visits. Second, the frequency of contacts with a general practitioner may proxy for the underlying health state and thereby the opportunity cost of receiving care. In other words, individuals who have more contacts with general practitioner visits may have lower underlying states of health which lowers their wage-earning capacity and thereby the opportunity cost of seeking care. Put another way, the costs of foregoing care are higher for those in poorer health.

I include an indicator for whether the individual did not use any care for all three types of care. This is included in the NHS/private variable for hospital days and tests and services and in the usage variable for general practitioner visits. Without including this, the resulting “price” would be conditional on use. Such a conditional price would be inappropriate to use in modeling individual decisions where the decision not to use care is not only possible but prevalent.

The BHPS includes a question about private insurance coverage starting with wave F: ‘w’hlcvr. This question includes a response for whether the insurance is in the individuals name or another family member’s name. I combined these categories to equal 1, indicating any insurance coverage. The percentage of individuals with insurance coverage is remarkably stable at about 20% from waves F through N. Given this stability, I apply the wave F response to individuals who appear in prior waves in order to consistently include this insurance variable in the price for all waves. Private medical insurance in the UK is typically for acute illness and covers hospitalization, specialist visits and tests; it does not generally cover general practitioner visits.[147] For this reason, the insurance variable is included as an instrument for hospital days and tests and services prices, but not for general practitioner price.

The BHPS question about job status, 'w'jbstat, asks the respondent for his or her *current* economic activity. I used current status rather than job status on September 1st , 'w'jbstatt, or job status a year ago September 1st 'w'jbstatl, to reduce any intentional or unintentional recollection errors and because the current status question has fewer missing values. However, this may result in mis-matched timing between medical care usage over the past year and any change in job status that occurred during the year. There are 10 categories: self employed, employed, unemployed, retired, maternity leave, family care, full-time student, long-term sick or disabled, government training service and other.

I create a categorical variable for wage, 'w'wage, for quartiles conditional on any wages from the question about labor income last month, 'w'fimnl. Thus, 'w'wage includes five categories, 0 indicating no wages, 1,2,3,and 4 indicating the quartile of wage. This question includes self-employment income but does not include any non-labor income from investments or government transfers as this income would not reflect

---

[147] See the consumer guide to private medical insurance produced by the Association of British Insurers: http://www.abi.org.uk/BookShop/ResearchReports/PMI%20Guide%20Web%20FINAL.pdf

any opportunity cost of time to seek medical care. I use the question for labor income last month rather than annual labor income to minimize reporting error and maximize variation in the data. However, as with the job status question there may be a mis-match in timing between current labor income and past use of medical care.

Finally, I include region, ‘w’region, as an instrument for the price of medical care. Following prior literature, the region variable can proxy for different distances individuals have to travel to receive care and the density of providers in different regions. However, other than “inner London” most of the regions as defined in the BHPS include both relatively urban as well as rural areas. The region variable may also reflect different costs-of-living as well as different local standards of care.

I follow the same methodology as with the MCA for the health state for the same reasons as stated in the prior section. I compute weights for each wave separately; all observations with complete data are used even if these observations are not included in the FULL or OSM samples; and the resulting index is standardized to be non-negative, but I impose no upper bound.

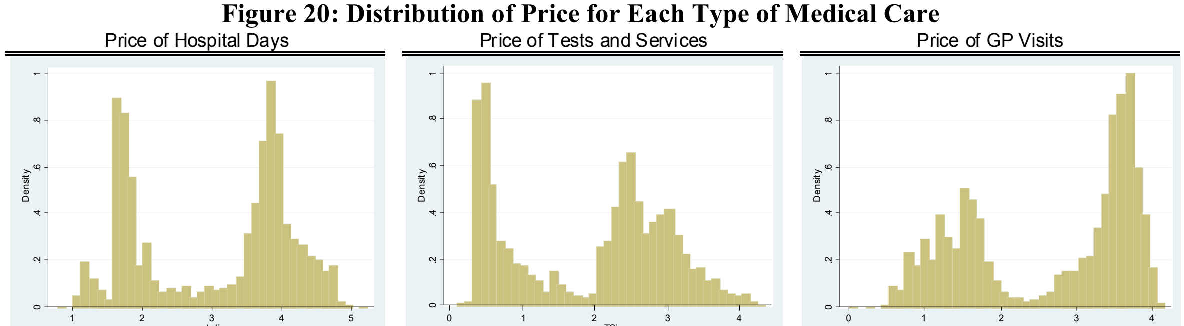
2. MCA Results for the Price of Medical Care

The weights associated with the price instruments appear intuitively reasonable.[148] Appendix D reports weights for all instruments for all waves for each category of medical

[148] The only exception was the weights for the price of tests and services for wave C only were initially all of the opposite sign than the other waves. I determined that wave C had a higher ratio of individuals with both zero wage and zero use of tests and services. This higher mass acted like an “outlier” to pull the MCA coordinates in a different direction. I eliminated 706 observations with both zero wave and zero usage from the MCA analysis for wave C only (these observations are included in the regression estimates). Doing so changed the sign of all of the weights to be consistent with the other waves. However, the resulting average price is higher (3.995 vs. 2.508 average for the remaining waves.) A sensitivity analysis of the NLSUR regression results using only waves D through N did not result in substantially different parameter estimates and did not alter any of the inferences.

care. For all services and all waves the weight for zero wage is negative, the weight for any positive wage is positive and the value of the weight increases monotonically with the wage quartile. This results in a price index that is bimodal: one mode for zero wage and a higher mode for positive wage as illustrated in the figure below. The values of the weights are similar across services and across waves. The standard deviation of the weights across waves is low for all services.

Figure 20: Distribution of Price for Each Type of Medical Care



The weights for different job status categories are also intuitively reasonable. For all services and all waves the weights for self-employed and employed are larger (reflecting a greater opportunity cost of care) than for other job status categories. The weight for employed is slightly larger than the weight for self-employed, perhaps reflecting greater job flexibility among the self-employed to schedule medical care appointment.[149] As expected, the weight for long-term disability is negative and of the highest absolute magnitude of all the categories: those on long-term disability have little if any opportunity cost of care and high cost to their health of not getting care. The magnitude of the weights is similar for hospital days and tests and services and only

---

[149] I initially tried to use the BHPS question about the time of day worked, 'w'jbtime, to reflect the assumption that those who worked during the day had a higher opportunity cost since most medical care would need to be scheduled during work hours. Unfortunately, this question did not include those self-employed and was not asked consistently across the waves.

slightly lower for general practitioner visits. GP visits typically require less time and are easier to schedule than hospitalizations or tests.

For hospital days and tests and services the weight for all NHS is negative and the weights for no service and all private are positive. The weight for “both” is positive for all waves of the tests and services price, but negative for waves E, I, J and K for hospital days. For tests and services the magnitude of the weight for “both” is always higher than that for all private; but this is not always the case for hospital days. The difference in weights for “both” between the two types of care may reflect the different “top up” rules of the NHS. It is easier to go outside the NHS for tests and services than for hospitalizations because once an individual leaves the NHS she cannot return for that “episode” of care. Also, individuals who only occasionally go outside the NHS compared to those who always use private services may choose to go outside for extraordinary circumstances that tend to be more expensive. Alternatively, the results may reflect reporting errors and/or errors in constructing the NHS index for tests and services.

The weight associated with having insurance is positive and of a greater absolute magnitude than the weight associated with no insurance. Again, the magnitudes of the weights are consistent across waves and between hospital days and tests and services (general practitioner visits does not include insurance as an instrument.) Since using the NHS is free with or without insurance, having insurance merely reduces the price of a private provider. The reduced price is still higher than the free NHS. In addition, insurance allows individuals in the UK to access care more quickly in the private provider market than from the NHS. The reduction in waiting time indicates a higher opportunity cost of time among the insured (either for economic or health reasons) that may be reflected in the higher positive weight for insurance.

Finally, the weights for region exhibit variation across regions as expected. The weights for London (inner and outer) are consistently positive across all services and all waves. This suggests that the region weights reflect cost-of -living rather than travel distance since London is presumed to be the most densely populated region with the greatest concentration of medical care providers. The only other region with consistently positive weights is the rest of South East. The regions with consistently negative weights across all waves and services are: West Midlands[150], Merseyside, Tyne & Wear, Wales and Northern Ireland (which did not join the survey until wave G). These are generally poorer regions. The other regions have some variation in the sign of their weights across the waves, but not among the services. Except for the rest of Yorks and Humber the general trend is for the weights to become more positive over time. Notably, this is the only perceived time trend among all the weights.

There is no apparent time trend in the mean computed prices, standard deviations or maximum computed prices across the waves for any of the services. Summary statistics for the computed price are in the figure below. This table shows the average *across waves* of the mean, maximum and standard deviation of computed prices. The minimum price for all services is constrained to zero.

Figure 21: Summary Price Statistics

Summary Price Statistics

	Hospital days	Tests & Services	GP Visits
Mean	2.975	2.508	2.841
Maximum	5.280	4.804	4.552
St. dev.	1.228	1.001	1.212

[150] Except for wave B for general practitioner visits only.

It is important to remember that these are arbitrary, not monetary units. The average "price" for hospital days is slightly higher than the other services, and the maximum price for hospital days is more significantly higher. It is surprising that the standard deviation for tests and services is the lowest among the three services given the higher price computed for wave C.

IV. Estimation Strategy

There are four major econometric challenges: joint estimation of demands, consistency with the economic restrictions, non-negative and highly skewed discrete dependent variables and unobservable heterogeneity. The primary estimation strategy is a pooled Poisson non-linear seemingly-unrelated system of demands with standard errors clustered by individual.[151] This strategy favors the ability impose the economic restrictions on a simultaneous model over the greater flexibility in modeling over dispersion and unobservable heterogeneity in single equation panel data models. This emphasis is consistent with the goal of the paper to test the theory rather than estimate parameters for forecasting.

The specification in equation (22) explicitly models the change in health and relationship between health and consumption that remain as "unobservable heterogeneity" in single period, single equation models. However, including the change in health and estimating a system of demands makes it more difficult to consistently model additional unobservable heterogeneity. Since health is only weakly exogenous a mean-differenced fixed effects estimator would be inconsistent. The demand system makes it difficult to use a first-differenced strategy because of the initial conditions

[151] All estimation is done in STATA 10.1, and the code is available upon request.

problem. Strategies to address the initial conditions problem for dynamic models with discrete dependent variables have been developed for single-equation models by Wooldridge (2005) extending Heckman (1981). The strategy to model the distribution of the heterogeneity conditional on the initial condition would need to be extended to model the joint distribution for the system of equations. This is significantly more complicated even if we make the simplifying assumption that the heterogeneity is distributed identically for each equation. In addition, it is difficult to avoid the incidental parameters problem in a count data model with over-dispersion as discussed in the empirical literature review. Extant work on fitting the overdisperson in count data models has been on single equation models; it is unclear how single equation negative binomial and finite mixture models generalize to systems of demands with varying dispersion parameters. Future research will work to overcome these significant econometric hurdles.

In the present work, I check for robustness of the results to these model misspecifications with single equation models for each demand. I estimate a tobit and random effects tobit for the continuous non-negative consumption expenditures. I estimate a random effects tobit because there is no sufficient statistic for a fixed effects tobit model. I estimate both pooled and fixed effects negative binomial models for hospital days. However, I do this skeptically in light of the incidental parameters problems discussed above. I estimate a pooled negative binomial for tests and services; a fixed effects negative binomial failed to converge. Finally, I estimate an ordered logit and random effects probit for the categorical general practitioner visits. I attempted to estimate a RUM Mixed Logit with random coefficients for the health state variables as described in Train (2003) using the STATA "mixlogit" command. However, this model also failed to converge.

Within the systems framework I check for robustness of the results to two additional econometric concerns. First, the demand for general practitioners is a categorical variable yet in the main specification I treat it as a Poisson distributed numerical variable. I also estimate a three-equation system without the categorical GP demands but retaining the price of GP visits as a regressor in the remaining three equations. Second, even though the primary specification is a pooled panel, the full sample is nonetheless unbalanced. I estimate the multiple and single equation models with both an unbalanced FULL and balanced OSM sample.

V. Empirical Results

A. Tests of Economic Restrictions

Parameter estimates for all of the models are presented in the Appendix F. In order to interpret the parameter estimates as reflecting a utility function envisioned by the theory, it is first necessary to check the economic restrictions. The unrestricted models are inconsistent with these restrictions. In each model the own-price coefficient for tests and services is positive rather than negative and the cross-price symmetry restrictions do not hold as indicated by the likelihood ratio tests in the figure below.

Figure 22: Liklihood Ratio Tests of Symmetry Restrictions

Model	Sample	LL restricted	LL unrestricted	LR	χ^2	Conclusion
4equ	FULL	-1116481	-1110660	11642	12.59	Reject
4 equ	OSM	-371287	-369160	4253	12.59	Reject
3 equ	FULL	-965752	-961665	8174	7.81	Reject
3 equ	OSM	-320400	-318936	2928	7.81	Reject

Nonetheless, when the symmetry restrictions are imposed all of the own-price coefficients are negative consistent with a negative semi-definite Slutsky matrix. Therefore, estimates from models with the symmetry restrictions imposed can be interpreted as consistent with utility maximization and thereby used to test the hypotheses associated with the specified theory.

Phlips (1990) notes that

> …there is no reason why measured behavior should obey (the theoretical restrictions) as theory is always a simplification of reality…All we can hope is that rough estimates, computed without imposing these constraints, will not be inconsistent with them.[152]

That the unconstrained estimates are inconsistent with the theory in itself raises interesting questions about observed behavior that are beyond the scope of this paper. I will discuss this further in the conclusion with respect to future research.

B. Parameter Estimates and Hypothesis Tests

The critical point for the present study is that the main coefficients of interest on the change in health and lagged health are robust to imposing the economic restrictions on the cross-price parameters. In all of the models the signs are the same and the values are quite close, in many cases identical to two decimal places. This indicates that inferences on these parameters are robust to the underlying inconsistency with the theory.

The most important finding is that across all the different specifications the coefficients on the change in health in the demand for medical care equations are negative and highly significant.[153] This suggests that the greater the decline in health the

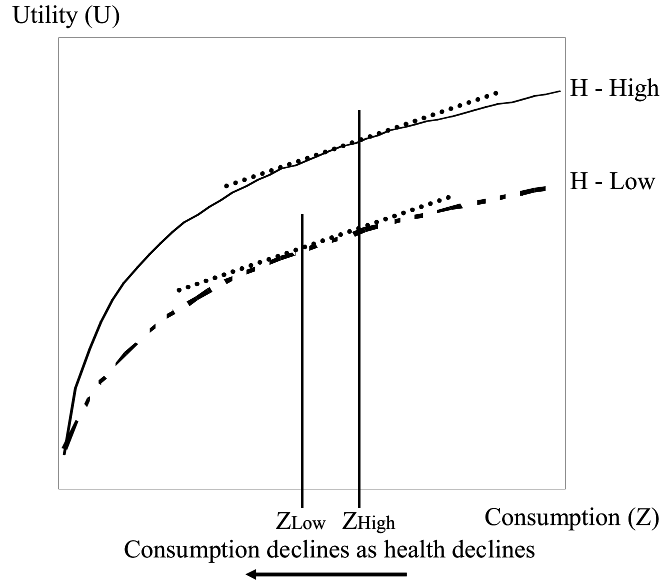
[152] Phlips (1990) p. 53 – 54.
[153] Unless otherwise noted, significance is measured at the 1% alpha.

greater the demand for medical care even after controlling for the lagged state of health. This relationship is significant for all three medical care demands for all of the system specifications. However, since the systems are estimated with a Poisson functional form, the large overdispersion, particularly for the hospital days equation, likely deflates the standard errors even with clustering and robust estimation (Cameron and Trivedi (2005)). The single-equation models better account for such overdispersion and similarly show highly significant negative coefficients for the change in health and lagged health for all of the medical care demands. Therefore, the conclusion that the change in health is significant does not appear to be an artifact of mis-specifying the underlying distribution.

While the coefficients on the change in health and lagged health are negative for the medical care demands, they are positive in the demand for consumption. Again, this finding is robust to imposing the economic restrictions on all the multiple equation models and also consistent with the single equation specifications. Positive coefficients on the change in health and lagged health in the demand for consumption equation suggest two things.[154] First, the statistical significance of these parameters indicates that the demands for health and consumption are not independent as is assumed in pure investment models of the demand for health and single equation models of the demand for medical care. Second, the positive signs of the coefficients on health in the demand for consumption equation suggest that health and consumption are complements. All else equal, when health declines, consumption declines and the higher the level of health the higher the demand for consumption. Holding price constant, this suggests that a decline in health makes the utility of consumption flatter as illustrated in the figure below.

[154] Both coefficients being of the same sign is consistent with the theoretical expectations as discussed in Part I.

Figure 23: Decline in Health, Decline in Marginal Utility of Consumption



When health declines individuals reduce the demand for consumption to maintain the equality between the marginal benefits from consumption and the marginal cost reflected by the constant price. This suggests that the sign of the second derivative of utility with respect to health and consumption is positive: $U_{ZH} > 0$. I will explore the implications of this finding in the next section.

The estimated relationships between the demands for consumption and medical care in general support the theoretical implication from the time paths that these demands are not independent. The cross-price coefficients for consumption and hospital days and consumption and tests and services are statistically significant in most of the specifications; however, the sign of these coefficients is not consistent. The sign often flips from negative to positive when the symmetry restrictions are imposed. The cross-price relationship between general practitioner visits and hospital days is only significant in the four-equation restricted models. Similarly, the relationship between general practitioner visits and consumption is only significant in the full sample unrestricted models. However, the relationship between general practitioner visits and tests and services is always significant and positive except for the unrestricted tests and services equations. The signs for these cross-price parameters are statistically significant, but also inconsistent in sign in the single equation models. Given this inconsistency, the only conclusion is that these demands are not independent; to determine whether they are complements or substitutes requires a more precise measurement of price and greater attention to the model specification issues.

Finally, the cross-equation correlations in the four and three equation specifications are reported in Appendix F. The estimated cross-equation correlations between the demands for consumption and medical care range from approximately .23 for consumption and hospital days to .06 for consumption and tests and services with GP in between. I would expect a larger cross-equation correlation between consumption and medical care if the consumption variable had included more discretionary consumption items. As expected, there is a much greater positive correlation among the medical care demands than between the demand for consumption and medical care. The largest correlation is between hospital days and tests and services and the smallest between hospital days and general practitioner visits with the correlation between tests and services and GP visits in between. The relatively lower correlations with GP visits is inconsistent with a "gatekeeper" model that suggests general practitioners control access to hospitals and specialists. However, the low correlations may also be due to the categorical nature of the GP variable, measurement error in the price parameter, and lack of data on discretionary consumption as discussed above.

VI. Implications and Discussion

The finding that the change in health is statistically significant and negative in the demand for medical care equations supports the theory that the greater the decline in health the greater the demand for medical care. It also offers support for Grossman's derived demand hypothesis. I offer an explanation for why prior empirical tests of this hypothesis found a negative relationship between contemporaneous health and the demand for medical care. Single-period cross-section estimates of the demand for medical care confound the effects of the change in health and the state of health. Moreover, single-period estimates risk bias by omitting the change in health, thereby reflecting an inherently dynamic demand process.

The finding that the change in health and the state of health are statistically significant and positive in the demand for consumption has implications for modeling and interpreting the health-wealth gradient and the time paths of health and consumption. Estimates of the statistical value of a life year typically assume that health and consumption are complements and rise and decline together over the lifecycle. For example, Murphy and Topel (2006) assume the complementarity of health and consumption in their model and use this relationship to calibrate the time path of health based on the observed time path of consumption.[155] As a result, the rate of change in health is high between ages 50 and 70 and asymptotically declines at the end of life. This drives their conclusion that the value of a statistical life year peaks at age 50. However, neither is consistent with the observation that individuals remain healthier longer, experience steeper declines in their health at much more advanced ages and appear willing to pay very high amounts on medical care at the end of life. More importantly,

[155] See Murphy and Topel (2006) p. 877 for the model and p. 887 for the time path of health.

linking the time paths of health and consumption together a priori does not illuminate the mechanism that ties these two time paths together. This underlying mechanism may be essential to understanding the impacts of various policies, particularly in light of the finding that the demands for consumption and medical care are not separable.

The time path of consumption combined with the empirical findings that the cross partials are positive (health and consumption are complements) suggest a mechanism whereby the time paths for health and consumption can diverge due to the interaction with the time path of medical care demand. Referring to equation (21), if the cross-partials between consumption and health and consumption and the change in health are positive as suggested by the empirical results, then the first term of the time path of consumption is consistent with Murphy and Topel (2006) in suggesting that as health declines consumption declines. However, the findings also suggest that the decline in health is associated with an increase in the demand for medical care. A positive $\dot{m}$ in the second term would mitigate the decline in consumption over time. In other words, there is both a direct (negative) and indirect (positive) relationship between the time paths of health and consumption. While we observe the two as complements, their rates of change are also impacted by the time path of medical care demand. If the time paths for health and consumption diverge such that health falls at a slower rate earlier in life and a faster rate later, then this would change the calculations of the value of a statistical life year in Murphy and Topel (2006.) Looking at their equation (9) p. 879, the time path for the value of a life year is directly proportional to the time path for $\dot{H}$, so if $\dot{H}$ falls at a slower rate, so does the value of a life year. Moreover, their equation (9) also suggests that when consumption is greater than income, which their figure 2b indicates happens between ages 65 and 80, a decline in health would also have an indirect positive impact on the value of a life year. Both observations imply a higher level of investment in life later in life consistent with observations of significant investment at older ages and among those in poorer health.

Returning to the time path for medical care demand in (18) offers another perspective on the relationship between medical care and consumption. Again, given empirical findings consistent with $U_{Z\alpha} > 0$, an increase in consumption which we observe earlier in the lifecycle would be consistent with an increase in the demand for medical care among those who may be younger and in relatively good health. A significant increase in the demand for medical care may occur if consumption increases significantly (e.g. job promotion, inheritance wealth transfers) even if health does not decline. However, an increase in both consumption and medical care could violate a binding budget constraint, particularly earlier in the lifecycle in the absence of significant savings. Yet, it may be consistent with the theory of Nyman (2003) who suggests that individuals demand health insurance not for certainty, but because they demand a wealth transfer in sick states. Such a wealth transfer would support increasing both consumption and medical care. The empirical results presented here combined with the underlying theory suggest that the demand for wealth transfer may be associated with particularly large declines in health at younger ages and/or in the absence of savings.

Finally, jointly estimating the demands for consumption and medical care offers an alternative way to model the health/wealth gradient. There is a substantial literature on the relationship between health and wealth and investigation into the inequality of medical care use associated with socio-economic status.[156] Most estimate single-equation models for health (typically with self-reported health as a dependent variable) or medical care. A notable exception is Mangalore (2006) who estimates a three-period three equation model for income, heath and health care. However by estimating income rather than consumption this model does not capture the trade-off between purchasing medical care and purchasing other goods that may contribute to health (e.g. food, shelter, exercise and leisure). By relaxing the separability assumption between health and wealth and estimating consumption and medical care jointly over a long time period allows the data to reflect both the effect of health on consumption and the effect of consumption on future health.

Conclusion

The dissertation makes several contributions to the theoretical and empirical literature on the demand for medical care. First, it offers an alternative explanation for why individuals spend so much on medical care when survival and/or longevity prospects are poor. Second, it illustrates that the restrictive multiplicative functional form for depreciating assets drives many of the key implications of extant theories. Third, it offers several testable hypotheses about the demand for medical care. Specifically, the model suggests that:

1. Individuals will demand more medical care the greater their decline in health at any level of health.

[156] For recent empirical work see Jones and Wildman (2008) and Mangalore (2006) and references therein for reviews and additional literature in this area. See Deaton (2002) for an analysis of the policy implications of the health-wealth gradient.

2. The greater the change in health, the greater the increase in the demand for medical care over time.

3. An increase in the marginal productivity of medical care is associated with an increase in the demand for medical care.

4. The negative effect of an increase in price on the demand for medical care decreases over an individual's lifecycle.

5. The value of longevity (life extension) decreases at the end of life; and the lower the interest rate the earlier in the lifecycle the value of longevity begins to decline.

6. The relative values of longevity and quality of life towards the end of life are a function of all the model parameters including the initial values for health and wealth.

In addition, the theoretical model assumes that the utility of health and wealth are not separable. If the change in health is relevant to utility and not independent of consumption, then the model implies that the demands for medical care and consumption are not separable.

The dissertation contributes to the empirical literature by jointly estimating demands for medical care and other consumption and imposing economic restrictions such that the results can be interpreted as consistent with the underlying theory. In addition, the dissertation illustrates the use of a multiple correspondence analysis to identify instruments for the unobservable health state and price of medical care. The resulting health index captures more dimensions of health that may be relevant to the demand for medical care than the most commonly used indicator of self-assessed health. Empirical findings support the hypothesis that the change in health is relevant and the assumption that health and consumption and thereby the demands for medical care and consumption are not separable. The dissertation offers empirical evidence that health and consumption are complements and extends this finding to a broad age spectrum.

The contributions of the dissertation have several policy implications. First, empirical findings suggest that single equation and single period estimates of the demand for medical care omit significant variables that capture individuals dynamic decision-making process and trade-offs individuals make between medical care and consumption. Therefore, price-elasticity estimates from such single period, single equation estimates are may be biased. Furthermore, the theoretical implication that the effect of price declines over the lifecycle suggests that the often cited -.2 price elasticity estimate from the RHIE (which did not include any individuals over 65) may not hold for those at older ages who account for a high proportion of top spenders. This suggests that policies that aim to reduce medical care spending by increasing consumers' point of purchase price (either through medical savings accounts or less generous insurance coverage) may have less effect on over-all health care spending that currently calculated by applying a -.2 price elasticity to the entire spending distribution.

Second, the finding of a significant relationship between the demands for medical care and other consumption combined with the finding that health and wealth are complements has implications for the value of a life year. The decline in health has a direct negative effect on consumption, but an indirect positive effect associated with the increase in medical care demand. If the value of a life year is based on the utility of consumption as in Murphy and Topel (2006), then the dissertation's findings suggests a higher value longer into the life-cycle which in turn supports a higher level of investment in the health of individuals at advanced ages.

Finally, the dissertation provides a rich foundation for future research. First, the deterministic model can be extended to reflect several sources of uncertainty including uncertain health shocks and uncertain productivity of medical care. While this has been done with the Grossman model (Cropper (1977), Dardanoni and Wagstaff (1990), Ehrlich (2000)), it has not been done with a model that includes the change in health and a more general functional form for health transition. Since these changes have made significant changes to the implications of the Grossman model, they may make similarly significant changes to stochastic specifications as well.

A second theoretical area left unexplored in the dissertation is the comparative dynamics of the model. Such comparative dynamics can offer further testable hypotheses and implications for policy. In particular, comparative dynamics may offer insights about the relationship between the model's parameters and the choice of terminal time (longevity) that are inaccessible from the equilibrium demand for health or the time path of medical care demand. However, the methodology for performing analytic comparative dynamic analysis on a two-state optimal control model is not settled. Existing literature uses the methodology of Oniki (1973) and Frisch decision functions (Ried (1989)); both have limitations for the current model.[157] Caputo (2005) offers a dynamic envelope method that easily handles multiple states without restricting the form of the objective function. However, Caputo develops the method for models with fixed terminal time and time-invariant parameters. I would like to extend the dynamic

[157] Oniki's method for comparative dynamics is based on path analysis, and it is questionable whether this method can be extended to a problem with two state dimensions. Moreover, Oniki requires analysis from a zero parameter value, which is not applicable to price, health, depreciation or other parameters of interest. While Ehrlich and Chuma (1990) reference Oniki, they have provided no derivation of their comparative dynamics to support their conclusions. Frisch decision functions used by Reid (1989) require strict separability of health and consumption in utility, which would not be appropriate for the present model.

envelope method for endogenous terminal time and time-varying parameters to fit the present model and address may interesting questions about the effect of time-varying prices, wages and health shocks on longevity.

The dissertation leaves most of the testable implications of the model untested. Testing these requires different datasets more closely aligned with the different research questions. For example, testing the hypothesis that advances in medical technology are associated with increased demand for medical care may be possible by comparing the demand of individuals with different medical conditions associated with different rates of technological advancement. There have been tremendous advances in some form of cancers over others, and more advancement in the treatment of heart disease over the latter part of the 20th century than practically any other major ailment. Estimation of a parameter for this component of medical care demand may help not only to explain the cross-sectional variation in demand but also to quantify the contribution of medical technology to the increase in medical care spending.[158] Testing the hypothesis regarding quantity vs. quality of life may be possible with a dataset with more nuanced information about medical care choices among individuals at the end of their lives.

The dissertation raises several contemporary econometric issues. Extant literature has not yet extended the discussion of the incidental parameters problem for count data models with unobservable heterogeneity to multiple equation specifications. Moreover, the weak exogeneity of the health and wealth states raises the additional issue of initial

[158] There is a growing literature that attempts attribute a share of medical care cost inflation to the advancement of medical technology. See for example Weisbrod (1991) and Newhouse (1992). However, this is a particularly difficult econometric issue because of the difficulty of estimating the counterfactual growth rate without the new technology. Structural estimation derived from the theory that specifically parameterizes such technological advancement may be able to estimate the primatives associated with preferences for technology (see in particular Gilleskie (2008)).

conditions for such multiple equation models. Addressing the overdispersion of medical care counts and accounting for unobservable heterogeneity in a consistent way is essential to fit the data and use the model for forecasting.

Fitting the top tail of the spending distribution is critical for several avenues of future research. Kean (2005) argues persuasively that it is difficult to have confidence in a model unless it is shown to fit historical data to some subjective degree of reasonableness. A reasonable fit combined with the theoretical explanation of high spending would support using the model as a baseline measure of individual medical care demand. Other sources of demand including physician and insurance induced demand (moral hazard) could then be quantified relative to such a baseline similar to the way so-called "abnormal returns" are quantified relative to the returns predicted by the Capital Asset Pricing Model (CAPM).[159]

Finally, I conclude by returning to Arrow's (1968) observation that the "special characteristics of the medical care market" is due to the "non-marketability of the bearing of suitable risks." The greatest financial risk is in the tail of the spending distribution where individual demand is most likely to outstrip available resources. A model that can forecast the tail of the spending distribution and identify the levers of tail demand could underlie efforts to design new financial products to manage this risk. Managing this risk with tradable products on a secondary market rather than with the law of large numbers in large diversified risk pools can open up new options for the organization of medical care that may overcome some of the non-optimality that Arrow observed forty years ago and that we continue to experience today.

[159] This is done despite the untenable assumptions that underlie the CAPM (quadratic utility or multivariate normality of asset returns) and the significant literature that fails to find empirical support for the CAPM predictions.

Appendix A: List of Symbols

The following is notation for the theoretical model
U = Utility function; LU stands for lifetime utility.
Z = Consumption commodities (all consumption other than medical care, including leisure time and non-medical health consumption such as diet and exercise.)
H = State of health.
m = Medical care including pharmaceuticals, procedures (e.g. heart surgery, X-rays) and physician visits but not including un-prescribed nutrition, exercise, supplements.
r = Constant real rate of interest.
α = A deterministic function that maps time, units of medical care, the state of health and health depreciation to units of health change.
δ = A deterministic function mapping time to the amount of health change.
R = Full wealth.
w = A deterministic function that maps the state of health and time to income including both wages and transfer payments (e.g. government disability, health insurance.)
P = Price per unit of medical care as a function of time.
T = Terminal time (time of death.)
T_{max} = maximum biological lifespan
H_{min} = Minimum health stock necessary to sustain life.

Following is additional notation used in the empirical specification:
Z = Consumption
m_j = Medical care $j \in \{1,2,3\}$ where 1 =hospital days (HD); 2 = general practitioner visits (GP); and 3 = other health tests and services (TS).
p_j^m = *unobservable* price of medical care for type of care j
p_i^Z = price of consumption proxied by the British CPI for all goods and services.
$NI_t - S_t$ = Period t budget constraint of net income (including transfers, after taxes) minus (plus) savings (debt).
H_t^* = *Unobserved* state of health.
couple = married or living as a couple. Single, whether never married, divorsed or widowed is the omitted group.
edu = highest categories of attained education. Lower categories are the omitted group.

Appendix B: Derivation of the Equilibrium Condition and Time Paths

Equilibrium Demand for Health

Following Ehrlich and Chuma (1990), the current stock equilibrium is derived from the necessary condition (6) as follows:

$$\dot{\lambda}^H = -\frac{\partial V}{\partial H} = -\left[\left(e^{-rt}U_H + U_\alpha \alpha_H\right) + \lambda^H \alpha_H + \lambda^R w_H\right]$$

$$\frac{\dot{\lambda}^H}{\lambda^R} = -\frac{e^{-rt}\left(U_H + U_\alpha \alpha_H\right)}{\lambda^R} - \frac{\lambda^H \alpha_H}{\lambda^R} - w_H$$

$$\dot{g} + \frac{\lambda^H \dot{\lambda}^R}{\lambda^{R2}} = -\frac{e^{-rt}\left(U_H + U_\alpha \alpha_H\right)}{\lambda^R} - g(t)\alpha_H - w_H; \dot{\lambda}^R = -r\lambda^R(t); \lambda^R(t) = e^{-rt}\lambda^R(0)$$

$$\dot{g} - rg(t) + g(t)\alpha_H = -\frac{e^{-rt}\left(U_H + U_\alpha \alpha_H\right)}{e^{-rt}\lambda^R(0)} - w_H$$

$$g(t)\left[r - \alpha_H - \tilde{g}(t)\right] = \frac{\left(U_H + U_\alpha \alpha_H\right)}{\lambda^R(0)} + w_H$$

Where the third line substitutes from the time derivative: $\dot{g} = \frac{\dot{\lambda}^H}{\lambda^R} - \frac{\lambda^H \dot{\lambda}^R}{\lambda^{R2}}$, the fourth line substitutes from the necessary conditions for $\dot{\lambda}^R$ as shown, and the second to last line is multiplied by -1 before factoring g(t) in the last line.

Time Path of $\dot{m}$ *:*

Necessary condition (8)

$$\frac{\partial V}{\partial m} = e^{-rt}U_\alpha \alpha_m + \lambda^H \alpha_m - \lambda^R P(t) = 0 \tag{23}$$

Time derivative of necessary condition (8):

$$\begin{aligned} \dot{V}_m &= -re^{-rt}U_\alpha \alpha_m + e^{-rt}\left[\dot{U}_\alpha \alpha_m + U_\alpha \dot{\alpha}_m\right] \\ &\quad + \dot{\lambda}^H \alpha_m + \lambda^H \dot{\alpha}_m - \dot{\lambda}^R P(t) - \lambda^R \dot{P} \end{aligned} \tag{24}$$

Expand time derivatives of U, and α functions as follows:

$$\begin{aligned} \dot{U}_\alpha &= U_{\alpha Z}\dot{Z} + U_{\alpha H}\dot{H} + U_{\alpha\alpha}\dot{\alpha} \\ &= U_{\alpha Z}\dot{Z} + U_{\alpha H}\dot{H} + U_{\alpha\alpha}\left[\alpha_t + \alpha_m \dot{m} + \alpha_H \dot{H} + \alpha_\delta \dot{\delta}\right] \end{aligned} \tag{25}$$

Where the second line expands $\dot{\alpha}$. And:

$$\dot{\alpha}_m = \alpha_{mt} + \alpha_{mm}\dot{m} + \alpha_{mH}\dot{H} + \alpha_{m\delta}\dot{\delta} \tag{26}$$

Substitute (27) and (28) into (29):

$$\begin{aligned} \dot{V}_m &= -re^{-rt}U_\alpha \alpha_m + e^{-rt}\begin{bmatrix} \left(U_{\alpha Z}\dot{Z} + U_{\alpha H}\dot{H} + U_{\alpha\alpha}\left[\alpha_t + \alpha_m \dot{m} + \alpha_H \dot{H} + \alpha_\delta \dot{\delta}\right]\right)\alpha_m \\ + U_\alpha\left(\alpha_{mt} + \alpha_{mm}\dot{m} + \alpha_{mH}\dot{H} + \alpha_{m\delta}\dot{\delta}\right) \end{bmatrix} \\ &\quad + \dot{\lambda}^H \alpha_m + \lambda^H\left(\alpha_{mt} + \alpha_{mm}\dot{m} + \alpha_{mH}\dot{H} + \alpha_{m\delta}\dot{\delta}\right) - \dot{\lambda}^R P(t) - \lambda^R \dot{P} \end{aligned} \tag{30}$$

Move $\dot{m}$ terms to lhs and factor $\dot{m}$. There are two sets of terms: one associated with the utility function and the other associated with the health transition constraint.

$$-\dot{m}\left[e^{-rt}\left(U_{\alpha\alpha}\alpha_m^2+U_\alpha\alpha_{mm}\right)+\lambda^H\alpha_{mm}\right]=-re^{-rt}U_\alpha\alpha_m$$
$$+e^{-rt}\begin{bmatrix}\left(U_{\alpha Z}\dot{Z}+U_{\alpha H}\dot{H}+U_{\alpha\alpha}\left[\alpha_t+\alpha_H\dot{H}+\alpha_\delta\dot{\delta}\right]\right)\alpha_m\\+U_\alpha\left(\alpha_{mt}+\alpha_{mH}\dot{H}+\alpha_{m\delta}\dot{\delta}\right)\end{bmatrix} \tag{31}$$
$$+\dot{\lambda}^H\alpha_m+\lambda^H\left(\alpha_{mt}+\alpha_{mH}\dot{H}+\alpha_{m\delta}\dot{\delta}\right)-\dot{\lambda}^R P(t)-\lambda^R\dot{P}$$

Substitute for $\dot{\lambda}^H$ and $\dot{\lambda}^R$ from the necessary conditions:

$$-\dot{m}\left[e^{-rt}\left(U_{\alpha\alpha}\alpha_m^2+U_\alpha\alpha_{mm}\right)+\lambda_H\alpha_{mm}\right]=-re^{-rt}U_\alpha\alpha_m$$
$$+e^{-rt}\begin{bmatrix}\left(U_{\alpha Z}\dot{Z}+U_{\alpha H}\dot{H}+U_{\alpha\alpha}\left[\alpha_t+\alpha_H\dot{H}+\alpha_\delta\dot{\delta}\right]\right)\alpha_m\\+U_\alpha\left(\alpha_{mt}+\alpha_{mH}\dot{H}+\alpha_{m\delta}\dot{\delta}\right)\end{bmatrix} \tag{32}$$
$$+\left(-\left[e^{-rt}\left[U_H+U_\alpha\alpha_H\right]+\lambda^H\alpha_H+\lambda^R w_H\right]\right)\alpha_m$$
$$+\lambda^H\left(\alpha_{mt}+\alpha_{mH}\dot{H}+\alpha_{m\delta}\dot{\delta}\right)-\left(-\lambda^R r\right)p(t)-\lambda^R\dot{p}$$

Simplify with $D=\left[-e^{-rt}\left(U_{\alpha\alpha}\alpha_m^2+U_\alpha\alpha_{mm}\right)-\lambda^H\alpha_{mm}\right]$ and combine like terms:

$$\dot{m}=\frac{1}{D}\begin{Bmatrix}+e^{-rt}\alpha_m\left(U_{\alpha Z}\dot{Z}+U_{\alpha H}\dot{H}+U_{\alpha\alpha}\left(\alpha_t+\alpha_H\dot{H}+\alpha_\delta\dot{\delta}\right)\right)\\+\left(e^{-rt}U_\alpha+\lambda_H\right)\left(\alpha_{mt}+\alpha_{mH}\dot{H}+\alpha_{m\delta}\dot{\delta}\right)\\-\alpha_m\left[e^{-rt}\left(rU_\alpha+U_H+U_\alpha\alpha_H\right)+\lambda^H\alpha_H+\lambda^R w_H\right]\\+\lambda^R\left(rP(t)-\dot{P}\right)\end{Bmatrix} \tag{33}$$

Time path of $\dot{Z}$ *:*

Time derivative of necessary condition:

$$\dot{V}_Z=-re^{-rt}U_Z+e^{-rt}\left(U_{ZZ}\dot{Z}+U_{ZH}\dot{H}+U_{Z\alpha}\dot{\alpha}\right)-\dot{\lambda}^R=0 \tag{34}$$

Move $\dot{Z}$ term to rhs, expand $\dot{\alpha}$ and substitute for $\dot{\lambda}_R$:

$$-e^{-rt}U_{ZZ}\dot{Z}=-re^{-rt}U_Z+e^{-rt}\left[U_{ZH}\dot{H}+U_{Z\alpha}\left(\alpha_t+\alpha_m\dot{m}+\alpha_H\dot{H}+\alpha_\delta\dot{\delta}\right)\right]-\left(-\lambda^R r\right)=0 \tag{35}$$

Simplify:

$$\begin{aligned}\dot{Z}&=\frac{\left[U_{ZH}\dot{H}+U_{Z\alpha}\left(\alpha_t+\alpha_m\dot{m}+\alpha_H\dot{H}+\alpha_\delta\dot{\delta}\right)\right]}{-U_{ZZ}}+\frac{rU_Z-e^{rt}r\lambda^R}{U_{ZZ}}\\&=\frac{\left[U_{ZH}\dot{H}+U_{Z\alpha}\left(\alpha_t+\alpha_m\dot{m}+\alpha_H\dot{H}+\alpha_\delta\dot{\delta}\right)\right]}{-U_{ZZ}}\end{aligned} \tag{36}$$

Where the last term simplifies by substituting the FOC (9) for λ^R.

Appendix C: Sufficiency Conditions

The following analysis is for the Mangasarian sufficiency conditions, not the more general Arrow sufficiency conditions. Relaxing the assumption of diminishing returns to scale for health production would require a more complex proof of Arrow sufficiency. The Mangasarian conditions require the Hamiltonian to be concave in all state and control variables. Since the sum of a concave function is concave, each term of the Hamiltonian is considered in turn.

Hessian of the utility component: $e^{-rt}U\left[Z(t),H(t),\alpha(t)\right]$

The discount factor is positive and thereby not necessary to determine the signs of the determinants. The wealth state, R, does not appear in the utility function. Therefore, the Hessian for this component is a 3x3 matrix:

$$\begin{vmatrix} U_{ZZ} & U_{ZH} & U_{Zm} \\ U_{HZ} & U_{HH} & U_{Hm} \\ U_{Zm} & U_{mH} & U_{mm} \end{vmatrix}$$

$$\begin{vmatrix} U_{ZZ} & U_{ZH}+U_{Z\alpha}\alpha_H & U_{Z\alpha}\alpha_m \\ U_{HZ}+U_{\alpha Z}\alpha_H & U_{HH}+2U_{H\alpha}\alpha_H+U_\alpha\alpha_{HH} & U_{H\alpha}\alpha_m+U_{\alpha\alpha}\alpha_m\alpha_H+U_\alpha\alpha_{Hm} \\ U_{Z\alpha}\alpha_m & U_{H\alpha}\alpha_m+U_{\alpha\alpha}\alpha_m\alpha_H+U_\alpha\alpha_{Hm} & U_{\alpha\alpha}\alpha_m^2+U_\alpha\alpha_{mm} \end{vmatrix}$$

For the function to be concave the first principal minors (U_{ZZ}, U_{HH}, U_{mm}) must all be negative. U_{ZZ} and U_{HH} have already been assumed negative, and U_{mm} is unambiguously negative by specified assumptions that $U_{\alpha\alpha}<0; U_\alpha>0; \alpha_{mm}<0$.

The determinant of the second principal minor must be positive:

$$\begin{vmatrix} U_{ZZ} & U_{HZ} \\ U_{ZH} & U_{HH} \end{vmatrix}$$

$$U_{ZZ}\left(U_{HH}+2U_{H\alpha}\alpha_H+U_\alpha\alpha_{HH}\right)-\left(U_{ZH}+U_{\alpha Z}\alpha_H\right)^2 \geq 0$$

This holds if the first term, which is positive by the multiplication of two negative second derivatives, is greater or equal to the second term which is positive since squared.

Finally, the determinant of the third principal minor, or the full Hessian, must be negative:

$$\begin{aligned} & U_{ZZ}\left[\left(U_{HH}+2U_{H\alpha}\alpha_H+U_\alpha\alpha_{HH}\right)\left(U_{\alpha\alpha}\alpha_m^2+U_\alpha\alpha_{mm}\right)-\left(U_{H\alpha}\alpha_m+U_{\alpha\alpha}\alpha_m\alpha_H+U_\alpha\alpha_{Hm}\right)^2\right] \\ & -\left(U_{ZH}+U_{Z\alpha}\alpha_H\right)\left[\left(U_{HZ}+U_{\alpha Z}\alpha_H\right)\left(U_{\alpha\alpha}\alpha_m^2+U_\alpha\alpha_{mm}\right)-\left(U_{Z\alpha}\alpha_m\right)\left(U_{H\alpha}\alpha_m+U_{\alpha\alpha}\alpha_m\alpha_H+U_\alpha\alpha_{Hm}\right)\right] \\ & +U_{Z\alpha}\alpha_m\left[\left(U_{HZ}+U_{\alpha Z}\alpha_H\right)\left(U_{H\alpha}\alpha_m+U_{\alpha\alpha}\alpha_m\alpha_H+U_\alpha\alpha_{Hm}\right)-\left(U_{Z\alpha}\alpha_m\right)\left(U_{HH}+2U_{H\alpha}\alpha_H+U_\alpha\alpha_{HH}\right)\right] \\ & \leq 0 \end{aligned}$$

This sufficiency requirement is highly ambiguous and can hold with a number of different relative magnitudes given the stated assumptions on the cross-partial derivatives.

Hessian of the health transition component: $\alpha\left[t, m(t), H(t), \delta(t)\right]$

Since the depreciation function is neither a state nor a control, it is not directly included in the Hessian. The co-state variable λ^{H} must be positive over the planning horizon for a maximization problem and is omitted in the following computation of determinants to simplify notation.

$$\begin{vmatrix} \alpha_{mm} & \alpha_{mH} \\ \alpha_{mH} & \alpha_{HH} \end{vmatrix}$$

The health production function is assumed to be diminishing in returns to both m and H ensuring that the first principal minors are negative.

$$\alpha_{HH}\alpha_{mm} - \left(\alpha_{mH}\right)^{2} \geq 0$$

α_{mH} can take any sign since it is squared; however the magnitude of the squared term must be less than the product of the second derivatives.

Hessian of the wealth transition component: $rR(t) + w(H(t)) - p(t)m(t) - Z(t)$

Again, The co-state variable λ^{R} must be positive over the planning horizon for a maximization problem and so is omitted in the Hessian as it plays no role in determining the signs.

$$\begin{vmatrix} w_{HH} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{vmatrix}$$

The only second derivative associated with a state or control variable is the second derivative of the wealth function with respect to health. It is negative by assumption of diminishing returns (zero in the empirical specification.) All other principal minors are zero, consistent with the sufficiency conditions.

Appendix D: BHPS variables

For more detail on each variable please see
http://www.iser.essex.ac.uk/ulsc/bhps/doc/volb/allterms.php
'w' indicates the wave with wave exclusions as noted. Unless otherwise noted all variables are individual responses available for all waves.

Dependent Variables

'w'hospd	Hospital inpatient days in the past year, numerical from 0 to 365.
'w'hl2gp	Number of times the individual has talked to or visited a general practitioner for own health. 1 = one or two; 2 = three to five; 3 = six to ten; 4 = more than ten.

Variables used to compute index for tests and services (TS) coded 1 = yes; 0 = no.

'w'hlck#	Health checks for # = a through h for waves A through H, 'w'hlcki, blood tests, is added from wave E.
'w'hlsv#	Health services for # = a through i for waves A through D. 'w'hlckl, consultations, and 'w'hlckm, family planning, are added at wave E. I do not include 'w'hlckj or 'w'hlckk listed as "other" because they are more likely to include welfare rather than health services. However, I do include 'w'hlsvc, meals on wheels, and 'w'hlcvd, social worker visits because they are more likely to indicate limitations with daily activities and/or social and emotional impairments associated with health.

Variables used for aggregate consumption

Note: all of these variables are household responses.

'w'xpfood:	Weekly categorical expenditure on food converted to numerical using the maximum of each category and annualized by x 52.
'w'xphsn	Net (of government subsidy) monthly rent or mortgage. Continuous monetary variable with zero value for homes with no mortgage.
'w'xpfuel	Monthly numerical expenditure on oil, gas and electricity annualized by x 12 waves A – E. for eaves G – N fuel expense is an aggregate of: 'w'xpgasy, 'w'xpoily and 'w'xplecy. Annual fuel cost for wave F was linearly interpolated from waves E and G by individual pid since there is no cross-wave household indicator.
'w'cdnuxp	Annual amount spent on consumer durables for the list of products given. Waves A – F
'w'cd#cst	Amount spent on consumer durables for # = 1 through 12 aggregated to 'w'cdcost for waves G – N.

Variables used in the MCA analysis for the health state:

'w'sex	1 = male; 2 = female
'w'age	Age at date of interview.
'w'hlstat	Self-reported health on a scale from 1 = excellent to 5 = very poor. 'w'hlsf1 are used for waves I and N with the same coding.

'w'hlghq1	Likert scale of subjective well-being from 1 to 36 with higher values denoting lower levels of health. I use the 36 point likert scale rather than the 12 point caseness scale to maximize the variation in the data.
'w'nxdts	Number of accidents over the past year categorical from 0 to 4+
'w'hldsbl	Whether the respondent is registered disabled. lhldsbl1, considers oneself disabled, used for waves L and N.
'w'hlprb#	Health problems # = a through m for all waves.
'w'hllt#	Health limitations # = a through e plus wa, limitations with work for all waves except I. All values are 1 = yes and 0 = no except 'w'hlttwa which has four categories from 0, indicating that the prior question about work limitations was answered no, 1 = a lot to 4 = not at all. The following variables are used for health limitations for waves I and N: 'w'hlsf1, 'w'hlsf3a through 'w'hlsf3j, 'w'hlsf4a through 'w'hlsf4d, 'w'hlsf6 to 'w'hlsf8.
'w'mastat	Marital status values 1 – 6 representing: married, living as a couple, separated, divorced, widowed, never married. There are no individuals under 16 in the sample. I use this variable instead of 'w'mlstat because it includes living as a couple.
'w'smoker	Current smoking status: 1 = yes; 2 = no. ismnow used for wave I.

<u>Variables used for the price of medical care</u>

'w'hospnhs	Categorical variable for hospital days with 1 = all NHS, 2 = all private; 3 = both private and NHS.
'w'hlck#n	Health checks for # = a through h for waves A through H, 'w'hlcki, blood tests, is added from wave I. 1 = NHS, 2 = private; 3 = both.
'w'hlsv#n	Health services for # = a through i for waves A through D. 'w'hlckl, consultations, and 'w'hlckm, family planning, are added at wave E. . 1 = NHS, 2 = private; 3 = both. Note: I use the same variables that comprise the TS index and create a variable 'w'TSNHS for whether the services were 1 = all NHS, 2 = all private, and 3 = both. An individual would be coded both if only one services was noted as either NHS or both. I add a category 0 indicating no tests or services reported.
'w'hl2gp	I use the index for GP use and include 0 indicating no GP use.
'w'hlcvr	Categorical variable indicating private insurance coverage. I combine categories 1, yes in own name, and 2, yes in other family members name to 1 = yes with 2 = no private insurance. This variable is only available after wave F. However, the ratio of insured to uninsured is stable in the data at approximately 80% uninsured. I apply the insurance status from wave F to individuals in the data for prior waves only for the purpose of determining cost of service weights. Insurance status is not otherwise modeled and/or used as an independent variable. Insurance status is not used as an instrument for the price of general practitioner visits because most visits are through the NHS.

‘w’jbstat	Categorical variable 1 – 10 with categories: self employed, employed, unemployed, retired, maternity leave, family care, FT student, LT sick or disabled, government training program, other.
‘w’fimnl	Labor income last month. I use the imputed variable that includes self-employment income. I create a categorical variable with 0 and four quartiles greater than 0 called ‘w’wage to use in the MCA analysis.
‘w’region	Categorical from 1 – 18 indicating the different regions of the UK. We use this variable to control for unobservable cost-of-living differences across regions.

Other independent variables

‘w’qfedhi	highest educational attainment. Top five categories: higher degree, first degree, teaching QF, Other Higher QF and Nursing QF are grouped as = 1 in the variable edu. Lower levels of educational attainments are the omitted category.

Appendix E: Instrument Weights for Health State and Price of Medical Care

Health State Weights:

	Weight for each wave															
Instrument	A	B	C	D	E	F	G	H	I	J	K	L	M	N	mean	st. dev.
Sex																
Male	0.226	0.253	0.262	0.266	0.288	0.281	0.264	0.262	0.161	0.243	0.246	0.236	0.24	0.253	0.249	0.030
Female	-0.197	-0.217	-0.223	-0.228	-0.246	-2.41	-0.226	-0.22	-0.155	-0.203	-0.203	-0.198	-0.201	-0.21	-0.367	0.588
Age																
16	1.318	1.918	1.347	1.295	1.138	1.371	1.263	1.332	1.011	1.42	1.486	1.426	1.373	0.943	1.332	0.230
36	0.583	0.693	0.663	0.836	0.533	0.349	0.284	0.476	0.431	0.484	0.668	0.399	0.664	0.694	0.554	0.157
56	-0.369	-0.819	-0.761	-0.606	-0.648	-0.728	-0.264	-0.476	-0.449	-0.984	-0.821	-0.333	-0.346	-0.287	-0.564	0.233
76	-2.137	-3.304	-2.456	-1.529	-1.872	-1.766	-1.983	-2.484	-1.376	-2.463	-2.704	-2.767	-1.779	-1.651	-2.162	0.551
Oldest	-5.162	-20.175	-8.414	-1.7	-7.028	-7.264	-9.08	-0.708	-3.03	-6.23	-1.713	-3.714	-10.79	-5.549	-6.468	4.951
Oldest age	96	97	93	96	95	96	97	97	95	99	96	96	97	98	96.286	
Self-reported health																
excellent	1.093	1.145	1.117	1.109	1.081	1.118	1.132	1.195	1.04	1.252	1.246	1.242	1.234	0.999	1.143	0.080
good	0.563	0.569	0.598	0.621	0.625	0.627	0.665	0.712	0.788	0.736	0.726	0.715	0.705	0.736	0.670	0.070
fair	-0.994	-1.091	-0.073	-1.031	-0.947	-0.959	-0.759	-0.867	0.261	-0.871	-0.928	-0.938	-0.969	0.242	-0.709	0.474
poor	-4.349	-3.831	-4.089	-4.051	-3.794	-3.722	-3.429	-3.232	-1.427	-3.319	-3.68	-3.549	-3.711	-1.457	-3.403	0.883
very poor	-7.322	-7.403	-6.995	-7.095	-7.438	-6.86	-6.478	-6.279	-3.984	-6.051	-5.973	-6.306	-6.241	-3.58	-6.286	1.175
Mental & emotional health																
0	1.299	0.782	0.619	0.604	0.887	0.881	1.391	1.084	1.083	1.003	1.025	1.09	0.902	0.991	0.974	0.222
5	1.006	0.816	0.832	0.922	0.801	0.98	0.916	0.935	0.674	0.917	1.015	0.978	0.893	0.701	0.885	0.107
15	-0.747	-0.374	-0.654	-1.261	-7.42	-0.831	-0.673	-0.477	-0.529	-0.744	-0.779	-0.829	-0.913	-0.68	-1.208	1.800
25	1.905	-2.071	-2.316	-1.478	-2.26	-2.461	-0.995	-3.278	-1.842	-2.302	-3.169	-2.226	-2.646	-1.666	-1.915	1.256
36	-11.535	-6.022	-2.532	-4.716	-3.425	-6.601	-4.405	-4.05	-2.828	-5.067	-4.598	-4.342	-3.218	-2.774	-4.722	2.297
Reported health problems																
arms	-2.242	-2.262	-2.295	-2.252	-2.177	-2.217	-2.093	-1.987	-1.6	-1.965	-2.023	-2.015	-2.029	-1.65	-2.058	0.216
no problem	0.659	0.657	0.73	0.735	0.712	0.764	0.786	0.789	0.627	0.799	0.784	0.802	0.777	0.646	0.733	0.063
sight	-2.325	-2.911	-3.199	-3.375	-3.351	-3.188	-2.995	-3.093	-1.962	-3.191	-2.879	-2.793	-2.909	-1.946	-2.866	0.468
no problem	0.177	0.164	0.149	0.158	0.137	0.139	0.165	0.151	0.101	0.167	0.146	0.153	0.159	0.108	0.148	0.022
hearing	-2.104	-1.975	-2.057	-2.135	-2.174	-2.043	-1.787	-2.007	-1.239	-1.978	-1.965	-1.882	-1.835	-1.417	-1.900	0.268
no problem	0.17	0.162	0.17	0.18	0.172	0.16	0.164	0.184	0.126	0.183	0.175	0.179	0.167	0.131	0.166	0.018
skin/allergy	-0.169	-0.332	-0.216	-0.321	-0.267	-0.292	-0.339	-0.398	-0.195	-0.41	-0.46	-0.336	-0.45	-0.358	-0.325	0.090
no problem	0.02	0.037	0.027	0.038	0.031	0.039	0.047	0.059	0.026	0.056	0.059	0.048	0.062	0.048	0.043	0.014
chest/breathing	-1.992	-2.119	-2.003	-1.858	-1.709	-1.468	-1.743	-1.753	-1.335	-1.83	-1.933	-1.785	-1.82	-1.314	-1.762	0.241
no problem	0.226	0.253	0.256	0.255	0.24	0.212	0.28	0.276	0.23	0.302	0.298	0.285	0.29	0.206	0.258	0.032
heart/blood pressure	-2.364	-2.697	-2.557	-2.6	-2.51	-2.586	-2.288	-2.337	-1.624	-2.272	-2.221	-2.113	-2.129	-1.571	-2.276	0.339
no problem	0.313	0.367	0.36	0.363	0.359	0.382	0.385	0.409	0.314	0.443	0.445	0.447	0.451	0.338	0.384	0.048
stomach/digestion	-2.15	-2.624	-2.77	-2.714	-2.803	-2.495	-2.661	-2.504	-1.748	-2.311	-2.276	-2.124	-1.968	-1.521	-2.334	0.394
no problem	0.125	0.165	0.177	0.179	0.179	0.168	0.219	0.211	0.138	0.2	0.214	0.196	0.173	0.136	0.177	0.030
diabetes	-2.968	-3.199	-3.157	-3.151	-3.002	-2.715	-2.924	-3.066	-1.888	-2.685	-2.681	-2.562	-2.434	-1.731	-2.726	0.455
no problem	0.05	0.061	0.063	0.066	0.062	0.057	0.077	0.082	0.06	0.085	0.095	0.097	0.099	0.073	0.073	0.016
anxiety/depression	-2.368	-2.839	-2.768	-2.788	-2.832	-2.663	-2.557	-2.655	-1.673	-2.585	-2.672	-2.445	-2.416	-1.633	-2.492	0.386
no problem	0.124	0.179	0.169	0.179	0.182	0.185	0.219	0.229	0.143	0.235	0.263	0.227	0.221	0.148	0.193	0.040
alcohol/drugs	-1.307	-3.25	-2.097	-2.317	-1.956	-2.918	-2.937	-2.433	-1.219	-1.73	-2.209	-2.342	-2.134	-1.538	-2.171	0.603
no problem	0.005	0.01	0.008	0.008	0.008	0.011	0.016	0.011	0.007	0.007	0.012	0.013	0.013	0.008	0.010	0.003
epilepsy	-1.277	-1.98	-1.543	-1.283	-1.177	-1.639	-1.393	-1.819	-1.272	-2.17	-2.199	-1.918	-2.024	-1.138	-1.631	0.382
no problem	0.008	0.015	0.012	0.01	0.009	0.012	0.011	0.014	0.01	0.02	0.02	0.015	0.017	0.009	0.013	0.004
migraine	-0.256	-0.718	-0.863	-0.735	-0.85	-0.986	-1.021	-0.795	-0.538	-0.863	-1.112	-0.897	-0.923	-0.762	-0.809	0.213
no problem	0.022	0.066	0.081	0.065	0.073	0.102	0.108	0.082	0.051	0.086	0.098	0.085	0.077	0.063	0.076	0.022
other	-1.889	-2.294	-2.191	-2.254	-2.417	-2.37	-2.45	-2.112	-1.5	-1.713	-3.052	-1.509	-1.67	-1.197	-2.044	0.491
no problem	0.096	0.104	0.085	0.083	0.098	0.117	0.118	0.122	0.088	0.108	0.045	0.061	0.075	0.058	0.090	0.024

(Continued on the next page)

Instrument Weights for Health State Continued

	Weight for each wave															
Instrument	A	B	C	D	E	F	G	H	I	J	K	L	M	N	mean	st. dev.
Reported activity limitations[1]																
housework	-7.898	-7.816	-7.53	-7.051	-7.265	-7.197	-6.222	-6.37	-6.01	-6.14	-5.841	-5.984	-6.112	-6.108	-6.682	0.740
no problem	0.38	0.359	0.374	0.403	0.406	0.415	0.475	0.464	1.71	0.468	0.502	0.474	0.47	1.683	0.613	0.461
climbing stairs	-7.684	-7.709	-7.607	-7.234	-7.184	-7.29	-6.137	-6.509	-7.386	-6.149	-5.895	-6.237	-6.138	-7.725	-6.920	0.699
no problem	0.442	0.41	0.414	0.437	0.438	0.434	0.501	0.471	0.574	0.507	0.53	0.489	0.494	0.75	0.492	0.088
dressing	-10.821	-10.38	-10.404	-9.654	-9.782	-9.172	-8.06	-8.286	-9.053	-7.977	-7.329	-7.497	-7.521	-8.891	-8.916	1.172
no problem	0.134	0.138	0.147	0.161	0.169	0.176	0.2	0.195	0.305	0.194	0.216	0.203	0.218	0.237	0.192	0.045
walking	-7.846	-7.674	-7.373	-7.193	-7.192	-7.22	-6.235	-6.362	-6.705	-6.13	-5.854	-6.087	-6.077	-7.054	-6.786	0.658
no problem	0.428	0.416	0.419	0.445	0.433	0.441	0.486	0.488	0.671	0.495	0.534	0.496	0.499	0.695	0.496	0.087
other	-1.497	-1.615	-1.533	-1.473	-1.441	-1.332	-1.097	-1.153	-1.839	-1.033	-0.953	-1.055	-1.112	-2.083	-1.373	0.331
no problem	0.076	0.09	0.079	0.071	0.073	0.073	0.064	0.07	0.675	0.074	0.072	0.074	0.077	0.974	0.182	0.279
Work limitattions																
no problem	0.722	0.724	0.71	0.723	0.714	0.735	0.766	0.756	0.768	0.783	0.815	0.799	0.787	0.746	0.753	0.034
no limitations	-0.779	-0.745	-0.857	-0.716	-0.925	-0.474	-0.55	-0.383	-0.382	-0.294	-0.479	-0.334	-0.513	-0.399	-0.559	0.207
little	-2.278	-2.444	-2.235	-2.017	-1.906	-1.866	-1.58	-1.57	-1.585	-1.533	-1.439	-1.523	-1.578	-1.973	-1.823	0.327
somewhat	-3.803	-3.663	-3.672	-3.451	-3.239	-3.504	-2.826	-2.947	-2.66	-2.831	-2.65	-2.86	-2.958	-2.984	-3.146	0.399
a lot	-7.552	-7.249	-7.174	-6.777	-6.98	-6.694	-6.02	-5.857	-4.193	-5.74	-5.626	-5.726	-5.702	-3.892	-6.084	1.088
Disabled																
yes	-6.546	-6.729	-6.815	-6.631	-6.732	-6.5	-5.561	-5.887	-3.691	-5.615	-4.493	-4.567	-4.581	-3.266	-5.544	1.217
no	0.259	0.271	0.275	0.273	0.29	0.291	0.328	0.33	0.268	0.359	0.389	0.564	0.558	0.385	0.346	0.101
Accidents																
0	-0.02	0.026	0.03	0.006	0.024	0.027	0.026	0.026	0.024	0.039	0.041	0.02	0.037	0.044	0.025	0.016
1	0.218	-0.097	-0.142	0.01	-0.16	-0.127	-0.076	-0.141	-0.084	-0.125	-0.148	-0.09	-0.279	-0.304	-0.110	0.123
2	0.223	-0.534	-0.628	-0.495	0.069	-0.283	-0.241	-0.391	-0.589	-0.997	-1.055	-0.577	-0.693	-0.97	-0.512	0.374
3	-0.533	-1.373	-3.326	0.611	-1.978	-2.014	-2.857	-1.501	-1.294	-3.522	-2.725	-1.681	-0.786	-1.959	-1.781	1.122
4+	-3.155	-5.859	-4.909	2.54	-3.128	-4.745	-4.632	-2.988	-1.21	-2.56	-4.228	-1.259	-2.555	-2.837	-2.966	2.083
Marital status																
married	0.059	0.038	0.015	-0.027	-0.013	-0.051	-0.039	-0.006	-0.008	-0.025	-0.058	-0.033	-0.035	-0.007	-0.014	0.033
living as couple	0.844	0.771	0.78	0.775	0.813	0.734	0.761	0.752	0.671	0.811	0.782	0.778	0.72	0.49	0.749	0.086
widowed	-2.604	-2.592	-2.501	-2.494	-2.668	-2.66	-2.404	-2.478	-1.772	-2.287	-2.207	-2.39	-2.523	-1.997	-2.398	0.258
divorced	-0.74	-0.735	-0.8	-0.694	-0.67	-0.804	-0.777	-0.816	-0.585	-1.098	-1.112	-1.022	-0.961	-0.679	-0.821	0.165
separated	-0.31	-0.228	-0.55	-0.132	-0.014	-0.058	-0.062	-0.242	-0.192	-0.372	-0.505	-0.535	-0.142	-0.061	-0.243	0.185
never married	0.776	0.81	0.811	0.863	0.797	0.864	0.825	0.795	0.619	0.822	0.853	0.808	0.802	0.598	0.789	0.081
Smoker																
yes	-0.036	-0.065	-0.094	-0.103	-0.113	-0.087	-0.156	-0.138	0.033	-0.141	-0.155	-0.134	-0.124	-0.052	-0.098	0.053
no	0.015	0.026	0.035	0.093	0.044	0.034	0.065	0.056	-0.027	0.056	0.061	0.05	0.044	0.018	0.041	0.028
Health average[2]	89.817	88.427	85.363	86.105	85.625	85.827	84.941	82.869	81.011	83.982	82.613	80.057	80.828	77.322	83.913	3.384
Health standard deviation	9.553	11.253	13.363	13.816	13.532	12.905	15.280	15.775	18.813	14.582	15.613	15.338	15.195	20.169	14.656	2.708
number of observations[3]	10264	9845	9600	9481	9249	9438	11193	10906	15623	15606	18867	16597	16238	15791		

1. Health limitation questions are changed and disaggregated for waves I and N. The full list of questions were used in the MCA analysis, and results are reported separately.
For the purpose of this table the following comparisons are use Housework = "a lot" from vigorous activity + bending and kneeling
Climbing stairs ="a lot" + "a little" from climbing one flight of stairs.If individuals have a lot of trouble climbing one flight they will have a lot of trouble climbing several flights
Walking = "a lot" + "a little" from walking 1/2 Other = "a little" from moderate activity
2. The Health state is predicted using row scores based on the indicator matrix.
3. Number of observations for predicted health exceeds the number of observations underlying the MCA which counts observations with identical attributes once.

Gender Differences In Reported Health Problems

Proportion of Each Gender Reporting Each Health Problem in Each Wave

	A		B		C		D		E		F		F		G	
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
arms	0.208	0.249	0.202	0.247	0.211	0.275	0.215	0.279	0.217	0.282	0.229	0.289	0.229	0.289	0.247	0.304
sight	0.067	0.078	0.047	0.064	0.043	0.055	0.041	0.055	0.036	0.049	0.037	0.051	0.037	0.051	0.049	0.060
hearing	0.088	0.067	0.092	0.068	0.090	0.068	0.092	0.068	0.091	0.065	0.086	0.068	0.086	0.068	0.095	0.079
skin/allergy	0.078	0.124	0.081	0.116	0.091	0.127	0.084	0.120	0.089	0.115	0.096	0.134	0.120	0.134	0.093	0.143
chest/breathing	0.105	0.102	0.108	0.110	0.112	0.117	0.115	0.127	0.124	0.128	0.120	0.134	0.117	0.134	0.137	0.140
heart/blood pressure	0.108	0.126	0.110	0.128	0.115	0.135	0.111	0.134	0.112	0.139	0.117	0.143	0.059	0.143	0.134	0.157
stomach/digestion	0.053	0.057	0.055	0.065	0.055	0.070	0.058	0.065	0.060	0.064	0.059	0.070	0.023	0.070	0.067	0.087
diabetes	0.019	0.016	0.020	0.018	0.024	0.018	0.022	0.021	0.022	0.021	0.023	0.022	0.047	0.022	0.027	0.026
anxiety/depression	0.033	0.068	0.041	0.077	0.039	0.080	0.041	0.080	0.045	0.083	0.047	0.086	0.006	0.086	0.054	0.102
alcohol/drugs	0.005	0.003	0.005	0.002	0.006	0.002	0.006	0.001	0.007	0.003	0.006	0.002	0.008	0.002	0.009	0.003
epilepsy	0.007	0.006	0.008	0.008	0.008	0.007	0.008	0.008	0.008	0.008	0.008	0.008	0.048	0.008	0.008	0.009
migraine	0.043	0.109	0.043	0.119	0.046	0.119	0.041	0.116	0.040	0.116	0.048	0.132	0.048	0.132	0.054	0.129
other	0.034	0.065	0.034	0.058	0.029	0.053	0.027	0.048	0.030	0.056	0.034	0.064	0.034	0.064	0.035	0.061
N	4,815	5,416	4,608	5,196	4,459	5,115	4,432	5,035	4,309	4,926	4,434	5,001	4,434	5,001	5,206	5,977

	H		I		J		K		L		M		N		Average		
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Difference
arms	0.255	0.313	0.233	0.286	0.260	0.316	0.254	0.301	0.251	0.308	0.245	0.302	0.244	0.308	0.233	0.290	-0.057
sight	0.044	0.055	0.042	0.055	0.044	0.060	0.044	0.057	0.050	0.057	0.047	0.054	0.047	0.058	0.045	0.057	-0.012
hearing	0.097	0.074	0.102	0.068	0.107	0.070	0.097	0.069	0.104	0.072	0.098	0.069	0.096	0.072	0.095	0.070	0.025
skin/allergy	0.100	0.151	0.086	0.136	0.089	0.145	0.087	0.134	0.100	0.136	0.093	0.136	0.089	0.129	0.092	0.132	-0.041
chest/breathing	0.131	0.142	0.129	0.136	0.139	0.144	0.129	0.139	0.134	0.138	0.128	0.140	0.125	0.140	0.123	0.132	-0.008
heart/blood pressure	0.135	0.163	0.143	0.160	0.149	0.173	0.161	0.174	0.166	0.178	0.167	0.178	0.172	0.181	0.131	0.154	-0.024
stomach/digestion	0.066	0.087	0.061	0.072	0.071	0.088	0.074	0.094	0.074	0.093	0.070	0.089	0.070	0.089	0.061	0.077	-0.016
diabetes	0.029	0.026	0.034	0.026	0.037	0.028	0.041	0.030	0.045	0.030	0.045	0.034	0.049	0.036	0.032	0.025	0.007
anxiety/depression	0.050	0.105	0.047	0.090	0.056	0.111	0.061	0.113	0.053	0.111	0.059	0.105	0.058	0.102	0.046	0.093	-0.047
alcohol/drugs	0.007	0.003	0.006	0.002	0.007	0.002	0.009	0.003	0.008	0.003	0.009	0.003	0.007	0.003	0.007	0.003	0.004
epilepsy	0.007	0.009	0.007	0.009	0.008	0.010	0.009	0.009	0.007	0.009	0.008	0.008	0.009	0.008	0.011	0.008	0.002
migraine	0.047	0.131	0.044	0.117	0.046	0.128	0.041	0.111	0.043	0.116	0.037	0.105	0.035	0.102	0.044	0.119	-0.075
other	0.042	0.070	0.044	0.065	0.046	0.076	0.011	0.018	0.030	0.050	0.032	0.058	0.037	0.062	0.033	0.058	-0.025
N	5,048	5,848	7,178	8,429	7,162	8,422	8,628	10,237	7,619	8,964	7,521	8,713	7,260	8,529	5,808	6,721	-913

Index Weights for the Price of Medical Care

Hospital Days

MCA Instrument Weights for Price of Hospital Days

Instrument	Weight for Each Wave														
	B	C	D	E	F	G	H	I	J	K	L	M	N	mean	st. dev.
NHS															
no usage	0.108	0.124	0.136	0.136	0.122	0.142	0.156	0.162	0.16	0.15	0.138	0.16	0.146	0.142	0.017
all NHS	-1.175	-1.23	-1.321	-1.38	-0.1259	-1.434	-1.517	-1.355	-1.402	-1.364	-1.306	-1.455	-1.35	-1.263	0.353
all private	1.611	1.136	0.468	1.582	0.334	0.91	1.459	1.145	0.98	1.679	0.798	0.75	0.632	1.037	0.445
both	0.573	0.411	0.428	-0.142	1.56	1.059	0.729	-0.14	-0.192	-0.206	1.461	0.497	1.041	0.545	0.612
Insurance															
yes	1.165	1.111	1.103	1.095	1.048	1.266	1.271	1.396	1.403	1.442	1.342	1.333	1.296	1.252	0.133
no	-0.253	-0.239	-0.239	-0.236	-0.225	-0.252	-0.266	-0.242	-0.248	-0.24	-0.241	-0.239	-0.226	-0.242	0.011
Job Status															
self-employed	0.902	0.961	1.1	1.154	1.105	1.192	1.123	1.023	1.015	0.991	0.945	0.967	0.937	1.032	0.093
employed	1.269	1.273	1.253	1.231	1.253	1.24	1.22	1.29	1.279	1.305	1.285	1.272	1.275	1.265	0.024
unemployed	-1.857	-1.705	-1.712	-1.6	-1.659	-1.601	-1.517	-1.612	-1.605	-1.694	-1.674	-1.638	-1.631	-1.654	0.081
retired	-1.908	-1.928	-1.921	-1.922	-1.954	-1.832	-1.852	-1.791	-1.8	-1.746	-1.816	-1.824	-1.854	-1.858	0.064
maternity	0.37	0.226	0.134	0.179	-0.295	0.433	0.076	0.449	0.114	-0.07	0.29	-0.304	-0.205	0.107	0.259
family care	-1.805	-1.8	-1.823	-1.837	-1.867	-1.752	-1.724	-1.61	-1.674	-1.694	-1.763	-1.743	-1.859	-1.765	0.077
ft. student	-0.947	-0.839	-0.675	-0.804	-0.873	-0.613	-0.611	-0.643	-0.705	-0.582	-0.686	-0.65	-0.62	-0.711	0.116
long term sick	-2.148	-2.211	-2.173	-2.189	-2.233	-2.211	-2.22	-2.023	-2.049	-2.023	-2.015	-2.112	-2.051	-2.128	0.085
government training	-0.088	-0.075	-0.681	-0.619	-0.878	-0.858	-1.102	-1.093	-1.046	-0.483	-0.871	-1.047	-1.027	-0.759	0.356
other	0.021	-1.053	-1.066	-2.162	-0.621	-0.692	0.702	-1.062	-0.9	-0.88	-0.881	-0.831	-1.286	-0.824	0.664
Wage quartiles															
no wages	-1.874	-1.877	-1.892	-1.896	-1.916	-1.835	-1.842	-1.74	-1.76	-1.73	-1.79	-1.8	-1.831	-1.829	0.062
1	0.629	0.626	0.615	0.613	0.54	0.618	0.61	0.689	0.712	0.71	0.687	0.695	0.69	0.649	0.052
2	1.14	1.14	1.113	1.115	1.12	1.102	1.084	1.121	1.109	1.132	1.131	1.095	1.124	1.117	0.017
3	1.29	1.278	1.257	1.221	1.294	1.27	1.23	1.33	1.291	1.31	1.285	1.271	1.276	1.277	0.030
4	1.536	1.55	1.535	1.539	1.548	1.562	1.534	1.615	1.631	1.641	1.599	1.59	1.574	1.573	0.038
region															
Inner London	0.381	0.453	0.202	0.289	0.39	0.253	0.255	0.536	0.642	0.399	0.755	0.613	0.671	0.449	0.180
Outer London	0.358	0.212	0.321	0.409	0.397	0.524	0.538	0.615	0.629	0.789	0.531	0.486	0.443	0.481	0.150
Rest of SE	0.41	0.462	0.511	0.525	0.486	0.544	0.573	0.759	0.65	0.801	0.821	0.734	0.641	0.609	0.136
South West	-1.23	-0.048	-0.035	-0.011	-0.092	0.112	-0.003	0.167	0.142	0.188	0.207	0.18	0.14	-0.022	0.377
East Angila	-1.89	-0.284	-0.129	-0.163	-0.269	-0.11	-0.055	0.169	0.067	0.128	-0.035	-0.196	0.137	-0.202	0.529
East Midlands	-0.077	-0.146	-0.116	-0.168	-0.171	-0.05	-0.015	0.156	0.236	0.236	0.264	0.291	0.253	0.053	0.187
West Midlands	-0.447	-0.467	-0.586	-0.555	-0.654	-0.498	-0.399	-0.221	-0.262	-0.381	-0.322	-0.552	-0.584	-0.456	0.133
Rest of West Midlands	0.326	0.159	0.241	0.3	0.406	0.339	0.246	0.346	0.298	0.355	0.412	0.506	0.373	0.331	0.088
Greater Manchester	0.01	0.02	-0.211	-0.177	-0.151	-0.071	-0.078	0.128	0.048	0.154	0.251	0.245	0.353	0.040	0.178
Merseyside	-0.674	-0.76	-0.789	-0.564	-0.725	-0.851	-0.587	-0.31	-0.242	-0.7	-0.063	-0.178	-0.186	-0.510	0.274
Rest of North West	-0.056	0.087	0.121	-0.04	-0.129	0.029	0.184	0.122	0.278	0.314	0.278	0.234	0.3	0.132	0.148
South Yorkshire	-0.396	-0.384	-0.177	-0.294	-0.265	-0.371	-0.173	0.079	-0.068	0.138	0.284	0.282	0.445	-0.069	0.287
West Yorkshire	-0.223	-0.247	-0.298	-0.29	-0.258	-0.287	-0.303	-0.228	-0.2	-0.049	0.165	0.168	0.099	-0.150	0.181
Rest of Yorks & Humber	0.305	0.386	0.291	0.051	0.108	-0.101	-0.199	-0.077	0.053	0.09	-0.09	-0.191	0.011	0.049	0.187
Tyne & Wear	-0.344	-0.459	-0.34	-0.57	-0.519	-0.604	-0.511	-0.69	-0.544	-0.41	-0.401	-0.146	-0.315	-0.450	0.145
Rest of North	-0.146	-0.119	-0.206	-0.102	-0.115	-0.336	-0.345	-0.077	-0.26	-0.169	-0.5	0.059	0.12	-0.169	0.166
Wales	-0.589	-0.629	-0.647	-0.742	-0.515	-0.709	-0.778	-0.545	-0.473	-0.378	-0.432	-0.371	-0.348	-0.550	0.146
Scotland	-0.181	-0.137	-0.231	-0.141	-0.062	-0.32	-0.378	-0.146	-0.166	-0.042	-0.056	-0.039	0.028	-0.144	0.116
Northern Ireland						-0.278	-0.593	-0.479	-0.33	-0.407	-0.378	-0.363	-0.424	-0.407	0.097
Mean Price	3.038	3.369	3.116	3.183	3.121	2.857	2.824	2.996	2.570	2.389	3.014	2.939	3.258		
St. dev.	1.415	1.385	1.351	1.331	1.303	1.275	1.227	1.197	1.169	1.125	1.091	1.056	1.039		
Minimum	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		
Maximum	5.172	5.810	5.478	5.697	5.533	5.138	5.424	5.300	4.723	4.786	5.222	5.026	5.325		

Tests and Services

MCA Instrument Weights for Price of Tests and Services

	Weight for Each Wave														
Instrument	B	C[1]	D	E	F	G	H	I	J	K	L	M	N	mean	st. dev.
NHS															
no usage	0.051	1.796	0.038	0.087	0.294	0.337	0.514	0.243	0.390	0.350	0.415	0.422	0.582	0.425	0.445
all NHS	-0.231	-0.044	-0.309	-0.338	-0.380	-0.397	-0.412	-0.400	-0.420	-0.441	-0.427	-0.427	-0.441	-0.359	0.112
all private	0.849	0.668	0.963	1.020	0.919	1.138	1.034	1.171	1.192	1.280	1.232	1.206	1.107	1.060	0.175
both	1.326	0.976	1.441	1.525	1.701	1.621	1.629	1.686	1.675	1.691	1.606	1.532	1.759	1.551	0.210
Insurance															
yes	1.229	0.922	1.242	1.202	1.168	1.384	1.382	1.530	1.516	1.524	1.436	1.452	1.406	1.338	0.177
no	-0.265	-0.208	-0.269	-0.258	-0.250	-0.276	-0.291	-0.265	-0.268	-0.253	-0.259	-0.261	-0.245	-0.259	0.019
Job Status															
self-employed	0.904	0.922	1.094	1.197	1.126	1.215	1.150	1.045	1.041	1.017	0.997	0.981	0.975	1.051	0.099
employed	1.262	1.152	1.228	1.195	1.212	1.202	1.189	1.241	1.227	1.244	1.219	1.225	1.215	1.216	0.028
unemployed	-1.953	-1.814	-1.755	-1.683	-1.668	-1.589	-1.526	-1.627	-1.596	-1.662	-1.597	-1.596	-1.576	-1.665	0.116
retired	-1.874	-1.968	-1.838	-1.820	-1.843	-1.726	-1.746	-1.682	-1.705	-1.648	-1.719	-1.735	-1.748	-1.773	0.089
maternity	0.784	0.748	0.761	0.614	0.189	1.010	0.762	1.186	0.628	0.463	0.734	0.311	0.221	0.647	0.292
family care	-1.728	-1.830	-1.704	-1.732	-0.176	-0.652	-1.665	-1.533	-1.563	-1.587	-1.605	-1.600	-1.730	-1.470	0.486
ft. student	-1.065	-0.100	-0.859	-0.976	-0.990	-0.752	-0.799	-0.082	-0.879	-0.747	-0.853	-0.874	-0.781	-0.751	0.308
long term sick	-2.010	-2.163	-2.109	-2.158	-2.140	-2.030	-2.061	-1.901	-1.939	-1.912	-1.919	-1.953	-1.979	-2.021	0.097
government training	-0.829	-0.799	-0.945	-0.967	-1.017	-0.964	-1.446	-1.197	-1.210	-0.581	-0.979	-1.298	-1.162	-1.030	0.230
other	0.068	-1.031	-1.429	-2.257	-0.199	-0.645	0.688	-1.159	-0.895	-0.810	-0.919	-0.784	-1.133	-0.808	0.718
Wage quartiles															
no wages	-1.843	-1.941	-1.816	-1.814	-1.834	-1.739	-1.762	-1.661	-1.679	-1.647	-1.696	-1.710	-1.738	-1.760	0.086
1	0.567	0.478	0.505	0.506	0.427	0.468	0.467	0.542	0.583	0.576	0.543	0.536	0.553	0.519	0.048
2	1.092	1.018	1.054	1.044	1.025	0.986	0.989	1.005	0.991	1.013	0.999	1.009	1.008	1.018	0.030
3	1.300	1.233	1.229	1.219	1.277	1.244	1.233	1.298	1.265	1.269	1.247	1.232	1.227	1.252	0.027
4	1.608	1.428	1.660	1.650	1.636	1.702	1.648	1.727	1.713	1.714	1.688	1.687	1.668	1.656	0.077
region															
Inner London	0.378	0.558	0.229	0.360	0.415	0.340	0.274	0.527	0.666	0.388	0.880	0.612	0.731	0.489	0.193
Outer London	0.438	0.161	0.352	0.451	0.459	0.591	0.589	0.663	0.659	0.793	0.586	0.484	0.478	0.516	0.159
Rest of SE	0.452	0.439	0.601	0.633	0.561	0.618	0.640	0.838	0.733	0.909	0.915	0.837	0.755	0.687	0.159
South West	-0.123	-0.060	0.015	0.005	-0.056	0.192	0.048	0.248	0.224	0.224	0.282	0.296	0.206	0.115	0.147
East Angila	-0.238	-0.387	-0.187	-0.241	-0.341	-0.186	-0.140	0.022	-0.014	0.013	-0.151	-0.185	-0.256	-0.176	0.126
East Midlands	-0.120	-0.154	-0.155	-0.243	-0.240	-0.161	-0.084	0.107	0.154	0.146	0.147	0.200	0.151	-0.019	0.170
West Midlands	-0.497	-0.584	-0.528	-0.542	-0.604	-0.492	-0.286	-0.215	-0.315	-0.295	-0.231	-0.439	-0.563	-0.430	0.141
est of West Midlands	0.325	0.093	0.248	0.246	0.392	0.296	0.145	0.317	0.268	0.343	0.357	0.414	0.325	0.290	0.091
Greater Manchester	-0.070	-0.065	-0.296	-0.236	-0.161	-0.100	-0.049	0.032	-0.047	0.089	0.239	0.222	0.364	-0.006	0.192
Merseyside	-0.717	-0.718	-0.858	-0.598	-0.870	-0.917	-0.674	-0.360	-0.363	-0.135	-0.157	-0.298	-0.280	-0.534	0.279
Rest of North West	-0.132	-0.015	0.085	-0.110	-0.172	-0.021	0.067	0.155	0.195	0.251	0.243	0.155	0.190	0.069	0.146
South Yorkshire	-0.471	-0.237	-0.312	-0.401	-0.372	-0.475	-0.249	0.076	-0.160	-0.007	0.124	0.137	0.280	-0.159	0.255
West Yorkshire	-0.232	-0.274	-0.367	-0.284	-0.338	-0.325	-0.412	-0.344	-0.358	-0.155	0.009	0.161	0.088	-0.218	0.187
st of Yorks & Humber	0.381	0.291	0.233	0.037	0.054	-0.149	-0.283	-0.146	-0.036	0.023	-0.136	-0.241	-0.037	-0.001	0.202
Tyne & Wear	-0.284	-0.355	-0.455	-0.657	-0.547	-0.632	-0.536	-0.723	-0.622	-0.472	-0.512	-0.264	-0.306	-0.490	0.151
Rest of North	-0.147	-0.105	-0.166	-0.009	-0.061	-0.305	-0.274	-0.067	-0.264	-0.138	-0.064	0.073	0.091	-0.110	0.124
Wales	-0.547	-0.379	-0.652	-0.720	-0.503	-0.621	-0.708	-0.496	-0.430	-0.381	-0.425	-0.364	-0.342	-0.505	0.134
Scotland	-0.197	-0.062	-0.276	-0.221	-0.135	-0.399	-0.413	-0.208	-0.173	-0.076	-0.073	-0.054	0.028	-0.174	0.133
Northern Ireland						-0.307	-0.552	-0.275	-0.209	-0.368	-0.366	-0.379	-0.436	-0.362	0.104
Mean Price	2.731	3.995	2.602	2.653	2.802	2.988	2.466	2.646	2.509	2.046	2.171	2.037	2.447		
St. dev.	1.190	2.075	1.080	1.054	1.057	1.052	1.076	0.982	0.970	0.919	0.885	0.860	0.890		
Minimum	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		
Maximum	5.397	6.496	5.194	5.167	5.159	5.284	4.575	5.061	4.498	4.450	4.341	4.092	4.432		

1. For wave C 706 observations with zero usage and zero wage are eliminated. These "outlier" observations shifted the mass such that all of the signes were reversed. The elimination of these outliers with zero usage makes signs consistent with other waves, but increases the average price for wave C.

General Practitioner Visits

MCA Instrument Weights for Price of General Practitioner Visits

	Weight for Each Wave														
Instrument	B	C	D	E	F	G	H	I	J	K	L	M	N	mean	st. dev.
Usage															
1	0.494	0.497	0.643	0.544	0.603	0.636	0.571	0.636	0.688	0.696	0.644	0.695	0.711	0.620	0.074
2	0.341	0.387	0.291	0.383	0.319	0.401	0.408	0.424	0.394	0.4	0.403	0.372	0.313	0.372	0.042
3	-0.261	-0.344	-0.293	-0.238	-0.318	-0.3	-0.177	-0.274	-0.252	-0.26	-0.29	-0.342	-0.379	-0.287	0.053
4	0.889	-0.713	-0.861	-0.908	-0.641	-0.828	-0.924	-0.854	-0.862	-0.788	-0.911	-0.945	-0.934	-0.714	0.490
5	-1.673	-1.67	-1.536	-1.608	-1.637	-1.594	-1.721	-1.65	-1.62	-1.636	-1.554	-1.558	-1.593	-1.619	0.053
Job Status															
self-employed	0.933	0.988	1.068	1.098	1.069	1.148	1.1	0.982	0.961	0.961	0.947	0.956	0.916	1.010	0.076
employed	1.188	1.155	1.137	1.1	1.097	1.099	1.08	1.133	1.127	1.146	1.123	1.119	1.108	1.124	0.029
unemployed	-1.342	-1.315	-1.259	-1.268	-1.391	-1.264	-1.212	-1.287	-1.242	-1.333	-1.302	-1.321	-1.342	-1.298	0.049
retired	-1.643	-1.673	-1.679	-1.655	-1.715	-1.614	-1.631	-1.592	-1.596	-1.541	-1.615	-1.644	-1.655	-1.635	0.045
maternity	-0.057	0.127	-0.042	-0.007	-0.456	-0.109	-0.028	0.15	0.17	0.104	0.277	-0.091	-0.023	0.001	0.180
family care	-1.548	-1.561	-1.564	-0.592	-1.617	-1.527	-1.589	-1.413	-1.468	-1.496	-1.539	-1.541	-1.606	-1.466	0.269
ft. student	-0.707	-0.538	-0.55	-0.674	-0.67	-0.424	-0.412	-0.41	-0.518	-0.362	-0.405	-0.389	-0.306	-0.490	0.131
long term sick	-2.242	-2.33	-2.365	-2.414	-2.372	-2.297	-2.303	-2.124	-2.186	-2.162	-2.152	-2.147	-2.111	-2.247	0.106
government training	-0.0453	-0.383	-0.339	-0.461	-0.604	-0.762	-0.63	-0.727	-0.593	-0.008	-0.625	-0.647	-0.881	-0.516	0.262
other	-0.603	-0.582	-0.82	-1.777	-0.793	-0.46	0.281	-0.977	-0.668	-0.756	-0.765	-0.731	-1.047	-0.746	0.448
Wage quartiles															
no wages	-1.583	-1.61	-1.641	-1.655	-1.696	-1.627	-1.654	-1.55	-1.569	-1.541	-1.598	-1.609	-1.627	-1.612	0.044
1	0.686	0.632	0.607	0.563	0.5	0.616	0.617	0.673	0.69	0.693	0.688	0.691	0.701	0.643	0.061
2	1.143	1.109	1.052	1.055	1.035	1.039	1.024	1.067	1.067	1.077	1.044	1.035	1.037	1.060	0.034
3	1.205	1.196	1.183	1.138	1.184	1.157	1.127	1.191	1.17	1.19	1.167	1.158	1.136	1.169	0.025
4	1.323	1.307	1.297	1.281	1.272	1.272	1.246	1.295	1.295	1.306	1.282	1.275	1.251	1.285	0.022
region															
Inner London	0.179	0.383	0.23	0.232	0.246	0.2	0.18	0.465	0.591	0.359	0.684	0.519	0.636	0.377	0.184
Outer London	0.224	0.055	0.241	0.358	0.258	0.384	0.387	0.412	0.472	0.529	0.337	0.304	0.236	0.323	0.123
Rest of SE	0.282	0.336	0.354	0.425	0.329	0.421	0.448	0.573	0.509	0.608	0.606	0.531	0.47	0.453	0.109
South West	0.008	0.018	-0.001	0.03	-0.052	0.075	0.023	0.211	0.163	0.178	0.196	0.14	0.139	0.087	0.088
East Angila	-0.11	-0.164	-0.023	-0.072	-0.202	-0.1	-0.049	0.177	0.099	0.189	-0.02	-0.132	-0.109	-0.040	0.124
East Midlands	-0.043	-0.077	-0.065	-0.114	-0.111	0.006	0.03	0.168	0.216	0.205	0.262	0.274	0.262	0.078	0.155
West Midlands	0.407	-0.488	-0.577	-0.566	-0.612	-0.535	-0.414	-0.272	-0.298	-0.455	-0.423	-0.632	-0.727	-0.430	0.283
est of West Midlands	0.29	0.188	0.264	0.312	0.368	0.338	0.187	0.358	0.305	0.301	0.313	0.437	0.376	0.311	0.070
Greater Manchester	0.061	-0.001	-0.105	-0.125	-0.039	0.008	-0.081	0.114	0.066	0.096	0.213	0.219	0.303	0.056	0.132
Merseyside	-0.4	-0.518	-0.569	-0.354	-0.522	-0.629	-0.404	-0.146	-0.184	-0.11	-0.054	-0.15	-0.15	-0.322	0.199
Rest of North West	-0.086	0.059	0.075	-0.053	-0.132	-0.007	0.109	0.104	0.191	0.265	0.185	0.199	0.2	0.085	0.124
South Yorkshire	-0.345	-0.351	-0.138	-0.343	-0.221	-0.306	-0.168	0.08	-0.119	0.017	0.13	0.159	0.269	-0.103	0.214
West Yorkshire	-0.219	-0.237	-0.324	-0.342	-0.227	-0.319	-0.285	-0.276	-0.22	-0.109	0.055	0.065	0.083	-0.181	0.154
st of Yorks & Humber	0.338	0.27	0.191	0.017	0.108	-0.063	-0.127	-0.073	0.125	0.114	-0.068	-0.152	0.009	0.053	0.153
Tyne & Wear	-0.3	-0.356	-0.276	-0.353	-0.395	-0.497	-0.409	-0.59	-0.42	-0.34	-0.387	-0.18	-0.231	-0.364	0.108
Rest of North	-0.068	-0.06	-0.199	-0.154	-0.091	-0.279	-0.283	-0.062	-0.185	-0.142	0.001	0.165	0.091	-0.097	0.132
Wales	-0.447	-0.494	-0.523	-0.704	-0.438	-0.511	-0.585	-0.469	-0.448	-0.315	-0.329	-0.313	-0.289	-0.451	0.120
Scotland	-0.139	-0.085	-0.12	-0.082	0.047	-0.213	-0.263	-0.1	-0.1	0.027	-0.015	0.011	0.065	-0.074	0.099
Northern Ireland						-0.189	-0.468	-0.378	-0.252	-0.337	-0.301	-0.278	-0.338	-0.318	0.084
Mean Price	2.838	2.880	2.850	2.923	2.856	3.212	3.182	2.436	2.868	2.743	2.624	2.542	2.974		
St. dev.	1.368	1.353	1.333	1.321	1.293	1.260	1.219	1.191	1.160	1.116	1.078	1.044	1.025		
Minimum	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		
Maximum	5.120	4.487	4.594	4.661	4.378	4.901	4.815	4.117	4.537	4.515	4.315	4.174	4.563		

Appendix F: Estimation Results

Multiple Equation Results
NLSUR with Poisson form $\mathbf{y} = \exp(\mathbf{X}'\boldsymbol{\beta}) + \boldsymbol{\varepsilon}$
Coefficient Estimates and p-values reported in parentheses
* Indicates cross-price coefficients restricted for symmetry.

Hospital Days (Hdays)

	4 Equation Model				3 Equation Model			
	Full Sample		OSM Sample		Full Sample		OSM Sample	
coefficient	unrestricted	restricted	unrestricted	restricted	unrestricted	restricted	unrestricted	restricted
bhπ (change in health)	-0.0258	-0.0285	-0.0279	-0.0287	-0.0253	-0.0268	-0.0269	-0.0255
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
bhH (lagged health)	-0.0229	-0.0236	-0.0267	-0.0251	-0.0223	-0.0219	-0.0254	-0.0215
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
chh	-1.8007	-1.6547	-1.8652	-1.7537	-1.8268	-1.6468	-1.9047	-1.84681
	(0.000)	(0.000)	0.0000	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
chz*	-0.8754	0.0890	-0.7165	0.0789	-1.0006	0.1016	-0.9358	0.0944
	(0.159)	(0.000)	(0.582)	(0.000)	(0.108)	(0.000)	(0.469)	(0.000)
chts*	0.8664	0.1071	1.0487	0.1157	0.8659	0.1801	1.0372	0.1890
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
chgp*	-0.0102	0.2640	0.0038	0.2808	-0.0356	0.0605	-0.0427	-0.0572
	(0.896)	(0.000)	(0.985)	(0.000)	(0.650)	(0.487)	(0.833)	(0.776)
bhh (net income)	0.0045	0.0037	0.0033	0.0016	0.0046	0.0040	0.0036	0.0057
	(0.002)	(0.061)	(0.607)	(0.853)	(0.001)	(0.008)	(0.569)	(0.601)
bhc (couple)	-0.4165	-0.3694	-0.5625	-0.5279	-0.4158	-0.3889	-0.5725	-0.6017
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
bhE (education)	0.0030	0.0448	0.0148	0.0319	0.0012	0.0342	0.0060	-0.0101
	(0.000)	(0.505)	(0.924)	(0.832)	(0.986)	(0.614)	(0.969)	(0.948)
bh (constant)	5.2973	4.2443	5.2443	4.4773	5.4643	4.3483	5.5364	4.7445
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Tests and Services (TS)

	4 Equation Model				3 Equation Model			
	Full Sample		OSM Sample		Full Sample		OSM Sample	
coefficient	unrestricted	restricted	unrestricted	restricted	unrestricted	restricted	unrestricted	restricted
bh□ (change in health)	-0.0079	-0.0109	-0.0081	-0.0112	-0.0079	-0.0094	-0.0080	-0.0099
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
bhH (lagged health)	-0.0109	-0.0144	-0.0110	-0.0147	-0.0109	-0.0130	-0.0110	-0.0135
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ctsts	0.0011	-0.2265	0.1244	-0.2261	0.1723	-0.0913	0.1313	-0.10566
	-0.0010	(0.000)	0.0000	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ctsz*	-2.5734	0.0385	-2.4127	0.0456	-2.5849	0.0365	-2.4137	0.0483
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ctsh*	-0.0004	0.1071	0.0389	0.1157	-0.0168	0.1801	0.0219	0.1890
	(0.943)	(0.000)	(0.000)	(0.000)	(0.006)	(0.000)	(0.042)	(0.000)
ctsgp*	-0.2681	0.0353	-0.2590	0.0310	-0.2639	-0.1904	-0.2512	-0.1774
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
btss (net income)	0.0011	0.0017	0.0015	0.0028	0.0013	0.0020	0.0018	0.0031
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
btsc (couple)	0.0714	0.0952	0.0462	0.0563	0.0736	0.0821	0.0481	0.0451
	(0.000)	(0.000)	(0.004)	(0.001)	(0.000)	(0.000)	(0.003)	-0.0090
btsE (education)	0.1513	0.1762	0.1393	0.1567	0.1539	0.1845	0.1401	0.1619
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
bts (constant)	5.0689	2.1289	4.8950	2.1682	5.0857	2.1293	4.8992	2.1609
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Multiple Equation Specification Results continued

Multiple Equation Results for NLSUR with Poisson form $\mathbf{y} = \exp(\mathbf{X}'\boldsymbol{\beta}) + \boldsymbol{\varepsilon}$
Coefficient Estimates and p-values reported in parentheses
* Indicates cross-price coefficients restricted for symmetry.

General Practitioner Visits (GP)

	4 Equation Model			
	Full Sample		OSM Sample	
coefficient	unrestricted	restricted	unrestricted	restricted
bgp□ (change in health)	-0.0036	-0.0042	-0.0043	-0.0046
	(0.000)	(0.000)	(0.000)	(0.000)
bgpH (lagged health)	-0.0044	-0.0051	-0.0053	-0.0055
	(0.000)	(0.000)	(0.000)	(0.000)
cgpgp	-0.4936	-0.4224	-0.4812	-0.4204
	(0.000)	(0.000)	0.0000	(0.000)
cgpz*	-0.4831	-0.0012	-0.2663	0.0072
	(0.000)	(0.847)	(0.000)	(0.518)
cgph*	0.1983	0.2640	0.2227	0.2808
	(0.000)	(0.000)	(0.000)	(0.000)
cgpts*	0.1685	0.0353	0.1484	0.0310
	(0.000)	(0.000)	(0.000)	(0.000)
bgpp (net income)	-0.0016	-0.0016	-0.0024	-0.0024
	(0.000)	(0.000)	(0.000)	(0.000)
bgpc (couple)	0.0276	0.0347	0.0552	0.0588
	(0.000)	(0.000)	(0.000)	(0.000)
bgpE (education)	0.0232	0.0285	0.0122	0.0158
	(0.000)	(0.000)	(0..174)	(0.079)
bgp (constant)	2.1796	1.5948	1.9259	1.5649
	(0.000)	(0.000)	(0.000)	(0.000)

Consumption (Z)

	4 Equation Model				3 Equation Model			
	Full Sample		OSM Sample		Full Sample		OSM Sample	
coefficient	unrestricted	restricted	unrestricted	restricted	unrestricted	restricted	unrestricted	restricted
bZ□ (change in health)	0.0031	0.0034	0.0039	0.0037	0.0035	0.0035	0.0037	0.0038
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
bZH (lagged health)	0.0060	0.0057	0.0057	0.0053	0.0059	0.0059	0.0055	0.0056
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
czz	-0.6876	-0.5723	-0.7968	-0.6836	-0.6843	-0.5634	-0.7931	-0.68698
	(0.000)	(0.000)	0.0000	(0.000)	(0.000)	(0.000)	(0.067)	(0.007)
czh*	0.0727	0.0890	0.0642	0.0789	0.0721	0.1016	0.0632	0.0944
	(0.000)	(0.000)	(0.026)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
czts*	0.0762	0.0385	0.0899	0.0456	0.0760	0.0365	0.0890	0.0483
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.010)	(0.000)	(0.000)
czgp*	-0.0257	-0.0012	-0.0257	0.0072	-0.0246	-0.0150	-0.0243	-0.0156
	(0.000)	(0.847)	(0.025)	(0.518)	(0.000)	(0.022)	(0.034)	(0.178)
bzz (net income)	0.0036	0.0031	0.0057	0.0057	0.0031	0.0031	0.0057	0.0057
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
bzc (couple)	0.1906	0.1919	0.3259	0.3258	0.1902	0.1914	0.3257	0.3251
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
bzE (education)	0.1598	0.1613	0.1172	0.1185	0.1600	0.1619	0.1174	0.1190
	(0.000)	(0.000)	(0.000)	(0.000)	(0.618)	(0.638)	(0.000)	(0.000)
bz (constant)	1.6182	1.4792	1.5762	1.4390	1.6219	1.4608	1.5831	1.4322
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Single Equation Results

Hospital Days (Hdays)

	Single Equation Models			
	Fixed Effect NB		NB	
coefficient	Full	OSM	Full	OSM
bh° (change in health)	-0.0191	-0.0259	-0.0459	-0.0526
	(0.000)	(0.000)	(0.000)	(0.000)
bhH (lagged health)	-0.0161	-0.0247	-0.0490	-0.0571
	(0.000)	(0.000)	(0.000)	(0.000)
chh	-2.6488	-2.6681	-2.8348	-2.6179
	(0.000)	(0.000)	0.0000	(0.000)
chz	-3.1696	-1.4338	-6.6457	-4.6802
	(0.163)	(0.000)	(0.000)	(0.000)
chts	2.1136	2.1272	3.1377	2.9508
	(0.000)	(0.000)	(0.000)	(0.000)
chgp	0.1061	0.1249	-0.6519	-0.6572
	(0.000)	(0.000)	(0.000)	(0.000)
bhh (net income)	0.0071	0.0098	0.0042	0.0020
	(0.000)	(0.000)	(0.027)	(0.502)
bhc (couple)	0.2749	0.2705	0.0396	0.0682
	(0.000)	(0.000)	(0.297)	(0.412)
bhE (education)	0.4284	0.3866	0.2636	0.2661
	(0.000)	(0.000)	(0.000)	(0.001)
bh (constant)	3.7546	2.2054	13.0438	11.1455
	(0.000)	(0.000)	(0.000)	(0.000)
N	52,712	22,224	119,750	40,896

The fixed effects model eliminates observations with zero hospital use over the entire sample period.
I estimate the fixed effects negative binomial for robustness to the overdispersion issue despite misgivings regarding the incidental parameters problem raised by Greene (2007).

Tests and Services (TS)

	Single Equation	
	NB	
coefficient	Full	OSM
bh≅ (change in health)	-0.0081	-0.0108
	(0.000)	(0.000)
bhH (lagged health)	-0.0110	-0.0144
	(0.000)	(0.000)
ctsts	0.1499	-0.2501
	(0.000)	(0.000)
ctsz*	-2.6443	0.0985
	(0.000)	(0.000)
ctsh*	0.0409	0.1233
	(0.000)	(0.000)
ctsgp*	-0.3057	0.0318
	(0.000)	(0.000)
btss (net income)	0.0011	0.0017
	(0.000)	(0.000)
btsc (couple)	0.0712	0.0946
	(0.000)	(0.000)
btsE (education)	0.1502	0.1775
	(0.000)	(0.000)
bts (constant)	5.1223	2.0607
	(0.000)	(0.000)

The fixed effects negative binomial model for tests and services did not converge. The function was not concave.

General Practitioner Visits (GP)

coefficient	Single Equation Models: Ordered Logit, Full	Ordered Logit, OSM	RE Ordered Probit, Full	RE Ordered Probit, OSM
bgp□ (change in health)	-0.0292	-0.0337	-0.0205	-0.0213
	(0.000)	(0.000)	(0.000)	(0.000)
bgpH (lagged health)	-0.0342	-0.0404	-0.0222	-0.0212
	(0.000)	(0.000)	(0.000)	(0.000)
cgpgp	-2.3320	-2.2553	-1.6249	-1.5773
	(0.000)	(0.000)	0.0000	(0.000)
cgpz	-2.1009	-1.0891	-1.5010	-1.2830
	(0.000)	(0.000)	(0.000)	(0.000)
cgph	0.8886	0.9871	0.3757	0.4290
	(0.000)	(0.000)	(0.000)	(0.000)
cgpts	0.8968	0.7933	0.7808	0.7279
	(0.000)	(0.000)	(0.000)	(0.000)
bgpp (net income)	-0.0042	-0.0075	-0.0005	-0.0019
	(0.000)	(0.000)	(0.143)	(0.003)
bgpc (couple)	0.1124	0.2741	0.0795	0.1221
	(0.000)	(0.000)	(0.000)	(0.000)
bgpE (education)	0.1649	0.0921	0.1473	0.0546
	(0.000)	(0.038)	(0..243)	(0.088)
bgp (constant)[1]	14.5730	13.8927	9.5375	9.0924
	(0.000)	(0.000)	(0.000)	(0.000)

The constant term is -/cut1 from the STATA output.
A RUM Mixed logit model for GP visits would not converge.

Consumption (Z)

coefficient	Single Equation Models: RE Tobit, Full	RE Tobit, OSM	Tobit, Full	Tobit, OSM
bh□ (change in health)	0.0108	0.0115	0.0138	0.0150
	(0.000)	(0.000)	(0.000)	(0.000)
bhH (lagged health)	0.0203	0.0180	0.0211	0.0197
	(0.000)	(0.000)	(0.000)	(0.000)
czz	-0.9879	-1.6219	-1.7532	-2.7047
	(0.000)	(0.000)	0.0000	(0.000)
czh	-0.3096	0.2403	0.3024	0.1849
	(0.000)	(0.000)	(0.000)	(0.000)
czts	0.0159	0.0518	0.4104	0.5833
	(0.579)	(0.235)	(0.000)	(0.000)
czgp	0.0646	0.0543	-0.0498	-0.0728
	(0.009)	(0.255)	(0.063)	(0.113)
bzz (net income)	0.0583	0.0601	0.1017	0.1102
	(0.000)	(0.000)	(0.000)	(0.000)
bzc (couple)	1.1061	1.6558	0.9799	1.4875
	(0.000)	(0.000)	(0.000)	(0.000)
bzE (education)	1.3561	1.1275	0.8266	0.5672
	(0.000)	(0.000)	(0.000)	(0.000)
bz (constant)	2.9199	3.3202	2.5594	3.0765
	(0.000)	(0.000)	(0.000)	(0.000)

There is no sufficient statistic to consistently estimate a fixed effects tobit model.

Cross-Equation Variance-Covariance Estimates

Cross-Equation Variance-Covariance Estimates

4 Equation FULL Sample

unrestricted estimates

Equation	consumption	hospital days	tests & services	GP visits
consumption	14.4979			
hospital days	0.1969	33.4166		
tests & services	0.0779	1.6368	3.9955	
GP visits	0.1170	0.5167	0.7744	0.8570

restricted estimates

Equation	consumption	hospital days	tests & services	GP visits
consumption	14.5120			
hospital days	0.2347	33.8277		
tests & services	0.0938	1.8855	4.3585	
GP visits	0.1235	0.5970	0.8575	0.8768

3 Equation FULL Sample

unrestricted estimates

Equation	consumption	hospital days	tests & services
consumption	14.4978		
hospital days	0.1958	33.4090	
tests & services	0.0706	1.6242	3.9943

restricted estimates

Equation	consumption	hospital days	tests & services
consumption	14.5112		
hospital days	0.2312	33.7241	
tests & services	0.0604	1.8092	4.2460

4 Equation OSM Sample

unrestricted estimates

Equation	consumption	hospital days	tests & services	GP visits
consumption	13.9841			
hospital days	0.1959	23.0411		
tests & services	0.0978	1.3325	3.7799	
GP visits	0.1315	0.4279	0.7309	0.8260

restricted estimates

Equation	consumption	hospital days	tests & services	GP visits
consumption	14.0059			
hospital days	0.2449	23.3457		
tests & services	0.1121	1.5020	4.1218	
GP visits	0.1447	0.4771	0.7921	0.8414

3 Equation OSM Sample

unrestricted estimates

Equation	consumption	hospital days	tests & services
consumption	13.9839		
hospital days	0.1950	23.0327	
tests & services	0.0913	1.3195	3.7789

restricted estimates

Equation	consumption	hospital days	tests & services
consumption	13.9986		
hospital days	0.2217	23.2474	
tests & services	0.0636	1.4218	4.0240

Bibliography

Acton, J.P. (1975) "Nonemonetary Factors in the Demand for Medical Services: Some Empirical Evidence," *The Journal of Political Economy* 83(3):595-614.

Akerlof, G.A. (1970) "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics* 84:488 – 500.

Alemayehu, B., and K.E. Warner (2004) "The Lifetime Distribution of Health Care Costs," *Health Services Research* 39(3):627-642.

Andersen, R., and J.F. Newman (1973) "Societal and Individual Determinants of Medical Care Utilization in the United States," *The Milbank Memorial Fund Quarterly: Health and Society* 51(1):95 – 124.

Arellano, M. and J. Hahn (2005) "Understanding Bias in Nonlinear Panel Models: Some Recent Developments." CEMFI Working Paper #0507.

Arrow, K.J. (1963) "Uncertainty and the Welfare Economics of Medical Care," *The American Economic Review* 53(5):851-883.

Arthur, W.B. (1981) "The Economics of Risks to Life," *The American Economic Review* 71:54-64.

Atkinson, S.E., and T.D. Crocker (1992): "Econometric Health Production Functions: Relative Bias from Omitted Variables and Measurement Error," *Journal of Environmental Economics and Management,* 22: 12 – 24.

Bardasi, E., S.P. Jenkins, H. Sutherland, H. Levy, and F. Zantomio (2007) *British Household Panel Survey Derived Current and Annual Net Household Income Variables, Waves 1-14, 1991-2005* [computer file]. *7th Edition.* University of Essex. Institute for Social and Economic Research, [original data producer(s)]. Colchester, Essex: UK Data Archive [distributor] SN: 3909.

Becker, G. S. (1964) *Human Capital* Columbia University Press for the National Bureau of Economic Research, New York.

Becker, G. S., K.M. Murphy, and T.J. Philipson, (2007) "The Value of Life Near its End and Terminal Care," NBER Working Paper No. W13333.

Ben-Porath Y. (1967) "The production of human capital and the life cycle of earnings," *Journal of Political Economy* 75:353-376.

Berger, M.C., G.C. Blomquist, D. Kenel and G.S. Tolley, (1987) "Valuing Changes in Health Risks: A Comparison of Alternative Measures," *Southern Economic Journal* 53: 967-84.

Blau, D.M. and D.B. Gilleskie (2006) "Health Insurance and Retirement of Married Couples," *Journal of Applied Econometrics* 21:935-953.

Blau, D.M. and D.B. Gilleskie (2008) "the Role of Retiree Health Insurance in the Employment Behavior of Older Men," *International Economic Review* 49(2):475-514.

Blaxter, M. (1985) "Self-definition of Health Status and Consulting Rates in Primary Care," *Quarterly Journal of Social Affairs* 1(2):131-171.

Bolin, K., L. Jacobson, and B. Lindgren (2002) "The Family as the Health Producer – When Spouses Act Strategically," *Journal of Health Economics* 21:475-495.

Bruni, F. (2006) "Be Merry, Not Ancient," *The New York Times*, April 9.

Burman, B., and G. Margolin (1992) "Analysis of the Association Between marital Relationships and Health Problems: An Interactional Perspective," *Psychological Bulletin* 112:39-63.

Cameron, A.C., P.K. F. Trivedi, F. Milne, and J. Piggott (1988) "A Microeconomic Model of the Demand for Health Care and Health Insurance in Australia," *The Review of Economic Studies* 55(1):85-106.

Caputo, M. (2005) *Foundations of Dynamic Economic Analysis: Optimal Control Theory and Applications*, Cambridge University Press: USA.

Carbone, J.C., S. Kverndokk, and O.J. Rogeberg (2005) "Smoking, Health, Risk, and Perception," *Journal of Health Economics* 24:631-653.

Cassileth, B. R, E.J. Lusk, T.B. Strouse, D.S. Miller, L.L. Brown, P.A. Cross and A.N. Tenaglia (1984) "Psychological Status in Chronic Illness. A Comparative Analysis of Six Diagnostic Groups," *New England Journal of Medicine* 311: 506 – 511.

Cauley, S.D. (1987) "The Time Price of Medical Care," *The Review of Economics and Statistics* 69(1):59-66.

Chang, F.R. (1996) "Uncertainty and Investment in Health," *Journal of Health Economics* 15:369-376.

Cohen, G., J. Forbes, and M. Garraway (1995) "Interpreting Self-Reported Limiting Long-term Illness," *British Medical Journal* 311:722-724.

Constantinides, G.M. (1990) "Habit Formation: A Resolution of the Equity Premium Puzzle," *The Journal of Political Economy* 98(3):519-543.

Contoyannis, P., A.M. Jones, and N. Rice (2004) "The Dynamics of Health in the British Household Panel Survey," *Journal of Applied Econometrics* 19:473-503.

Conwell, L.J., and J.W. Cohen (2005) “Characteristics of Persons with High Medeical Expenditures in the U.S. Civilian Noninstitutionalized Population, 2002, “ Statistical Brief #73 Agency for Health Care Research and Quality: Rockville, MD http://meps.ahrq.gov/mepsweb/data_files/publications/st73/stat73.pdf

Cropper, M.L. (1977) “Health, Investment in Health, and Occupational Choice,” *Journal of Political Economy* 85:273-294.

Dardanoni, V. (1986) “A Note on a Simple Model of Health Investment [On the Concept of Health Capital and the Demand for Health]” *Bulletin of Economic Research* 38(1):97-100.

Dardanoni, V. and A. Wagstaff (1990) “Uncertainty and the Demand for Medical Care,” *Journal of Health Economics* 9:23-38.

Deaton, A.S., and C.H. Paxson, (1998) “Ageing and Inequality in Income and Health,” *American Economic Review, Papers and Proceedings* 88:248-253.

Deb, P. and P.K. Trivedi (1997) “Demand for Medical Care by the Elderly: A Finite Mixture Approach,” *Journal of Applied Econometrics* 12:313-336.

Disney, R., C. Emmerson, and M. Wakefield (2006) “Ill Health and Retirement in Britain: A Panel Data-based Analysis,” *Journal of Health Economics* 25:621-649.

Dorfman, R. (1969) “An Economic Interpretation of Optimal Control Theory,” *The American Economic Review* 59(5): 817-831.

Duan, N., W.G. Manning, Jr., C.N. Morris, and J.P. Newhouse (1983) “A Comparison of Alternative Models for the Demand for Medical Care,” *Journal of Business & Economic Statistics* 1(2):115 – 126.

Ehrlich, I. (2000) “Uncertain Lifetime, Life Protection, and the Value of Life Saving,” *Journal of Health Economics* 19: 341-367.

Ehrlich, I. and G.S. Becker (1972) “Market Insurance, Self Insurance, and Self Protection,” *The Journal of Political Economy* 80:623-648.

Ehrlich, I. and H. Chuma (1990) “A Model of the Demand for Longevity and the Value of Life Extensions,” *The Journal of Political Economy* 98:761-782.

Ehrlich, I. and Y. Yin (2005) “Explaining Diversities in Age-Specific Life Expectancies and Values of Life Saving: A Numerical Analysis,” *The Journal of Risk and Uncertainty* 31:2; 129-162.

Eichner, M.J. (1998) “The Demand for Medical Care: What People Pay Does Matter,” *The American Economic Review* 88(2):117 - 121

Eisenring, C. (1999) "Comparative Dynamics in a Health Investment Model," *Journal of Health Economics* 18:655-660.

Ellis, R. (1986) "Rational Behavior in the Presence of Coverage Ceilings and Deductibles," *The Rand Journal of Economics*, 17: 158-175.

Erbsland, M., W. Ried, and V. Ulrich (1995) "Health, Health Care and the Environment: Econometric Evidence from German Micro Data," *Health Economics* 4:169-182.

Erickson, P., Wilson, R., and Ildy, S. (1995) "Years of Healthy Life" *Health People 2000 Statistical Notes* #7. Centers for Disease Control and Prevention and National Center for Health Statistics.

Evans, W.N., and W.K. Viscusi (1991) "Estimation of State-Dependent Utility Functions Using Suvey Data." *Review of Economics and Statistics* 73(1):94-104.

Finkelstein, A., E.F.P. Luttmer, E.F.P., and M.J. Notowidigdo (2008) "What Good is Wealth Without Health? The Effect of Health on the Marginal Utility of Consumption," Working paper presented at the 2009 AEA meetings.

Frederick, S., G. Lowenstein, and T. O'Donoghue (2002) "Time Discounting and Time Preference: A Critical Review," *Journal of Economic Literature* 40: 351-401.

Frisch, R. (1959) "A Complete Scheme for Computing All Direct and Cross Demand Elasticities in a Model with Many Sectors," *Econometrica* 27:2:177-196.

Gerdtham, U.G., M. Johannesson, L. Lundberg, and D. Isacson (1999) "The Demand for Health: Results From New Measures of Health Capital," *European Journal of Political Economy* 15:501-521.

Gilleskie, D. (1998) "A Dynamic Stochastic Model of Medical Care Use and Work Absence," *Econometrica* 66:1:1 – 45.

Gilleskie, D. (2008) "Health Capital: Theory and Empirical Evidence," in F.A. Sloan and H. Kasper, ed. *Incentives and Choice in Health Care* MIT Press: Cambridge, MA.

Gjerde, J., S. Grepperud, S. and K. Snorre (2005) "On Adaptation and the Demand for Health," *Applied Economics* 37:1283-1301.

Goodman, D.C., Fisher, E., Stukel, T.A., and Chang, C. (1997) "The Distance to Community Medical Care and the Likelihood of Hospitalization: Is Closer Always Better?" *American Journal of Public Health* 87:7:1144 – 1150.

Gould, M.I., and K. Jones (1996) "Analyzing Perceived Limiting Long-term Illness Using UK Census Microdata," *Social Science and Medicine* 42:857-869.

Greenacre, M. (2001) "Tying Up the Loose Ends in Simple Correspondence Analysis," Available at SSRN: http://ssrn.com/abstract=1001889

Greenacre, M. (2002) "Correspondence Analysis of the Spanish National Health Survey," *Gaceta Sanitari* 16(2):160 – 170.

Greene, W. (2003) *Econometric Analysis*, 5th Ed. Prentice Hall: Englewood Cliffs, NJ.

Greene, W. (2004) "The Behavior of the Maximum Likelihood Estimator of Limited Dependent Variable Models in the Presence of Fixed Effects," *Econometrics Journal* 7:98-119.

Greene, W. (2007) "Functional Form and Heterogeneity in Models for Count Data," Working Paper, Stern School of Business, New York University: NY.

Groot, W. (2000) "Adaptation and Scale of Reference Bias in Self-assessments of Quality of Life," *Journal of Health Economics* 19:403-420.

Grossman, M. (1972a) "On the Concept of Health Capital and the Demand for Health," *Journal of Political Economy* 80:223-255.

Grossman, M. (1972b) "The Demand for Health: A Theoretical and Empirical Investigation," Occasional Paper #119, National Bureau of Economic Research.

Grossman, M. (1999) "The Human Capital Model of the Demand for Health," Working Paper #7078, National Bureau of Economic Research.

Hadley, J., and T. Weidman (2006) "Health Insurance and Health at Age 65: Implications for Medical Care Spending on New Medicare Beneficiaries," *Health Services Research* 41(2)429-451.

Hanemann, W. M. (1991) "Willingness to Pay and Willingness to Accept: How Much Can They Differ?" *The American Economic Review* 81:635-647.

Harrington, W., and P. Portney (1987) "Valuing the Benefits of Health and Safety Regulation," *Journal of Urban Economics,* 22:101-112.

Hausman, J., B. Hall and Z. Griliches (1984) "Economic Models for Count Data with an Application to the Patents – R&D Relationship." *Econometrica* 52:909-938.

Heckman, J. (1976) "A Life-Cycle Model of Earnings, Learning, and Consumption," *Journal of Political Economy* 84(S4):S11-S44.

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

Himmerlstein, D.U., E. Warren, D. Thorne, and S. Woolhandler (2005) "Market Watch: Illness and Injury as Contributors to Bankruptcy," Health Affairs Health Tracking Market Watch: http://content.healthaffairs.org/cgi/content/full/hlthaff.w5.63/DC1

Hunt-McCool, J., B.F. Kiker and Y. Ng (1994) "Estimates of the Demand for Medical Care Under Different Functional Forms," *Journal of Applied Econometrics* 9(2):201-218.

Idler E.L (1992) "Self-assessed Health and Mortality: A Review of the Studies," in: Maes, S., Leventhal S. and M. Johnston Eds. *International Review of Health Psychology*. John Wiley & Sons Ltd: New York.

Idler, E.L. and S.V. Kasl (1995) "Self-ratings of Health: Do They Also Predict Change in Functional Ability?" *Journal of Gerontology* 50B:S344-S353.

Idler, E.L. and Y. Benyamini (1997) "Self-rated Health and Mortality: A Review of Twenty-seven Community Studies," *Journal of Health and Social Behavior* 38(1):21-37.

Janssen, R. (1992) "Time Prices and the Demand for GP Services," *Social Science & Medicine* 34(7):725 – 733.

Jenkins, S.P., J.A. Rigg and F. Devicienti (2001) "The Dynamics of Poverty in Britain," Research Report No. 157. Department for Work and Pensions http://www.dwp.gov.uk/asd/asd5/rport157/Main.pdf

Jones, A. and O. O'Donnell (2002) *Econometric Analysis of Health Data*. John Wiley & Sons, LTD. West Sussex, UK.

Kaiser Family Foundation (2007) Trends in Health Care Costs and Spending http://www.kff.org/insurance/upload/7692.pdf

Kamien, M.I. and N.L. Schwartz (1991) *Dynamic Optimization*, Elsevier North Holland, Amsterdam.

Keane, M.P. (2005) "Structural vs. Atheoretic Approaches to Econometrics," Keynote Address at the Duke Conference on Structural Models in Labor, Aging and Health, September 17, 2005.

Keeler, E.B., J.P. Newhouse and C.E. Phelps (1977) "Deductibles and the Demand for Medical Care Services: The Theory of a Consumer Facing a Variable Price Schedule Under Uncertainty," *Econometrica* 45(3):641-656.

Kiecolt-Glaser, J.K. and T.L. Newton (2001) "Marriage and Health: His and Hers," *Psychological Bulletin* 127(4):472-503.

Koc, C. (2004a) "The Effects of Uncertainty on the Demand for Health Insurance," *The Journal of Risk and Insurance* 71:41-61.

Koc, C. (2004b) "A Theoretical Rationale for an Inelastic Demand for Health Care," *Economics Letters* 82:9–14.

Koc, C. (2005) “Health-Specific Moral Hazard Effects,” *Southern Economic Journal* 72(1): 98-118.

Kolata, G. (2006) “So Big and Healthy Grandpa Wouldn’t Even Know You,” *The New York Times* accessed at: http://www.nytimes.com/2006/07/30/health/30age.html on 1/8/09.

Lancaster, T. (2000) “The Incidental Parameters Problem Since 1948.” *Journal of Econometrics* 95(2):391-414.

Landefeld, J.S., and E.P. Seskin, (1982) “The Economic Value of Life: Linking Theory to Practice,” *American Journal of Public Health* 72: 555-565.

Liljas, B. (1998) “The Demand for Health with Uncertainty and Insurance,” *Journal of Health Economics* 17:153-170.

Lillard. L.A., and Y. Weiss, (1997) “Uncertain Health and Survival: Effects of End-of-Life Consumption,” *Journal of Business and Economic Statistics* 15(2):254-268.

Linnerooth, Joanne (1979) “The Value of Human Life: A Review of the Models,” *Economic Inquiry* 17: 52-74.

MaCurdy, T. (1981) “An Empirical Model of Labor Supply in a Life-Cycle Setting,” *Journal of Political Economy* 89(6):1059-1085.

Manning, W., J.P. Newhouse and J.E. Ware, Jr. (1982) “The Status of Health in Demand Estimation: Beyond Excellent, Good, Fair, and Poor,” in *Economic Aspects of Health*, ed. Victor R. Fuchs, The University of Chicago Press: Chicago.

Manning, W., Newhouse, J., Duan, N., Keeler, E., Leibowitz, A., and Marquis, M. (1987) “Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment,” *American Economic Review*, 77:251-277.

Manning et. al. (1988) *Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment* Publication #R-3476-HHS Rand: Santa Monica, CA.

McGinley, L (2006) “Senior Operator: A Johns Hopkins Physician is Giving ‘Geriatric Surgery’ a Big Push,” *The Wall Street Journal*, October 6.

Mishan, E.J. (1971) “Evaluation of Life and Limb: A Theoretical Approach,” *The Journal of Political Economy* 79:687-05.

Mullahy, J., and P. Portney (1990): “Air Pollution, Cigarette Smoking, and the Production of Respiratory Health,” *Journal of Health Economics* 9:193-205.

Murphy, K.M. and R.H. Topel (2006) “The Value of Health and Longevity,” *The Journal of Political Economy*, 114(5):871–904.

Muurien, J. (1982) “Demand for Health: A Generalized Grossman Model,” *Journal of Health Economics* 1:5-28.

Newhouse, J.P. (1978) “Insurance Benefits, Out-of-Pocket Payments, and the Demand for Medical Care: A Review of the Literature,” RAND: Santa Monica, CA.

Newhouse, J.P. (1992) “Medical Care Costs: How Much Welfare Loss?” *Journal of Economic Perspectives* 6(3):3 – 21.

Newhouse, J.P.(1993), *Free for All?: Lessons from the RAND Health Insurance Experiment*. RAND: Cambridge, MA.

Newhouse, J.P., C.E. Phelps and M.S. Marquis (1980) “On Having Your Cake and Eating It Too: Econometric Problems in Estimating the Demand for Health Services,” *Journal of Econometrics* 13:365-390.

Neyman, J. and E. Scott (1948) “Consistent Estimates Based on Partially Consistent Observations,” *Econometrica* 16:1 – 32.

Nuland, S.B. (1995) *How We Die: Reflections on Life’s Final Chapter* Vintage Books: A Division of Random House: New York.

Nyman, J.A. (2003) *The Theory of Demand for Health Insurance* Stamford University Press: Stamford, CA.

O’Donnell, O., and C. Propper (1991) “Equity and the Distribution of UK National Health Service Resources,” *Journal of Health Economics* 49:634-641.

Oniki, H. (1973) “Comparative Dynamics (Sensitivity Analysis) in Optimal Control Theory” *Journal of Economic Theory* 6:265-283.

Pauly, M.V. (1968) “The Economics of Moral Hazard: Comment,” *American Economic Review* 58:531-537.

Petty W (1699) *Political Arithmetick, or a Discourse Concerning the Extent and Value of Lands, People, Buildings.* Robert Caluel: London.

Phelps, C.E., and J.P. Newhouse (1974) “Coinsurance, the Price of Time, and the Demand for Medical Services,” *The Review of Economics and Statistics* 56(3):334-342.

Phlips, L. (1990) *Applied Consumption Analysis, 2nd Edition*. North Holland: New York.

Pohlmeier, W., and V. Ulrich (1995) “An Econometric Model of the Two-Part Decision-making Process in the Demand for Health Care,” *The Journal of Human Resources* 30(2) 339-361.

Propper, C. (2000) "The Demand for Private Health Care in the UK," *Journal of Health Economics*. 19: 855-876.

Reinhardt, U.E. (1985) "The Theory of Physician Induced Demand: Reflections After a Decade," *Journal of Health Economics* 4(2):187-193..

Reinhardt, U.E., P.S. Hussey, and G.F. Anderson (2004) "U.S. Healthcare Spending in an International Context," *Health Affairs* 23(3):10-25.

Remler, D.K., and A.J. Atherly (2003) "Health Status and Heterogeneity of Cost Sharing Responsiveness: How Do Sick People Respond to Cost-sharing?" *Health Economics* 12:269-280.

Ried, M. (1998), "Comparative Dynamic Analysis of the Full Grossman Model," *Journal of Health Economics* 17:383-426.

Ross, C.E., J. Mirowsky and K. Goldsteen (1990) "The Impact of the Family on Health: The Decade in Review," *Journal of Marriage and the Family* 52:1059 – 1078.

Rust, J. and C. Phelan (1997) "How Social Security and Medicare Affect Retirement Behavior in a World of Incomplete Markets," *Econometrica* 55:999-1035.

Sackett, D. L. and G. W. Torrance (1978) "The Utility of Different Health States as Perceived by the General Public," *Journal of Chronic Diseases* 31: 697–704.

Schelling, T.C. (1968) , "The Life Your Save May Be Your Own," in *Problems in Public Expenditure*, S.B. Chase Jr., ed. Brookings Institute: New York.

Senior, M.L. (1998) "Area Variations in Self-Perceived Limiting Long-term Illness in Britain, 1992: Is the Welsh Experience Exceptional?" *Regional Studies* 32(3) 265-280.

Shlens, J. (2005) "A Tutorial on Principal Component Analysis, Version 2" http://www.snl.salk.edu/~shlens/pub/notes/pca.pdf

Sloan, F.A., W.K. Viscusi, H.W. Chesson, C.J. Conover, and K. Whetten-Goldstein, (1998) "Alternative Approaches to Valuing Intangible Health Losses: The Evidence for Multiple Sclerosis." *Journal of Health Economics* 7(4):475-497.

Smith, R.D. (2002) "Construction of the Contingent Valuation Market in Health Care: A Critical Assessment," *Health Economics* 12:8:609-628.

Steckel, R.H. (2008) "Biological Measures of the Standard of Living," *Journal of Economic Perspectives* 22(1): 129 – 152.

Sterns, S.C., and E.C. Norton (2004) "Time to Include Time to Death? The Future of Health Care Expenditure Predictions," *Health Economics* 13:315-327.

Sutton, M., R. Carr-Hill, H. Gravelle, and N. Rice (1999) "Do Measures of Self-reported Morbidity Bias the Estimation of the Determinants of Health Care Utilization?" *Social Science & Medicine* 49:867-878.

Taylor, M. F. (ed). with J. Brice, N. Buck and E. Prentice-Lane (2006) *British Household Panel Survey User Manual Volume A: Introduction, Technical Report and Appendices.* University of Essex: Colchester, UK.

Tolley, G., D. Kenkel, and R.Fabian, ed., (1994) *Valuing Health for Policy: An Economic Approach*, The University of Chicago Press: Chicago.

Van Doorslaer, E.K.A. (1987) *Health Knowledge and the Demand for Medical Care* Maastricht: Wolfeboro NH.

Van Doorslaer, E.K.A., A. Wagstaff, H. van der Burg, T. Christiansen, D. De Graeve, I. Duchesne, U-G. Gerdtham, M. Gerfin, J. Geurts, L. Gross, U. Häkkinen, J. John, J. Klavus, R.E. Leu, B. Nolan, O. O'Donnell, C. Propper, F. Puffer, M. Schellhorn, G. Sundberg and O. Winkelhake (2000) "Equity in the Delivery of Health Care in Europe and the U.S.," *Journal of Health Economics* 19:553-583.

Van Doorslaer, E.K.A., Jones A.M., Koolman, X. (2002) "Explaining Income-related Inequalities in Doctor Utilization in Europe: A Decomposition Approach. *ECuity II Project*, Working Paper #5.

Van Doorslaer, E.K.A., and U.G. Gerdtham (2003) "Does Inequality in Self-Assessed Health Predict Inequality in Survival by Income? Evidence from Swedish Data," *Social Science and Medicine* 57(9): 1621-1629.

Viscusi, W., and W. Evans (1990) "Utility Functions that Depend on Health Status: Estimates and Economic Implications," *American Economic Review* 80:353-374.

Wagstaff, A. (1986) "The Demand for Health: Some New Empirical Evidence," *Journal of Health Economics* 5:195-233.

Wagstaff, A. (1993), "The Demand for Health: An Empirical Reformulation of the Grossman Model," *Health Economics* 2:189-198.

Weisbrod, B.A. (1991) "The Health Care Quadrilemma: An Essay on Technological Change, Insurance, Quality of Care, and Cost Containment," *Journal of Economic Literature* 29:523-552.

Williams, K., and D. Umberson (2004) "Marital Status, Marital Transitions, and Health: A Gendered Life Course Perspective," *Journal of Health and Social Behavior* 45(1):81-98.

Wilson, C.M., and A.J. Oswald (2005) "How Does Marriage Affect Physical and Psychological Health? A Survey of the Longitudinal Evidence," IZA Discussion Paper No. 1619 http://ssrn.com/abstract=735205.

Wooldridge, J.M. (2005) "Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models with Unobserved Heterogeneity," *Journal of Applied Econometrics* 20:39-54.

Yang, Z., E.C. Norton and S.C. Sterns (2003) "Longevity and Health Care Expenditures: The Real Reasons Older People Spend More," *Journal of Gerontology* 58B(1):S2-S10.

Yi, W.W., and T.M. Ezzati-Rice (2005) "Concentration of Health Care Expenditures in the U.S. Civilian Non-institutionalized Population," *Statistical Brief #81* Agency for Health Care Research and Quality: Rockville, MD. http://www.meps.ahrq.gov/mepsweb/data_files/publications/st81/stat81.pdf

Zhang, J., P.P. Vitaliano, and H. Lin (2006) "Relations of Caregiving Stress and Health Depend on the Health Indicators Used and Gender," *International Journal of Behavioral Medicine* 13(2): 173 – 181.

Zweifel, P., and F. Breyer (1997) *Health Economics* Oxford University Press: New York.

Zweifel, P., S. Felder and M. Meiers (1999) "Ageing of Population and Health Care Expenditure: A Red Herring?" *Health Economics* 8(6):485-496.

Zweifel, P., and W.G. Manning (2000) "Moral Hazard and Consumer Incentives in Health Care," in *The Handbook of Health Economics*, J.A. Culyer and J.P. Newhouse, editors. Elsevier: Amsterdam.

VITA

Jennifer Leigh Kohn

1968	Born December 7 in Stamford, Connecticut.
1986	Graduated from Andrew Warde High School, Fairfield, Connecticut.
1986 – 1989	Attended the University of Virginia, Charlottesville, VA.; Echols Scholar.
1989 – 1990	Attended The University of Massachusetts at Amherst; majored in Social Thought and Political Economy.
1990	B.A. The University of Massachusetts at Amherst.
1990 – 1992	Consumer advocate for the New Jersey Public Interest Research Group.
1992 – 1994	Policy Analyst for the New York City Department of Consumer Affairs.
1994 – 1996	Director of Consumer Policy for the New York City Office of the Public Advocate.
1996 – 1998	Attended New York University Stern School of Business, New York, NY.; majored in Finance and Statistics.
1998	M.B.A., New York University Stern School of Business.
1997 – 2000	Senior Administrator for the Division of Cardiology, New York Presbyterian Hospital – Weil Cornell University Medical College.
2000	Son Max born September 27
2001 – 2005	Adjunct professor for business statistics and operations management, Montclair State University School of Business, Montclair, NJ.
2003	Son Sam born March 21
2005 – 2009	Attended Rutgers University School of Business, Ph.D. program in Management, emphasis in Finance and Economics, RA/TA.
2009	Ph.D. in Management, Rutgers University