Working Paper Series
Document de travail de la série

MAKING THE GROSSMAN MODEL STOCHASTIC: INVESTMENT IN HEALTH AS A STOCHASTIC CONTROL PROBLEM

Audrey Laporte, Brian Ferguson

Working Paper No: 170009

www.canadiancentreforhealtheconomics.ca

July 19, 2017

Canadian Centre for Health Economics
Centre canadien en économie de la santé
155 College Street
Toronto, Ontario

CCHE/CCES Working Paper No. 170009
July 19, 2017

Making the Grossman Model Stochastic: Investment in Health as a Stochastic Control Problem

Audrey Laporte[a,b], Brian Ferguson[b,c]

[a] *Institute of Health Policy, Management and Evaluation, University of Toronto*
[b] *Canadian Centre for Health Economics, Toronto, Canada*
[c] *Department of Economics and Finance, University of Guelph*

Abstract

It is well known that uncertainty is a key consideration in theoretical health economics analysis. The literature has shown that uncertainty is a multifaceted concept, with the individual's optimal response depending on the formal nature of the uncertainty and the time horizon involved. This paper extends the literature by considering uncertainty with regards to the cumulative effect on health capital of on-going health behaviours. It uses techniques of stochastic optimal control to analyze uncertainty which can be represented as a Weiner process and shows how, in a Grossman health investment framework, the optimal lifetime health investment trajectory might be affected.

JEL Classification: I1; I12

I. Introduction

Uncertainty has been a key element of theoretical health economics since Arrow's (1963) paper on health insurance. As Grossman's (1972) health capital framework came to prominence, the angle through which uncertainty was viewed shifted slightly, adding a production and investment perspective to the original insurance perspective. Whichever perspective a particular case happened to call for, though, one thing which was clear from the early days of the analysis of uncertainty in health economics was that the term uncertainty meant different things in different contexts, and that even when the literature referred to basically the same type of uncertainty, the context, or the structure of the model within which the uncertainty operated, played a significant role in determining the consequences of uncertainty.

Dardanoni and Wagstaff (1990) classified uncertainty as type 1 or type 2 depending on whether, as in the former case, it referred to uncertainty about the individual's pre-treatment state of health or, as in the latter case, it referred to uncertainty about the effectiveness of medical care. They also formalized the characterization of uncertainty, focusing on the Rothschild-Stiglitz (1970) concept of increasing or decreasing risk through mean preserving spreads or contractions, and also discussing combinations of changes in the mean and the spread of the distribution of the health variable. Their analysis set out the factors which affect the individual's optimal response to such changes as first order stochastically dominating shifts of the distribution of the random health variable, and points to the fact that third derivatives of the individual's utility function, U(C,H), where H is health and C is consumption of non-health related commodities, play a key role in determining the individual's response to changes in health risk. Following Dardanoni (1988), they interpret these third derivatives in terms of a two-good index of absolute risk aversion.

Dardanoni and Wagstaff introduced uncertainty into a one-period Grossman-type model. One point which the later literature has made clear is that the length of the individual's optimal planning horizon is critical to her response to uncertainty. Thus Eeckhoudt and Gollier (2005), working with a one period model and using Kimball's (1990) utility-based measure of prudence, which involves a third derivative of the utility function, find that greater prudence tends to be negatively associated with optimal preventive effort. Menegatti (2009), however, shows that the Eeckhodt and Gollier result depends on the one-period nature of the model and is reversed when you shift to a two-period framework, and, in a non-health context, Huggett and Vidon (2002) show that the results of a two-period model can be changed completely when the time horizon is extended to three periods. Clearly the answer to the question about the effect of uncertainty in a Grossman-type model depends crucially on the context surrounding the question.

With regards to the issue of time horizon, the Grossman model is often set up as a continuous-time optimal control problem, either with an infinite or a finite horizon. Cropper (1977) introduces uncertainty into an optimal control version of the health investment model. Here the nature of the uncertainty is a critical element: Cropper assumes that shocks are short term and mean reverting. She defines two possible outcomes – healthy and unhealthy – and makes a clear distinction between health capital and illness (Cropper (1977) pg. 1274). Her focus is on illnesses which do not have a permanent effect on the individual's health capital. Health Capital can be thought of as the ability of the body to resist disease: in Cropper's model it is not the individual's stock of health capital which is uncertain but rather her level of exposure to germs on any given day, meaning that on any given day, if we compare two individuals who have the same level of health capital, one might be well and the other sick, simply because of their degree of short term exposure to germs and viruses. The individual's stock of health capital defines her degree of resistance to germs: if on any given day the random variable which is her level of exposure to illness takes on a value which exceeds her degree of resistance to illness, she is classified as sick and her level of utility is set to zero for that instant. If her exposure is less than her resistance level she is classified as healthy and her utility at that instant is determined by a U(C,H) type utility function. Thus, although her expected degree of exposure to illness might prompt her to invest more heavily in H, raising her resistance level, the fact of being ill at one instant does not affect her level of health capital (or, indeed, her level of exposure) in the next instant.

Liljas (1998) extends Cropper's model by allowing for a continuum of health states rather than a simple healthy/sick dichotomy. Like Cropper he works in a continuous time framework, although because he sets the problem up as an isoperimetric problem he does not have variables representing the shadow price of health capital at each instant (co-state variables) and does not use a phase diagram representation. Liljas expresses his results in terms of Euler equations which define the optimal trajectory for the individual's stock of health capital. This approach highlights the difference between the Euler equations for health capital in the non-stochastic and extended-Cropper versions of the model, but has limitations when it comes to giving a sense of how the individual's optimal trajectory changes over the long run. Liljas's results also depend on the third derivative of the utility function, although he uses this primarily in a proof.

As noted above, Cropper-type intertemporal optimization models introduce uncertainty in a manner which affects the likelihood that the individual will be sick at any given point in time, and as a consequence have her instantaneous utility reduced, but which does not directly affect her stock of health capital in the future. In these models, the effect of illness on health capital operates through the individual's health investment decisions, in that awareness that she lives in an area where her exposure to germs is more likely to exceed her degree of resistance might prompt her to invest more in H. Thus Cropper-type models deal with minor illnesses, and investment in health can be thought of as a form of preventive care. Curative care, on the other hand, would be spending on health-related goods in response to an illness which did reduce H permanently and would only occur after a major illness had struck. In a Grossman framework we can still think in terms of the possibility of preventive health investment which reduces the likelihood of a falling victim to a major illness, the eventuality of which illness is a single realization of a stochastic health shock variable, but we must also recognize that once a major illness has struck and the individuals' stock of health capital has been permanently reduced, she must re-plan her post-illness health investment trajectory.

This type of major illness was studied in a simulation framework by Picone, Uribe and Wilson (1998), and modelled as a stochastic optimal control problem incorporating Poisson-type shocks to Health capital by Laporte and Ferguson (2007). Picone et al. (1998) simulate the optimal health investment trajectory for an individual who knows that at each instant in time there is a probability that she will suffer a major illness, but who does not actually do so. Laporte and Ferguson (2007), using a generalized Ito's lemma approach, incorporate Poisson-type shocks (in which there is a certain probability that the individual will be struck by a major illness at each instant in time but in which, as in most periods, if the illness does not strike, her health is unaffected) into the individual's lifetime optimal health investment problem and show how the phase diagram for the theoretical problem is modified, and also use the phase diagram approach to illustrate, qualitatively, how the occurrence of a major illness will cause her to alter her future investment trajectory.

Thus Cropper-Lijlas type shocks are experienced every day but are not cumulative in health capital while Poisson type shocks are experienced only rarely but, should they occur (which they might not) do have a permanent effect on one's stock of health capital. There is another form of stochastic element in health which in a sense combines aspects of the first two. These are cases where health behaviours which occur every day have a small impact on one's stock of health each day but potentially have a large cumulative effect over time, the magnitude of which is uncertain.

Consider for example, soft drinks and potato chips. An individual who consumes these items regularly knows that they will be bad for her in the long run but does not see a significant effect from day to day and also does not how bad they will be for her in particular in the long run. What she knows is in a sense is population level data-she has a sense of the population average impact on health of the daily consumption of soft drinks but she also knows that some people are more affected by them than others, and she does not know in advance where her particular metabolism will place her within the distribution of health damage. Thus we have a case where we know that regular daily consumption of a particular commodity will cause a downward trend in individual health capital and the spread around the trend gets wider the longer the process goes on, reflecting in a sense unobservable heterogeneity in metabolism. The same kind of uncertainty could surround positive health behaviours like eating healthy and exercising in the sense that again the individual may know the population average effect but does not know whether her eventual particular benefit will be at the high or low end of the distribution of benefit around the population mean. This is the type of health investment which we will be considering in this paper.

For our purposes the key feature of this sort of uncertainty is that not only does H cumulate over time as a result of the health investment process, but so too does uncertainty about H. The individual's expectation about the effects of her ongoing health investment activities are based on her reading of health-related articles in newspapers, but because she does not know exactly how her own physiology will respond to the investment activity, each new unit of health investment has its own zero-mean uncertainty element attached to it. While the variance of the instantaneous stochastic elements are unchanged from period to period, over time her uncertainty about the outcome of her activities will cumulate, so that, when she contemplates her lifetime investment trajectory from the point, early in life, when she first makes her health investment plan, the further into the future she looks the greater the uncertainty about what the cumulative effects of those activities will be. Thus as she looks further into the future she essentially sees her health as a random variable with an expectation and with a variance which is increasing as time passes, so long as she continues her activities. Mathematically we will represent this effect by using the techniques of stochastic control, with uncertainty represented by a Wiener process.

In this paper we explore the impact on the optimal health investment trajectory of adding this type of uncertainty to a Grossman model. We begin by setting out a deterministic optimal control version of the Grossman investment in health model. We then set out a stochastic control version incorporating a Weiner process and discuss why we think the Weiner process is the most suitable mathematical form to characterize this third type of uncertainty. Following that we solve the stochastic control problem and use a phase diagram to illustrate the comparative dynamics of the deterministic and stochastic problems in terms of the effect of Wiener type uncertainty on the optimal health investment trajectory. We then discuss the interpretation of the new terms which the stochastic process adds to the equation for the health investment trajectory and place our result in the context of the broader economic literature on choice under uncertainty.

II. Deterministic Version

We begin with a simple one-state variable case of the Grossman model, set within a continuous time optimal control framework. The individual aims to maximize her discounted lifetime utility:

(1) Max $\int_0^T U(C_t, H_t)e^{-\rho t}\,dt$ $U_C > 0$, $U_{CC} < 0$, $U_H > 0$, $U_{HH} < 0$, $U_{CH} \geq 0$

The individual's utility is a function of C, which stands for non-health-related consumption goods, and H, health capital. ρ is the individual's subjective discount rate. The stock of health capital evolves according to:

(2) $\dot{H} = I - \delta H$, $0 < \delta < 1$, $I \geq 0$

In equation (2), I is health investment, and δ is the rate of depreciation of health capital, which we assume is constant, so that when the individual undertakes no health investment activity (I=0) her stock of health capital will decline at a constant rate δ.

The individual is assumed to have an instantaneous budget constraint which must be satisfied each period:

(3) $Y_t = C_t + PI$

Here Y is instantaneous income which we treat as exogenous to the individual. Health investment, I, can be purchased in the market at a price P and the price of non-health related goods C is set to 1, so that income is measured in real consumption terms and P can be thought of as the relative price of health investment goods.

The Hamiltonian for this version of the Grossman problem is:

(4) $\mathcal{H} = U(Y_t - PI, H) + \Psi[I - \delta H]$

where we have used the budget constraint to substitute for C in the utility function. Here, Ψ is the co-state or shadow price of an additional unit of H.

We are interested in depicting the evolution of the individual's I and H over time - in other words in her optimal lifetime health investment trajectory. To solve this problem we begin by finding the first-order condition for the choice variable I, and then take advantage of the fact that the first order condition must hold at all t to generate an equation of motion for I to go with (2), the equation of motion for H. We then define the stationary loci for each of I and H. The stationary loci have the property that when the individual is on it there is no inherent tendency for the variable in question to change, so that along the H locus $\dot{H}$=0 and along the I locus $\dot{I}$=0. The advantage of defining stationary loci is expositional: it lets us draw a phase diagram in (I,H) space and divide the space into regions where I and H have intrinsic tendencies to rise or fall as time passes, which we identify on the phase diagram using phase arrows.[1]

The first order necessary condition for the problem (assuming an interior solution for I) is:

(5) $-pU_C + \Psi = 0$

Which can be written in marginal benefit equals marginal cost form:

(6) $\Psi = PU_C$

The left hand side of (6) Ψ, is the shadow price of another unit of health that comes from investing a unit of I. The right hand side is the marginal cost of an additional unit of I, which is the number of units of C that have to be given up in exchange for an additional unit of I multiplied by the utility value (U_c) of each unit of C.

The other condition which must be satisfied for dynamic optimization is the Pontryagin necessary condition which defines the equation of motion for Ψ:

(7) $\dot{\Psi} = \rho\Psi - \mathcal{H}_H$

$= \rho\Psi - [U_H - \delta\psi]$

$= [\rho + \delta]\psi - U_H$

Both conditions need to be satisfied and although (6) will hold for each t, we see from (7) that ψ changes over time which means that the value of I which causes (6) to hold will also change throughout the individual's life.

To derive an equation of motion for I we differentiate (6) with respect to t to yield:

(8) $P^2 U_{CC}\, \dot{I} - PU_{CH}\dot{H} + \dot{\psi} = 0$

From which we obtain:

(9) $\dot{I} = \frac{PU_{CH}\dot{H} - \dot{\psi}}{[P^2 U_{CC}]}$

The equation of motion for I in (9) tells us how I must change over time as H and ψ change over time in order that the necessary condition for I continues to hold for all t. Further, since (6) and (7) must hold for all t we can substitute for $\dot{H}$ and $\dot{\psi}$ giving:

(10) $\dot{I} = \frac{[U_H] - [\rho+\delta]PU_C + P[U_{CH}][I - \delta H]}{P^2 U_{CC}}$

[1] See Ferguson and Lim (1998) for an outline of the technique to generate and interpret phase diagrams.

Expressions (2) and (10) are used to derive the stationary loci for H and I respectively in the phase diagram depicted in Figure 1.

[FIGURE 1 ABOUT HERE]

Figure 1 shows the stationary loci for I and H and the phase arrows for both variables in the four segments of the diagram. It also shows the equilibrium for the system, at point E, and the stable and unstable branches pointing to and away from E. If ours were an infinite horizon problem, the individual's optimal trajectory would coincide with a stable branch to E. Since our individual is finite-lived, and knows it, her optimal trajectory will in fact be the one labelled A. We have drawn Figure 1 on the assumption that our individual is born healthy – i.e. that her initial level of health capital, H_0, is high. Her optimal trajectory will involve a relatively low initial level of health investment; below, in fact, the level necessary to hold H constant, so her optimal H will decline from its high initial level. Through the first part of her lifetime trajectory I will rise, slowing, but not halting, the decline in H then, when trajectory A cuts the stationary locus for I, her optimal level of health investment will begin to decline. The transversality conditions for the finite horizon version of the Grossman problem require that I reaches zero at or before T, the ending time for the problem[2]. Thus the optimal trajectory for an individual born healthy[3] takes the inverted-U shape of trajectory A in Figure 1. In going over to the stochastic control version of the problem we will use the phase diagram to show how, qualitatively, the addition of a certain type of uncertainty alters the shape of our individual's optimal lifetime health investment trajectory.

III. Stochastic Version

In this section we consider the model

$$(11) \qquad \underset{I}{Max} \; E[\int U(Y - PI, H)e^{-\rho t}dt]$$

Subject to

$$(12) \qquad dH = [I - \delta H]dt + \sigma_H dz_H$$

Here in addition to the standard Grossman notation, H is stochastic, obeying a Weiner process. Equation (12) is the stochastic counterpart of the $\dot{H}$ equation of the deterministic version of the model, and the first element in the dH equation is in fact the RHS of the $\dot{H}$ equation. In (12), that term is referred to as the drift – in this case a controlled drift since H could drift up or down depending on the value of I. In particular, if I = 0, H will drift down, in the same way as H declines over time in the

[2] On the analytical details of the non-stochastic Grossman model see Laporte (2015).

[3] If she were born unhealthy, so that H_0 was well to the left of point E, her initial I would be high, causing H to rise through the first part of her lifetime before beginning to decline so that the transversality condition would hold.

deterministic case. The new element in (12), $\sigma_H dz$, represents the Wiener process which characterizes the uncertainty associated with the evolution of H as time passes.

The Mathematics of a Wiener Process

We noted above that we want to consider the case in which not only is H stochastic but uncertainty about the level of H increases over time. To do this we introduce a Wiener process, or Brownian motion, to the evolution of H. The Wiener process is a continuous stochastic element and adding it essentially means that H is subject to uncertainty at every instant in time. The basic assumptions of a Wiener process are set out nicely by Mangel (1985)[4]. We define a stochastic process, Z(t), where

(i) Sample paths of Z(t) are continuous

(ii) Z(0) = 0, so we know the initial value of our variable with certainty

(iii) The increment $Z(\tau + s) - Z(\tau)$ is normally distributed with mean 0 and variance $\sigma^2 s$, where s is the length of the interval. Thus the variance of the increment in Z depends on, and increases with, the length of the period ahead, over which we are looking.

In continuous time applications of Wiener processes we let s = dt, a small increment in time. Then defining $dZ = Z(t+dt) - Z(t)$, we have

(13) $EdZ = 0,\ E(dZ)^2 = dt,\ EdtdZ = 0$

In (13) the Expectations operator, E, is present because the increment dZ is a normally distributed random variable. The third of the expressions in (13) arises because dZ is on the order of magnitude of the square root of dt. Along with the assumption that $(dt)^2 = 0$, which follows from the assumption that dt is an infinitesimal, (13) constitutes the rules of multiplication of Wiener terms.

Wiener processes are continuous time processes, but because they represent continuous shocks, and therefore a series of continuous, if infinitesimal, jumps, they are not differentiable using ordinary rules of calculus. In essence, this means that dZ/dt does not exist in the usual sense. A Wiener process basically represents a variable whose time path is all corners, or spikes. Thus we have to adopt the Ito Calculus for problems involving them.

In economic applications, we are generally not interested in Wiener processes per se, but in the behaviour of variables which are functions of Wiener processes – i.e. variables whose behaviour over time is subject to the continuous random shocks characterized by Brownian motion. Thus we can define a variable x such that

(14) $dx = \alpha dt + \sigma dz$

where dz represents the Wiener process. In the absence of the dz term we would have $dx = \alpha dt$, and dividing through by dt we would have $dx/dt = \alpha$, or, in the notation used for time derivatives, $\dot{x} = \alpha$. We cannot, however, simply divide through in (14) by dt, as in

(14a) $dx/dt = \alpha + \sigma dz/dt$

[4] See also Ferguson and Lim (1988)

because dz/dt does not exist in the usual sense. We can, however, use the Ito calculus to analyze the effect of the Brownian motion type of uncertainty driving the Wiener process on a variable which is itself a function of x – e.g. on $y = f(x,t)$. In doing so, note that we can take the expectation of (14):

$$(15) \qquad Edx = E\alpha dt + E\sigma dz = E\alpha dt = \alpha dt$$

Since σ, the variance scaling term, is non-stochastic, as are α and dt, and $Edz = 0$. Then we divide through by dt, giving an expression which we write as

$$(16)\ [Edx]/dt = \alpha.$$

Expression (16) refers to the expected instantaneous change in x as time passes, which we refer to as the drift term in x. Note that the drift in the stochastic x is the same as the actual time change in x if $\sigma = 0$, i.e. if x were non-stochastic. It is not, however, necessarily the actual change in x over an infinitesimal interval of time – that will be a combination of the drift plus noise.

In our particular application we have

(17) $EdH = [I - \delta H]dt$, so that $EdH/dt = [I - \delta H]$, the instantaneous drift in H and

$$(18) \qquad Var(dH) = E[dH - EdH]^2 = E[\sigma dz]^2 = \sigma^2 dt$$

Here equation (18) is the variance of the stochastic process which defines the instantaneous change in H. If we take the stochastic integral of our process, obtaining an expression for H as a function of elapsed time, we see why we say that the variance of H increases the further ahead we look.

The key to the use of the Ito calculus in stochastic control problems is that, although time derivatives may not exist in the usual sense, other derivatives do. Ito's approach to analyzing the behaviour of y is to take a second order Taylor Series expansion to $df(x,t)$:

$$(19)\ dy = df(x,t) = f_x(x,t)dx + \tfrac{1}{2} f_{xx}(x,t)(dx)^2 + f_t(x,t)dt + \tfrac{1}{2} f_{tt}(x,t)(dt)^2 + f_{xt}(x,t)dxdt$$

Next replace dx by $\alpha dt + \sigma dz$ everywhere, giving:

$$(20) \qquad dy = f_x(x,t)[\alpha dt + \sigma dz] + \tfrac{1}{2} f_{xx}(x,t)[\alpha dt + \sigma dz]^2 + f_t(x,t)dt + \tfrac{1}{2} f_{tt}(x,t)(dt)^2 + f_{xt}(x,t)[\alpha dt + \sigma dz]dt$$

Multiplying this expression out gives:

$$(21) \qquad dy = f_x\, \alpha dt + f_x\, \sigma dz + \tfrac{1}{2} f_{xx}\alpha^2[dt]^2 + \tfrac{1}{2} f_{xx}\, \sigma^2[dz]^2 + \tfrac{1}{2} f_{xx}\, [2\alpha\sigma\, dt\, dz] + f_t\, dt + \tfrac{1}{2} f_{tt}\, [dt]^2 + f_{xt}\alpha[dt]^2 + f_{xt}\sigma dz dt$$

Next, taking the expectations operator through and applying the rules of multiplication for Wiener processes, plus the fact that $[dt]^2$ is vanishingly small even in the absence of uncertainty, gives:

$$(22) \qquad Edy = f_x\, \alpha dt + \tfrac{1}{2} f_{xx}\, \sigma^2 dt + f_t\, dt$$

From which we obtain:

$$(23) \qquad [Edy]/dt = [f_x \alpha + ½ f_{xx} \sigma^2 + f_t]$$

Note that in the non-stochastic case we would have

$$(24) \qquad dy/dt = f_x \alpha + f_t$$

Where f_t represents any trend element in y which was not inherited from x. The drift term in (23) differs from the time trend in the non-stochastic case, as set out in (24), by the addition of the term $f_{xx} \sigma^2/2$. Assuming that both the non-stochastic and the stochastic examples start from the same point, the difference between the non-stochastic trend and the expected drift depends on the concavity or convexity of f. If $f_x > 0$ and $f_{xx} < 0$, so that f is concave in x, the effect of the Wiener process is to pull the drift down relative to the non-stochastic trend (note that the actual time path of the stochastic y will depend on the actual realizations of the Wiener process). Thus the presence of the Wiener process in x has an impact on the drift of y, but that impact will vary depending on the shape of the functional relation between x and y. In addition, even if α and f_t are zero, so there is no trend in x or in the deterministic equation for y, the drift term in the stochastic equation for y will be

$$(25) \qquad [Edy]/dt = ½ f_{xx} \sigma^2$$

so the presence of the Wiener process will add a drift, positive or negative depending on the sign of f_{xx}, with magnitude depending on the magnitudes of f_{xx} and of σ^2.

The fact that the presence of Wiener-type noise in x can produce drift in y even when there is no drift in x is a Jensen's Inequality effect. The upward and downward shocks to x are assumed to be identically distributed. If we assume that f(x) has the usual concave shape of a utility function, with $f_x > 0$, $f_{xx} < 0$, each upward step in x will produce an upward step in y and there will be the same number of upward as downward steps in y, but if we compare the effect of an upward step in x with that of a downward step of equal magnitude in x, the upward step in y produced by the upward step in x will be smaller than the downward step in y produced by the (equal magnitude) downward step in x. Thus over the long run, when x is hit by an equal number of identically distributed upward and downward shocks, y will have an equal number of upward and downward steps but the downward steps in y will tend to be larger than the upward steps. The result is that even though x will not have any particular drift over the long run (although it might appear to have drift in the short run if for example, it happens to be hit by a string of upward shocks), y will tend to have a downward drift over the long run.

To introduce this sort of uncertainty into our optimal control problem we adopt the approach in Pindyck (1982). This involves starting from what is essentially a dynamic programming approach to the problem and using either the standard calculus or the Ito calculus as appropriate in deriving what is in effect a stochastic Hamiltonian problem[5]. We can use the standard calculus in taking derivatives which do not involve the Wiener process: in effect we could take second order Taylor Series expansions in place of the first order ones which the standard calculus involves, then invoke the rules of multiplication from the Ito calculus to make the second order terms vanish. In effect, the Ito calculus extends the standard calculus to the stochastic case, and if there is no stochastic element in a particular differentiation, the two

[5] For an alternative method of defining a stochastic Hamiltonian equation see Malliaris and Brock (1982).

approaches give the same result. Broadly speaking, Pindyck's approach has two defining characteristics. One is that it replicates the steps used in deterministic optimal control analysis, giving a stochastic equation of motion for the addictive commodity, or drift term, for the control variable, which can be compared with the equation of motion derived from the deterministic case. The other is that, unlike the dynamic programming approach, it eliminates direct reference to the value function and therefore does not require an assumption as to a functional form for the value function. General assumptions still need to be made about the functional form of the instantaneous objective function, but we can, in a sense, focus on those rather than making them fit an explicitly solvable form for the value function.

IV. Solving the Stochastic Control Problem

We proceed by denoting expression (11) as J(H,t) and write, in dynamic programming form,

(26) $J(H,t) = \mathrm{Max}_I\, E[U(Y - PI, H)e^{-\rho t}dt + J(H+dH, t + dt)]$

$= \mathrm{Max}_I\, E[U(Y - PI, H)e^{-\rho t}dt + J(H, t) + J_H dH + ½\, J_{HH}(dH)^2 + J_t dt]$

Substituting and carrying the E operator through we have

(27) $0 = \mathrm{Max}_I\, [U(Y - pI, H)e^{-\rho t}dt + J_H[I - \delta H]dt + ½\, J_{HH}\, \sigma_H^2 dt + J_t dt]$

From which we can eliminate the dt terms:

(28) $0 = \mathrm{Max}_I\, [U(Y - PI, H)e^{-\rho t} + J_H[I - \delta H] + ½\, J_{HH}\, \sigma_H^2 + J_t]$

From (28) we find the FOC for I:

(29) $-PU_C\, e^{-\rho t} + J_H = 0$

Which we write as

(30) $J_H = PU_C\, e^{-\rho t}$

Taking the Ito derivative of (30) we have

(31) $\frac{1}{dt}\, E\, dJ_H = -\rho P U_C e^{-\rho t} + P e^{-\rho t} \frac{1}{dt} E dU_C$

which we will expand later.

Differentiating (28) with respect to H and invoking the envelope theorem gives

(32) $\frac{1}{dt} E dJ_H = \delta J_H - [U_H] e^{-\rho t}$

where we have used the fact that $J_{HH}[1 - \delta H] + ½J_{HHH}\sigma_H^2 + J_{Ht}$ can be written as $[EdJ_H]/dt$ to give us the left hand side of (32).

Substituting from (30) for J_H in (32) gives

(33) $\frac{1}{dt}EdJ_H = \delta PU_C e^{-\rho t} - U_H e^{-\rho t}$

Then equating (31) and (33) gives

(34) $-\rho PU_C e^{-\rho t} + Pe^{-\rho t}\frac{1}{dt}EdU_C = \delta PU_C e^{-\rho t} - U_H e^{-\rho t}$

Which can be re-written as

(35) $Pe^{-\rho t}\frac{1}{dt}EdU_C = [\rho + \delta]PU_C e^{-\rho t} - U_H e^{-\rho t}$

From which we can eliminate the $e^{-\rho t}$ terms:

(36) $P\frac{1}{dt}EdU_C = [\rho + \delta]PU_C - U_H$

In order to evaluate the left hand side of (36) we need to take a second order Taylor series expansion of $dU_C(Y - PI, H)$. To do this, let $F(I, H) = U_C(Y - PI, H)$. Then:

(37) $dF(I,H) = F_I dI + ½ F_{II}(dI)^2 + F_{IH}dIdH + F_H dH + ½ F_{HH}(dH)^2$

Looking at the component elements of dF we have

(38a) $F_I = -PU_{CC}$

(38b) $F_{II} = P^2U_{CCC}$

(38c) $F_{IH} = -PU_{CCH}$

(38d) $F_H = U_{CH}$

(38e) $F_{HH} = U_{CHH}$

Substituting into (37) we have

(39)

$$\frac{1}{dt}EdU_C =$$

$$-PU_{CC}\frac{1}{dt}EdI + \frac{P^2U_{CCC}}{2}\frac{1}{dt}E[dI]^2 - [PU_{CCH}]\frac{1}{dt}EdIdH + [U_{CH}]\frac{1}{dt}EdH + \frac{[U_{CHH}]}{2}\frac{1}{dt}\ E[dH]^2$$

Making the substitutions from the Ito rules for multiplication, we have:

(40) $\frac{1}{dt}EdU_C = -PU_{CC}\frac{1}{dt}EdI + \frac{P^2U_{CCC}}{2}\frac{1}{dt}E[dI]^2 - [PU_{CCH}]\frac{1}{dt}EdIdH + [U_{CH}][I - \delta H] + \frac{[U_{CHH}]}{2}\sigma_H^2$

Which from (36) must equal

(41) $\frac{[\rho+\delta]PU_C-[U_H]}{P}$

Thus we have

(42) $-PU_{CC}\frac{1}{dt}EdI + \frac{P^2U_{CCC}}{2}\frac{1}{dt}E[dI]^2 - [PU_{CCH}]\frac{1}{dt}EdIdH = \frac{[\rho+\delta]PU_C-[U_H]- P[U_{CH}][I-\delta H]}{P} - \frac{[U_{CHH}]}{2}\sigma_H^2$

Where we have isolated the remaining dI terms on the left hand side. Now write I as a function of H: I(H), and find the appropriate Taylor series expansions:

(43) $dI = I_H dH + ½ I_{HH}[dH]^2$

From which we find that

(44a) $[dI]^2 = I_H^2[dH]^2 = I_H^2\sigma_H^2\, dt$

and

(44b) $dIdH = I_H\sigma_H^2 dt$

Which lets us write

(45)

$$-PU_{CC}\frac{1}{dt}EdI = \frac{[\rho+\delta]PU_C-[U_H]- P[U_{CH}][I-\delta H]}{P} - \frac{[U_{CHH}]}{2}\sigma_H^2 - \frac{P^2U_{CCC}}{2}[I_H^2\sigma_H^2] + [PU_{CCH}]I_H\sigma_H^2$$

And, dividing through by $-PU_{CC}$:

(46) $\frac{1}{dt}EdI = \frac{[U_H]-[\rho+\delta]PU_C+P[U_{CH}][I-\delta H]}{P^2U_{CC}} + \frac{[U_{CHH}+]}{2PU_{CC}}\sigma_H^2 + \frac{P^2U_{CCC}}{2PU_{CC}}[I_H^2\sigma_H^2] - \frac{[U_{CCH}]}{U_{CC}}\ I_H\sigma_H^2$

which we rearrange as:

(47) $\frac{1}{dt}EdI = \frac{[U_H\ ]-[\rho+\delta]PU_C+P[U_{CH}][I-\delta H]}{P^2U_{CC}} + \frac{[U_{CHH}-2PU_{CCH}I_H]}{2PU_{CC}}\sigma_H^2 + \frac{U_{CCC}}{2PU_{CC}}\sigma_H^2[PI_H^2]$

V. Establishing the optimal trajectory under uncertainty.

The first element on the right hand side of (47) is the İ expression from the non-stochastic Grossman problem. This will give us a starting point for interpreting our results, both in terms of equation (47) and in terms of the phase diagram for the stochastic problem. The remaining two terms are effects which follow from the introduction of the stochastic element. They are made up of third derivatives of the utility function U(C,H), recognizing that C and H are linked by the budget constraint which ties the values of I and C together and by the production function for H which ties H and I together. The stochastic element in H is transmitted to I through the intertemporal optimal investment decision, and changes in I which are consequences of uncertainty about the inherent evolution of H necessarily have consequences for C. (This is a case of the situation discussed by Dardanoni (1988).) Because the utility function is a function of H and C alone – I does not appear in the utility function – we do not have any U_I terms, only U_C terms.

The presence of the often hard to sign third derivatives is a standard feature of problems of choice under uncertainty. One point to keep in mind in trying to interpret these terms is that they can also be read as second derivatives of marginal utility terms – thus U_{CCC} is both the third derivative of U(C,H) with respect to C and the second derivative of the marginal utility term, U_C(C,H) with respect to C. This latter interpretation is important since we are concerned with the level of a choice variable, and optimal levels are determined by marginal, more directly than total, utilities.

Consider first the last term on the RHS of (47), involving the expression U_{CCC}/U_{CC}. If C and H were independent in utility, so U_{CH} = 0 for all C and H, this term would be the only difference between the stochastic and non-stochastic equations which we are comparing (i.e. the İ and [1/dt]EdI equations). In that case of direct separability in utility between C and H (note that there would still be a connection between them through I and the budget constraint), if we set σ_H to zero, so we are in the non-stochastic case, the difference between the İ and the [1/dt]EdI equations would disappear, and if σ_H is non-zero but I_H = 0 so that health investment did not respond to changes in H, we would find no difference between the stochastic and non-stochastic equations even in the presence of uncertainty. It is also the case that in the stochastic case with I_H formally non-zero, if U_{CCC} = 0, the presence of uncertainty would have no effect on the optimal trajectory of I.

The interpretation of the term U_{CCC} has been approached from a number of directions in the literature. Kendall (1990) interprets the expression $-U_{CCC}/U_{CC}$ as the coefficient of prudence, a counterpart to the more familiar coefficient of risk aversion, $-U_{CC}/U_C$. The term U_{CCC} will clearly be involved in the way the coefficient of risk aversion changes as C changes, so U_{CCC} has also been discussed as an element in

Decreasing Absolute Risk Aversion. The U_{CCC} term has also been shown to be related to the individual's degree of downside risk aversion (see, for example, Liu and Meyer (2012)).

One way to look at the role of U_{CCC} is to consider a utility function which is a function of the random variable C alone, and use a third order Taylor series expansion to find an expected utility (EU(C)) expression. We do this taking EC, the expected value of C, as our point of expansion. First, we write:

(48) $U(C) = U(EC) + U_C(EC)[C – EC] + ½\, U_{CC}(EC)[C – EC]^2 + 1/6\, U_{CCC}(EC)[C – EC]^3$

Next, since we want to find an expression for EU(C), we apply the expectations operator to both sides of the Taylor series expansion. Note that because we are expanding at the exact point EC, the U(EC) and related terms are not stochastic, so the only stochastic elements on the RHS are those involving C directly. This gives:

(49 $EU(C) = U(EC) + U_C(EC)E[C – EC] + ½\, U_{CC}(EC)E[C – EC]^2 + 1/6\, U_{CCC}(EC)E[C – EC]^3$

In this expression the term $U_C(EC)E[C – EC]$ will vanish because U(EC) is non-stochastic and E[C-EC] – the expected value of the deviation from the mean of a value of the random variable C – will equal zero. This gives:

(50) $EU(C) = U(EC) + ½\, U_{CC}(EC)E[C – EC]^2 + 1/6\, U_{CCC}(EC)E[C – EC]^3$

In this expression we see that the deviation of the expected utility of C from the utility of the expected value of C (the Jensen's Inequality effect, sometimes referred to as the utility premium of uncertainty) depends on the second derivative of U, multiplied by the variance of C and the third derivative of U multiplied by $E[C – EC]^3$ (note that we are taking the expectation of a cube here, so that fact that $[C – EC]^3 = [C -EC][C – EC]^2$ does not mean that its expectation is equal to zero even though the expectation of [C – EC] = 0). The term $E[C – EC]^3$ is a measure of the degree of skewness of the distribution of the random variable C. If $E[C – EC]^3$ is negative, the distribution of C is skewed to the left – C has a long left, or down-side, tail – and if the expectation of the cube is positive, the distribution of C is skewed to the right.

If $U_{CCC} = 0$ for all values of C, our individual's utility is not affected by the presence of skewness in the distribution of her C. Note that, so long as we maintain the assumption of risk aversion, so that $U_{CC} < 0$, the presence of uncertainty will reduce her expected utility relative to the level of utility associated with her expected level of consumption, and the larger the variance of the distribution of C – loosely, the greater the degree of uncertainty she faces, the further EU(C) will be below (EC). Thus, in terms of our discussion of equation (47) above, if she is risk averse but has $U_{CCC} = 0$, our individual's utility will be reduced by the uncertainty but she will not adjust her quantitative consumption decision in response to it (although in an insurance model she will be prepared to buy insurance against the loss, if she is risk averse).

The usual assumption in the literature is that $U_{CCC} > 0$. In terms of equation (50) for EU(C), this means that our individual's expected utility will be reduced if the distribution of C is left skewed and increased if the distribution of C is right skewed. Thus she will generally be prepared to adjust her consumption to reduce the degree of left skew, but not to reduce the degree of right skew. (Extending the Taylor series expansion to the third order helps resolve the question of why the same risk-averse individual will be prepared to buy insurance against a loss and at the same time buy lottery tickets. In the case of insurance, she is facing the chance of a large downside shock, meaning a heavily left skewed element in the distribution of her wealth in any period. This prospect will reduce her expected utility so she will be willing to pay to avoid it. In the case of the lottery ticket she is contemplating a very right-skewed, albeit usually low probability, shock to her wealth, independent of any downside shocks that she might wish to insure against, and she would be willing to pay to acquire that. Thus, since the loss to be insured against and the possible gain from a winning lottery ticket are independent elements in the overall probability distribution function for her wealth; an individual with $U_{CCC} > 0$ would be prepared both to insure and to buy lottery tickets. The insurance/lottery paradox rests on assuming that only the mean and variance of wealth matter in the determination of expected utility, and that skewness does not. Note that if we take the Taylor series expansion out to fourth order, the degree of Kurtosis of the distribution of C, $E[C - EC]^4$, will enter the expression for expected utility. It has been suggested that this term plays a role in financial investment portfolio decisions along with the mean, variance and skewness of the returns on various assets.

Most individuals can safely be assumed to be averse to leftward skewness (i.e. downside risk averse) and open to rightward skewness in the distribution of things like wealth and health, so it is reasonably safe simply to assume that $U_{CCC} > 0$. The cross-third partial terms in (47) are rather more difficult to interpret. Dardanoni (1988), and Dardanoni and Wagstaff (1990), focus their discussion on measures of partial risk aversion (i.e. risk aversion in the direction of either one of the arguments of a two-argument utility function in isolation) and so combine own and cross third partials into terms which determine the sign of changes in risk aversion in the relevant direction. A discussion of the sign of a term like $U_{CHH}(C,H)$ necessarily starts with a discussion of U_{CH} since, as we noted above, if it is everywhere separable in C and H, the third cross-partials will always be zero.

The most common assumption (which, for example, Dardanoni and Wagstaff seem to make implicitly, although its violation would not vitiate their results) is probably that $U_{CH} > 0$: the marginal utility of C increases in H, in the sense that the healthier our individual is the more utility she gets from an additional unit of consumption even though the consumption commodity in question is not directly related to her stock of health capital (in this paper we made this assumption in relation to equation (1) above). Eeckhoudt and Schlesinger (2013), however, approaching the question from the perspective of uncertainty analysis, suggest that the concepts of risk apportionment and correlation aversion yield a different signing of U_{CH}. Consider an individual who is faced with a lottery under which she will face having to take hits to both her health and her wealth. These concepts mean that, weighing up the alternatives, she would generally prefer to take a hit to her wealth when she is in good health, and to take a hit to her health when her wealth is high rather than taking hits to both wealth and health at the same time. (In the lottery context she would prefer a lottery in which the outcome is that in the next period she has a fifty percent chance of taking a hit to her health and a fifty percent chance of taking a hit to her wealth, to a lottery in which in the next period she has a fifty percent chance of taking hits to both health and wealth and a fifty percent chance of taking hits to neither.) In this literature the individual is regarded as preferring to apportion the two hits by placing them in different states of nature or, alternatively, of preferring to combine good with bad, so that only one bad will occur in the next period rather than there being a chance of both bads occurring.) Eeckhoudt and Schlesinger note that this pattern of preferences translates to $U_{CH} < 0$. Obviously the way we think about her preferences will depend on what we envision she can spend her wealth on – what C can contain. We have been defining C as non-health related goods. If we think of a hit to C as being a hit to Y, and recognize that in the face of a bad health shock the individual will increase her spending on curative care, which will cut into her C, it would seem reasonable to assume that she would prefer not to take a hit to health and a hit to wealth at the same time. This, however, is not at the essence of the signing of the third cross derivatives, as Eeckhoudt, Rey and Schlesinger (2006) note. Correlation aversion, which implies $U_{CH} < 0$, is a general preference for combining good with bad rather than facing a chance of taking the bad all at once, with the notion being that higher consumption on non-health related goods will to some degree compensate for the reduction in health state, and vice versa.

Consider the Marginal Utility of C in the case where the utility function depends on C and H: $U_C(C,H)$. We assume that $U_C > 0$ and $U_{CC} < 0$. The third partials of the utility function relate to the concavity or convexity of the marginal utility functions. If $U_{CCC} > 0$, as we have assumed above, U_C is convex in the C direction. Reverting to the case of a single argument for a moment, let $V(C) = U_C(C)$. then, taking a second order Taylor series expansion of V(C) with EC as the point of expansion we have

(51) $V(C) = V(EC) + V_C(EC)[C - EC] + ½\, V_{CC}(EC)[C - EC]^2$

And taking expectations on both sides we have

(52) $EV(C) = V(EC) + ½\, V_{CC}(EC)E[C - EC]^2$

So whether the presence of uncertainty (a non-zero variance term, $E[C - EC]^2$) raises or lowers EV(C) relative to V(EC) will depend on whether $V_{CC} >< 0$. If $V_{CC}(C) > 0$, the presence of uncertainty will, by the Jensen's inequality effect with a convex function, pull EV(C) up relative to V(EC). However, since $V(C) = U_C(C)$, we can re-write (52) as

(53) $EU_C(C) = U_C(EC) + ½\, U_{CCC}(EC)\, E[C - EC]^2$

So $V_{CC}(C) > 0$ translates to $U_{CCC}(C) > 0$ and our assumptions about U(C) translate into marginal utility which is positive ($U_C > 0$), diminishing in C ($U_{CC} < 0$) and convex in the C direction ($U_{CCC} > 0$). Uncertainty about C pulls the individual's expected utility from consumption down but pulls her expected marginal utility up.

For the cross derivatives we are still looking at $U_C(C,H)$, but now looking at the shape of the marginal utility of C curve in the H direction. We still assume that $U_C > 0$, but now the sign of U_{CH} tells us whether the MU of C curve, $U_C(C,H)$, is upward or downward sloping when we plot it against H, and the sign of $U_{CHH}(C,H)$ tells us whether the U_C curve is convex or concave in the H direction. In terms of the effect on U_C of uncertainty about H, it is the case that regardless of whether U_{CH} is positive or negative, if U_{CHH} is positive (so U_C is convex in the H direction, regardless of whether the slope of U_C in that direction is positive or negative) uncertainty about H will pull the marginal utility of C up, and if $U_{CHH} < 0$, so U_C is concave in the H direction, uncertainty about H will pull the marginal utility of C down. The case of $U_{CHH} > 0$ is often referred to in the literature as the case of cross-prudence in consumption and the case where $U_{CHH} < 0$ is referred to as the case of cross-imprudence in C. Similarly, $U_{HCC} > 0$, the case where uncertainty about C pulls the marginal utility of H up, is often referred to as cross-prudence in Health and $U_{HCC} < 0$ is referred to as cross-imprudence in health[6].

VI. Comparative Dynamics in a Phase Diagram

In equation (47), the prudence term, written out in full, is $[U_{CCC}/2PU_{CC}]\ \sigma_H^2\ PI_H^2$. We have assumed that $U_{CCC} > 0$, and diminishing marginal utility gives us $U_{CC} < 0$. Since the remaining terms are all positive, if we look at this term alone (or equivalently, assume for the moment that $U_{HC} = 0$), [1/dt]EdI contains, in addition to $\dot{I}$, a negative term. In particular this means that for values of I and H for which $\dot{I} = 0$, [1/dt]EdI will be negative. In the phase diagram for the non-stochastic case, setting $\dot{I} = 0$ lets us find the stationary locus for I. In the stochastic case the set of (I,H) values for which $\dot{I} = 0$ will be associated with a negative expected instantaneous drift in I. If we want [1/dt]EdI = 0, given that the last term in (47) is negative, we must make the first term positive, i.e. we must be in a region on the (I,H) phase diagram in Figure 1 which is associated with $\dot{I} > 0$. The phase arrows for the non-stochastic case indicate that I will be increasing at points above the non-stochastic stationary locus, so we can say in general that the points for which the expected drift in I in the stochastic case will equal zero will lie above the stationary locus for I from the non-stochastic case. Thus we can draw a stationary-in-expectation locus for I in the stochastic case lying above the non-stochastic stationary locus for I. Note that the stationary locus for H will not be affected because it is a simple mathematical relation – given H and δ, what value would I have to take on to make $\dot{H} = 0$[7]. Behavioural factors are in the stationary locus for I. We have illustrated this case in Figure 2.

[FIGURE 2 ABOUT HERE]

In Figure 2, the trajectory labelled A was the optimal lifetime trajectory from Figure 1, the non-stochastic case, and the trajectory labelled B is its counterpart for the stochastic problem.

With regards to the shape of the new trajectory, since the last term in (47) is negative (and continuing to assume that $U_{CH} = 0$), we can also say that at any given (I,H) pair the expected drift in I in the stochastic case will be less than the change in I in the non-stochastic case. Thus I will, in expectation, rise more slowly and fall more quickly, and there will be a region on the diagram in which in the non-stochastic case I was rising but in the stochastic case I will be tending to drift down. With enough of an upward shift in the stationary locus in expectation, the segment of the original, non-stochastic trajectory along which I was rising will now be in the region of the phase diagram in which I is tending to drift down. This raises the question of how the individual's choice of initial value of I will be affected by the introduction of Weiner-type uncertainty. We note that we have not changed the planning horizon of the problem, and that the terminal transversality conditions for the stochastic problem will be the same, albeit in expectation, as those for the non-stochastic problem, meaning that the individual will not want to reach the horizontal axis significantly sooner in the stochastic case[8]. Thus, assuming that in the stochastic

[6] On the inspiration for the terminology see Eeckhoudt and Schlessinger (2013) and L. Eeckhoudt, B. Rey and H. Schlesinger (2006), and for an application of the concepts in health economics see M. Brianti, M. Magnani and M. Menegatti (2017).

[7] In addition, looking at equations (2) and (12) above we see that [EdH]/dt = 0 in (12) when $\dot{H} = 0$ in (2).

[8] On terminal transversality conditions in stochastic control problems see Malliaris and Brock (1982).

case the individual is born with the same initial H as she was in the non-stochastic case, we expect the starting point for her optimal lifetime trajectory to shift up. From our interpretation of (47) we also expect I to be rising more slowly in expectation, even though it is above the relevant stationary locus, so we expect the expectational trajectory to be flatter in the region where it is positively sloped, then to fall more steeply in the region where it is negatively sloped. Thus assuming that our individual starts with the same level of H and has the same planning horizon in both cases, we expect Wiener-type uncertainty about the effect of I over time to translate into a tendency to front-load I, then at some point to let it start falling faster than in the non-stochastic case.

We still have the question of the effect of the second term in (47) – the term involving cross-prudence effects – on the individual's lifetime trajectory. Here again, U_{CC} will be negative. U_{CHH} will be positive if our individual's marginal utility of C function is convex in H, so that uncertainty about H tends to cause the MU of C to drift up (in the Weiner process intertemporal case), regardless of the sign of U_{CH}. On the other hand, if we start from the assumption that U_{CH} is positive, so that being in better health increases the marginal utility which our individual derives from another unit of C, and if we assume that the effect of increases in H on the MU of C diminishes with increasing H, then U_{CHH} will be negative. Focusing on this first part of the second element on the RHS of (47), we have $[U_{CHH}/2PU_{CC}]\sigma_H^2$, which will be negative and hence reinforce the effect of the final term if $U_{CHH} > 0$.

Turning to the second part of the second element on the RHS of (47), we have $[-2PU_{HCC}I_H/2PU_{CC}]\sigma_H^2$. Here again we must consider the convexity or concavity of a marginal utility term – this time U_{HCC}, the convexity or concavity of the marginal utility of H with respect to C. Here, however, we have an additional term I_H, which reflects how increases in our individual's level of H affect her propensity to invest in H. It seems reasonable to assume that, the healthier our individual is, the lower the marginal benefit of additional units of H, especially if we assume there exists at any point in time some upper limit to H – some notion of perfect health. Thus we can make an argument for $I_H < 0$. Then the sign of this second element will depend on the sign of the cross-prudence term. If the MU of H is convex in C, and the MU of C is convex in the H direction, then the second element on the RHS of (47) will reinforce the third element in pulling the expectational stationary locus for I up and shifting the expectational trajectory for I up in the earlier part of our individual's lifetime plan.

VII. Conclusion

In this paper we have extended the analysis of investment in health under uncertainty to the case of cumulative uncertainty about the effects of health investments or, perhaps, health behaviours, which the individual engages in on a continuous basis. This would include cases relating to diet, exercise, or other health-related activities or habits as distinct from the treatment (whether preventive or curative) activities which have been the focus of most of the previous literature. We analyze the individual's problem using techniques of stochastic control theory in the case where uncertainty can be represented by a Wiener process, in this case, in the equation of motion for Health Capital.

As we would expect from the uncertainty literature broadly defined, our results depend on third derivatives of utility functions – on prudence and cross-prudence or cross-imprudence terms. While much of the literature interprets third derivatives directly – in terms of the curvature of the utility function and in particular in relation to Decreasing Absolute Risk Aversion (DARA) - we also note that the third derivative of a utility function can be interpreted as the second derivative of a marginal utility function. This is in some ways a more intuitive way of looking at the elements added to the equations which determine the choice of the level of health investment, since it is marginal, rather than total, utility which appears in the first order condition which is solved to determine the optimal level of health investment. The use of a continuous time stochastic control framework is also useful in terms of interpretation: in a static uncertainty model the effect of uncertainty is discussed in terms of whether the $EU_C(C)$ is greater or smaller than $U_C(EC)$, for example, in the case of the continuing processes invoked when we introduce Wiener processes into stochastic control problems the consequence of the continuous random-walk characteristic of the Wiener process is to cause a marginal utility term to continuously drift up or down as time passes, depending on the concavity or convexity of $U_C(C)$.

We note that the signing of these convexity/concavity terms is not necessarily easy – the signs which we might settle on instinctively for terms like $U_{CH}(C,H)$, for example – probably complementarity in utility, making the sign positive – is not necessarily the same sign we would settle on if we approach the question via the risk apportionment approach, as Eeckhoudt and Schlessinger (2013) have shown, and the difference can significantly affect individuals' long term health investment decisions. One advantage of the stochastic control approach which we have adopted here is that it allows for illustration of the effect of the introduction of uncertainty not just by setting out the equations representing the necessary conditions for the stochastic and deterministic cases and noting their differences; it also lets us discuss, at least in qualitative terms, using a phase diagram, how those differences can be expected to play out in individuals' optimal lifetime health investment trajectories.

In any given population there is likely to exist a mix of people who have $U_{CH} > 0$ and those who have $U_{CH} < 0$, as well as mixtures of different degrees of convexity or concavity of the U_C function with respect to H. This can be expected to be a complicating factor, which needs at least to be considered, even if it cannot be completely controlled for, in empirical implementation of Grossman's model of investment in health capital.

References

Arrow, K. H. (1963): "Uncertainty and the welfare economics of medical care" American Economic Review 53, 941-973

Brianti, M., M. Magnani and M. Menegatti (2017): "Optimal choice of prevention and cure under uncertainty on disease effect and cure effectiveness", Research In Economics, in press, doi:10.1016/j.rie.2017.03.005

Cropper, M. L. (1977): "Health, investment in health and occupational choice" Journal of Political Economy 85, 1273-1294

Dardanoni, Valentino (1988): "Optimal Choices under Uncertainty: The case of two-argument utility functions" Economic Journal 98, 429-450

Dardanoni, Valentino and Adam Wagstaff (1990): "Uncertainty and the Demand for Medical Care" Journal of Health Economics 9, 23-38

Eeckhoudt, Louis and Christian Gollier (2005): "The impact of prudence on optimal prevention" Economic Theory 26, 989-994

Eeckhoudt, Louis and Harris Schlesinger (2013) : "Higher-Order Risk Attitudes" Chapter 2 in G. Dionne (ed.) Handbook of Insurance Springer

Eeckhoudt, Louis, Beatrice Rey and Harris Schlessinger (2006): "A Good Sign for Multivariate Risk Taking" Management Science 52(1), February

Ferguson, B.S. and G.C. Lim (1998) Introduction to dynamic economic models Manchester University Press, Manchester.

Grossman, Michael (1972): "On the concept of health capital and the demand for health" Journal of Political Economy 80, 223-255

Huggett, Mark and Edward Vidon (2002): "Precautionary wealth accumulation: a positive third derivative is not enough" Economics Letters 76, 323-329

Kimball, Miles S. (1990) "Precautionary Saving in the Small and in the Large" *Econometrica* 58(1), January, 53-73.

Laporte, A. (2015): Should the Grossman Model of Investment in Health Capital Retain its Iconic Status? Canadian Centre for Health Economics Working Paper 2014-04, Revised January 2015: http://www.canadiancentreforhealtheconomics.ca/wp-content/uploads/2015/01/Laporte-R.pdf

Laporte, A. and B. S. Ferguson (2007): "Investment in health when health is stochastic" Journal of Population Economics 20(2), 423-444

Liljas, Bengt (1998): "The demand for health with uncertainty and insurance" Journal of Health Economics 17, 153-170

Liu, Liqun, and Jack Meyer (2012): “Decreasing absolute risk aversion, prudence and increases downside risk aversion” Journal of Risk and Uncertainty 44(3), 243-260

Malliaris, A.G. and W.A. Brock (1982): Sochastic Methods in Economics and Finance North Holland

Mangel, Marc (1985): Decision and Control in Uncertain Resource Systems, Academic Press, Orlando.

Menegatti, Mario (2009): “Optimal prevention and prudence in a two-period model” Mathematical Social Sciences 58, 393-397,

Picone, Gabriel, Martin Uribe, R. Mark Wilson (1998): “The effect of uncertainty on the demand for medical care, health capital and wealth” Journal of Health Economics 17, 171-185

Pindyck, Robert S. (1980) “Uncertainty and Exhaustible Resource Markets” *Journal of Political Economy* 88, December, 1203-25.

Pindyck, Robert S. (1982) “Uncertainty and the Behavior of the Firm” *American Economic Review* 72(3), June, 415-427.

Rothschild, M. and J. E. Stiglitz (1970): “Increasing risk: I. A definition” Journal of Economic Theory 2, 225-243

Figures

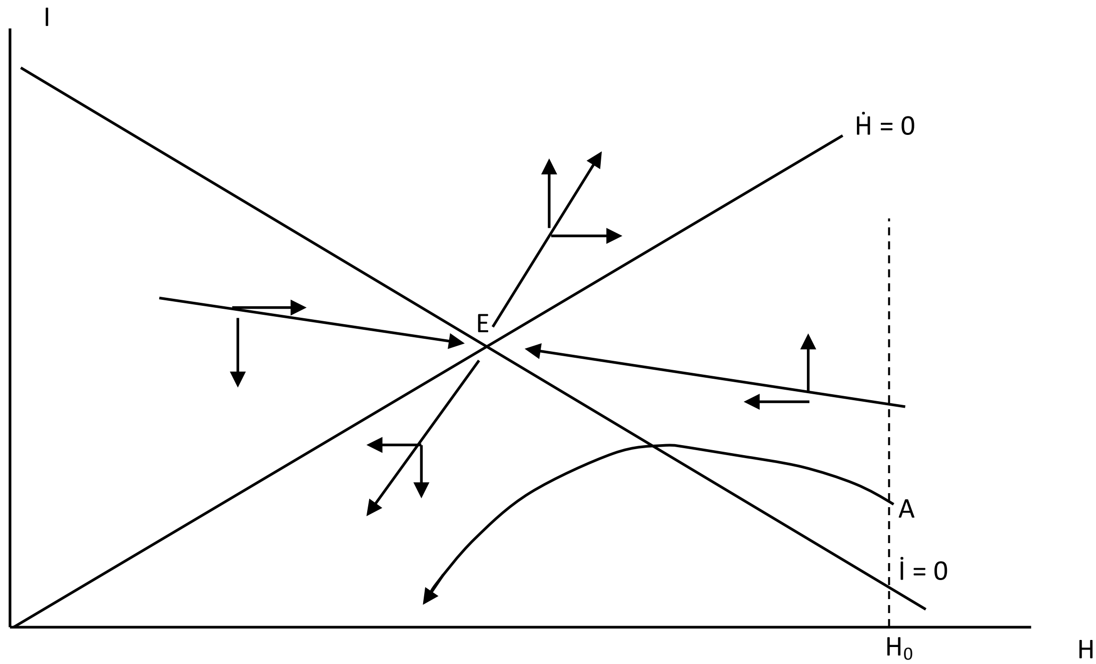

Figure 1 Phase Diagram for the Deterministic Case

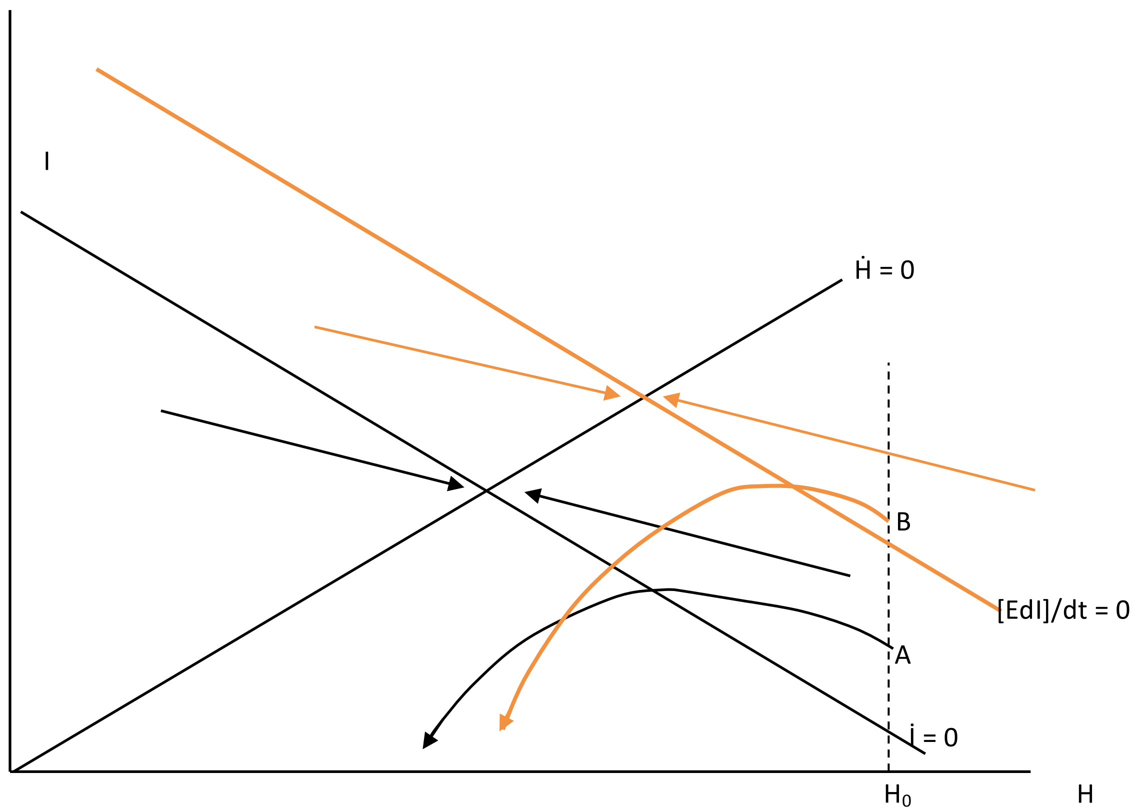

Figure 2 Phase Diagrams comparing Deterministic and Stochastic Case for Weiner Process