The Timing and Probability of Treatment Switch *

Felipa *de* Mello-Sampayo[†]

[†]ISCTE–IUL Institute University of Lisbon

Abstract

In most cases patients are not compelled to switch treatments at any specific moment, they hold an option to switch treatments that should only be exercised when it is optimal to do so. This paper sets up a stochastic model that provides an optimal rule for the timing of treatment switch. In the real world, cost fluctuations render the outcome of any treatment switch uncertain, so that patients might have to wait for more information before optimally switching treatments, especially when the incremental cost per quality-adjusted life year (QALY) gained cannot be fully recovered later on. The results of the model were then tested empirically with imatinib cost data. The empirical results support the findings of the analytical model.

JEL Classification: D81; I10; I12; I13.

Keywords: Irreversibility; Uncertainty; Cost-Benefit Analysis; Health; Medical Technology; Pharmaceutical.

*Financial support from Fundação para a Ciência e Tecnologia, under UNIDE-BRU, for the project entitled "Health and Economic Growth" PTDC/EGE-ECO/104157/2008 is also gratefully acknowledged. Corresponding author's email: `filipa.sampayo@iscte.pt`

1 Introduction

For some diseases, patients receive a sequence of treatments. These may involve different drugs or different dosages of the same drugs. The decision regarding whether to move a patient to the next treatment in a sequence may be based on patient characteristics or patient history, and therefore subject to variability. If it is accepted that adoption decisions should be made with consideration of the associated decision uncertainty, then we may say that models submitted to decision-makers should do two things: estimate expected net benefit (NB) and characterize decision uncertainty. If this dual purpose of models is accepted, failure to fulfil the latter requirement will limit its value for decision-making and leave the decision-maker without a key element of information.

The decision to adopt a particular technology should be based on expected net benefit so that, when comparing mutually exclusive treatment strategies for a particular disease, the optimal strategy is simply the one with the highest expected NB (Claxton, 1999). Nevertheless, decisions based on expected NB are only appropriate if there is also some consideration of whether current evidence is sufficient for allocating health-care resources, based on an assessment of the consequences of decision uncertainty (Sculpher and Claxton, 2005). If the decision uncertainty and the consequences of adopting a suboptimal treatment strategy are large, the decision-maker may require further evidence on which to base the adoption decision (Claxton, Eggington, Ginnelly, Griffin, McCabe, Philips, Tappenden and Wailoo, 2005).

For example, adopting some medical technologies restricts the use of certain medical technologies in the future, and explains the lack of consensus about when to start therapy in HIV patients (Cohen, 2000, Harrington and Carpenter, 2000). Some advocate fighting HIV with a powerful combination of drugs as early as possible in the course of the disease in order to prevent the disease from progressing. Others are concerned that starting therapy at early stages, may lead to the development of viral resistance to these drugs and related compounds. Although the disease may progress to an advanced stage more rapidly, while other clinicians advocate waiting until the disease reaches a more advanced stage to initiate treatments so that future options can be preserved. This problem of current decisions affecting future options has received considerable theoretical attention in the literature on economic investments. The higher the uncertainty about future outcomes, the more individuals will gain from waiting for more information before committing to investment (or dis-investment) whenever there are significant sunk costs (Pindyck, 1988). This result is a prediction of the "option-pricing" approach to the analysis of irreversible investment under uncertainty (Dixit, 1989*a*,*b*, Dixit and Pindyck, 1994). Numerous empirical studies have shown that ignoring this investment option value can be an error (see, for example, McDonald and Siegel (1985), Pindyck (1988), Pennings (2000), Brito and *de* Mello-Sampayo (2005), Yu, Chang and Fan (2006), *de* Mello-Sampayo, *de* Sousa-Vale and Camões (2011)). Analogously, the benefits associated with actions that preserve treatment choices in the future, above and beyond the direct value associated with those actions, are referred to as the option value of the intervention.

For many physicians the observation that current medical treatment decisions have repercussions for the treatment of health conditions in the future is an obvious one that is often considered in their clinical decision making. Such considerations form no part of health-care technology assessment calculations, leading to potentially significant mischaracterizations of the treatment value. Whilst it is difficult to systematically assess the size of the bias induced from ignoring option values, the only empirical study in the health domain found an increase in consumer willingness-to-pay of approximately 53% when option values were considered

(Smith, 2007).

Using the "option-pricing" approach to the analysis of irreversible treatment choices under uncertainty is important since the health sector is one in which there is tremendous uncertainty about the demand for future medical technologies. When we begin treating a population of individuals, we do not know what additional conditions they will develop in the future. Since new diseases are constantly emerging, we do not even necessarily know the nature of these future conditions. Higher life expectancy prospects for new conditions to arise, especially those associated with aging such as cancer and dementia, make the option value of the interventions a key variable of the valuation equation. Ignoring option values during the drug approval and reimbursement setting process could result in disincentives to create socially valuable technologies. Finally, unlike many private investment decisions, decisions taken by national health systems may be effectively irreversible for political reasons. Palmer and Smith (2011) focus on the timing of health investments and whether it makes sense to delay adoption of a new technology in anticipation of the exogenous arrival of new information about its value. While the prospects for delaying investments has potentially important implications for decision making, delay is often not feasible in this setting, especially on the time scale under which we expect new information to arrive. When analyzing situations where current treatment decisions have irreversible implications for the treatment of future diseases, and decision makers are choosing between competing interventions with differing temporal consequences, Zivin and Neideill (2009) find that irreversibility raises the value of treatment modalities that preserve future treatment options. However, introducing some reversibility can either increase or decrease the option value, depending on the distribution of patient types. These authors also examine the relationship between these values and the biological and economic parameters that characterize any given set of technologies.

In this paper we consider the problem of a patient that is using a specific treatment but is contemplating using higher doses of the same treatment and then is provided with a more advanced drug. The patient will use the new drug only if such a move is deemed beneficial in the medium and long term. That, in turn, will depend on the perceived evolution of cost. The higher the uncertainty regarding the cost of a new treatment, the more likely it is that a favourable situation will turn into an unfavourable one, and the more the patient will gain from waiting for more information before committing to the new treatment whenever the incremental cost per quality-adjusted life year (QALY) gained cannot be fully recovered later on. This result is a prediction of the option pricing approach to the analysis of irreversible investment under uncertainty (Dixit, 1989*a*,*b*, Dixit and Pindyck, 1994).

With the aim of empirically testing this study's "option-pricing" model, an econometric application uses data from a modelling exercise that compared alternative treatment pathways for patients with unresectable gastrointestinal stromal tumour (GIST) who failed to respond to imatinib 400 mg/day (Hislop, Quayyum, Elders, Fraser, Jenkinson, Mowatt, Sharma, Vale and Petty, 2011). The study of Hislop et al. (2011) assessed the effectiveness and cost-effectiveness of imatinib at escalated doses of 600 and 800 mg/day following progression of disease at a dose of 400 mg/day, compared with sunitinib, or the provision of best supportive care (BSC) only for patients with unresectable and/or metastatic GISTs. Several studies have reported further disease control after progression on an initial imatinib dose of 400 mg/day with dose escalation of imatinib to 800 mg/day, and this has also become common practice (Zalcberg, Verweij, Casali, Cesne, Reichardt, Blay and *et* al., 2005, Blanke, Demetri, *von* Mehren, Heinrich, Eisenberg, Fletcher and *et* al., 2008). However, it should be noted that current NICE guidelines for imatinib do not actually recommend dose escalation for patients with unresectable and/or metastatic GISTs who progress on an initial dose of 400 mg/day (NICE, 2004) but suggest that clinical decisions are made on an

individual case-by-case basis, reflecting uncertainty regarding optimal practice.

Three studies (Wilson, Connock, Song, Yao, Fry-Smith, Raftery and Peake, 2005, Huse, *von* Mehren, Lenhart, Joensuu, Blanke, Feng and et al., 2007, Mabasa, Taylor, Chu, Moravan, Johnston, Peacock and *et* al., 2008) compared imatinib with BSC. The study by Wilson et al. (2005) used the manufacturer submissions (Novartis model) and compared imatinib and BSC, but in the imatinib group allowed for escalation of doses from 400 to 600 mg/day for those who failed to respond or were intolerant to imatinib at the 400 mg/day dose. The study by Mabasa et al. (2008) noted that patients included from retrospective cohorts in their analysis were given imatinib 400 mg/day until disease progression, and later were allowed escalated doses of between 600 and 800 mg/day. Six out of 56 patients in the imatinib groups of patients considered in this economic evaluation were then allowed to switch to sunitinib therapy. The economic evaluation by Huse et al. (2007) considered imatinib at 400 mg/day. Two studies (Contreras-Hernandez, Mould-Quevedo, Silva, Salinas-Escudero, Villasis-Keever, Granados-Garcia and *et* al., 2008, Teich, Hashizume and Follador, 2009) compared both imatinib and sunitinib with BSC for patients who had failed or become resistant to imatinib 400 mg/day.

The empirical application of this study assumes that patients start with 400 mg/day and the second-line treatment consists of dose escalation of imatinib to 600 mg/day followed by sunitinib. The empirical results suggest that as sunitinib becomes less cost competitive, the cost uncertainty becomes more dominant. With limited substitutability, higher quality of sunitinib will increase the demand for sunitinib disregarding the cost uncertainty. Several key insights emerge. The existence of an option value means that imatinib at 400 mg/day may be the better choice when considering lifetime welfare. Thus, under irreversibility, low-risk patients must begin the second-line treatment as soon as possible, which is precisely when the second-line treatment is least valuable. As the costs of reversing current treatment impacts fall, it becomes more feasible to provide the option-preserving treatment to these low-risk individuals later on.

This paper is organized as follows. The following section develops the stochastic option-pricing model, specifying the two feasible treatments and examining the impact of cost shocks on both the timing of treatment switching and the NB of each treatment. Section 3 presents the probability and expected time of treatment switch. The discussion of the results with an empirical application are presented in Section 4. Conclusions are discussed in Section 5.

2 The Model

In this section we develop a simple model to illustrate the role that uncertainty and irreversibility can play in determining the decision regarding whether to move a patient to the next treatment in a sequence. Consider a chronic disease that occurs in period 1 and lasts two periods. Suppose that there are two choices in treating this disease, denoted by T_1 and T_2, respectively. Patients start being treated with T_1 and patients treated with T_2 in the second period attain a higher benefit but pay more.

The value of each technology is expressed in terms of net benefits, i.e. the value of health benefits generated by treatment (B) minus the costs of that treatment (C). The net benefits from treatment with T_1, at time t, is denoted by NB_{1t} and the net benefits from treatment with T_2 is denoted by NB_{2t}.

Population individuals, $N_t = \{1, 2, ...n\}$, are endowed with one unit of labour which is inelastically supplied to firms, and receive wage income at the rate w_t. We assume that the individual's health level augments with the public provision of health care services (see

Chakraborty (2004)). The benefit from treatment at time t, B_t, depends upon her health capital, h_t. We assume that the per capita health investment at time t, h_t, is financed at a balanced budget with a (constant) wage income tax, $0 < \tau < 1$, that is:

$$h_t = \tau w_t. \tag{1}$$

Patients observe and decide the viability, utility, and characteristics of health care goods and services only after using those products or services. Thus, the quality of health care good or service can only be ascertained upon their consumption. In such cases, a drop in price is often interpreted by the prospective consumer as a drop in quality or utility of the product or service. Indeed, it is possible for the demand curve for medical care to be upward sloping, even though medical care is a non-inferior good[1], a relationship that has some empirical support (Hoi and Robson, 1981, Hau, 2008, Dusansky and Cagatay, 2010). Under this hypothesis the demand for medical care is given by:

$$T = B_t h_t^{\varphi}, \tag{2}$$

where T is the total quantity demanded of the treatment at time t, B_t is the benefit gain at time t, and φ is the parameter for the elasticity of demand. We consider that medical care operates where patient's demand for treatments is inelastic[2], $0 < \varphi < 1$. In this model set-up, the mechanism that leads health to be a Giffen good also involves a wealth consideration; when the price of treatment falls, the patient is effectively wealthier, he can afford more treatments generally and so, he needs fewer treatments of this kind. From Equation (2), the benefit function of a representative patient is given by:

$$B_t = \frac{T}{h_t^{\varphi}}. \tag{3}$$

The benefit increases with the number of treatments, but it is inversely related with the per capita health investment. The higher patient's health capital, the fewer treatments he needs for the same benefit. The cost of the treatment, C_t, is a nonlinear function of patient's particular characteristic, x, that evolves over time:

$$C_t = rkx_t^2, \tag{4}$$

where r is the interest rate, k is the capital invested in the treatment and x is the level of the state variable that represents the random shock of the cost side at time t. For analytical tractability, the state variable is assumed to evolve according to a geometric Brownian motion:

$$dx = \alpha x dt + \sigma x dz, \tag{5}$$

where $dz = \varepsilon_t \sqrt{dt}$ is the increment of a Wiener process and

$$\varepsilon_t \curvearrowright N(0,1), \quad E(\varepsilon_t, \varepsilon_s) = 0 \text{ for } s \neq t.$$

[1]Housing is another example of a non-inferior good whose own-demand can be upward sloping (see Dusansky and Ç. Koç (2007)).

[2]An example of perfectly inelastic demand would be a life saving drug that people will pay any price to obtain. Even if the price of the drug were to increase dramatically, the quantity demanded would remain the same.

Equation (5) implies that the current value of the random shock is known, but the future values are log-normally distributed with a variance growing linearly with the time horizon[3].

We calculate net benefit for the patient using $T_i, i = \{1, 2\}$:

$$NB_{it} = B_{it} - C_{it}. \tag{6}$$

Firms producing treatments are identical and competitive. Treatment is produced using a Cobb-Douglas production technology:

$$T_{it} = k^{\theta}, i = 1, 2 \tag{7}$$

where $k = K_t/N_t$ is capital per worker and $0 < \theta < 1$. Profit maximisation implies that factor inputs are paid by their marginal products, that is,

$$\begin{array}{rl} r = & \theta k^{\theta-1}, \\ w = & (1-\theta)\, k^{\theta}. \end{array} \tag{8}$$

We assume treatment to be unit price, and capital totally depreciates at the end of every period.

2.1 Expected Net Benefit

By maximising the net benefit at time t, NB_{1t} can be derived as:

$$NB_{1t} = x^{\delta}\left[\frac{\theta^{\frac{\varphi-1}{\varphi}}}{(1-\theta)}\frac{\varphi\,(1-\varphi)^{\frac{1-\varphi}{\varphi}}}{\tau}\right], \tag{9}$$

where $\delta = 2(\varphi-1)/\varphi$. Using Ito's lemma, it can be confirmed that x^{δ} also follows a geometric Brownian motion: $dx^{\delta} = (\alpha\delta - \frac{\delta(\varphi-2)}{\varphi}\sigma^2)x^{\delta}dt + \delta\sigma x^{\delta}dz$.

We now calculate net benefit for the patient engaged in T_2. Here we assume that $B_1 = \pi B_2$, $\pi < 1$ and that $C_1 = \phi C_2$, $\phi < 1$. In other words, T_2 yields better results but is more expensive. Thus, there is an option value associated with using T_2 but that option comes at a cost.

By maximising the NB at time t, NB_{2t} can be derived as:

$$NB_{2t} = x^{\delta}\left[\frac{\theta^{\frac{\varphi-1}{\varphi}}}{(1-\theta)}\frac{\varphi\,(1-\varphi)^{\frac{1-\varphi}{\varphi}}}{\tau}\frac{\phi^{\frac{1-\varphi}{\varphi}}}{\pi^{\frac{1}{\varphi}}}\right]. \tag{10}$$

Proposition 1: *Higher volatility of patient´s particular characteristic reduces the expected NB associated with both technologies.*

Proof:

Comparing the state variables in Equations (15) and (17) in Appendix A, the expected NB is proportional to x^{δ} under both technologies. Recalling that medical care operates where the patient's demand for treatments is inelastic ($\varphi < 1$), it follows that $\delta < 0$. Consequently, cost uncertainty reduces expected NB under both treatments[4].

[3] We may also assume that x varies across patients, such that x_i may also follow a Brownian motion, for $i = 1, 2, ..., n$, $dx_i = \alpha_i x_i dt + \sigma_i x_i dz$. The expected value of dx_i across all patients, $E[dx_i]$, cannot be used to derive the expected value of C_t; an estimate of $E[dx_i^2]$ is required because $E[dx_i]^2 \neq E[dx_i^2]$. Thus, an analysis which failed to account for the variability in x across patients would provide biased estimates of expected costs. Such multi-patients processes are important in health economics, but we will not develop this more advanced application in this study.

[4] Under our model set-up, the magnitude of this adverse effect is identical for both technologies.

The second-line treatment consists of using a higher dosage of T_1 and then starting with T_2. When the patient is using exclusively T_1 her decision as to whether or not to use the second-line treatment constitutes an optimal stopping problem for which the relevant Bellman equation is:

$$V^1(x,t) = Max\{V^2; NB_1 + \frac{1}{t}E[dV^1]\}, \tag{11}$$

where $V^1(x,t)$ is the option value of intervention associated with using the second-line treatment, V^2 accounts for the expected patient's value gain that results from switching treatment and starting the second-line treatment, and the second term in curly brackets yields the time-discounted expected increment in the value of the option that arises from keeping the option unexercised for an additional lapse of time, dt. The range of values for which the second term in curly brackets is greater than the first defines the continuation region, where it is optimal not to exercise the option.

Proposition 2: *The patient will only use T_2 if the patient´s value associated with using the second-line treatment exceeds that of a situation of using T_1 , i.e.*

$$\widetilde{x} = [\frac{\beta_1}{\beta_1 - \delta} \times \frac{\mu - \alpha\delta - \frac{\delta(\varphi-2)}{2\varphi}\sigma^2}{\phi^{\frac{1-\varphi}{\varphi}}[e^{-\mu(\tilde{t}+1)}] - \pi^{\frac{1}{\varphi}}} \times \frac{\pi^{\frac{1}{\varphi}}\theta^{\frac{1-\theta}{\theta}}(1-\theta)\tau}{\varphi(1-\varphi)^{\frac{1-\varphi}{\varphi}}} \times \frac{IC_{QALY}}{\mu}(1 - e^{-\mu\tilde{t}})]^{\frac{1}{\delta}}. \tag{12}$$

Proof: See Appendix A.

In Equation (12), $\widetilde{x}$ is the critical value[5], $\widetilde{t}$ denotes the time at which T_2 will be used, and IC_{QALY} stands for the incremental cost of quality-adjusted life year (QALY[6]) gain of using a higher dosage of T_1. IC_{QALY} measures the QALY value wasted when using a higher dosage of T_1. It follows from Equation (12) and the previous assumptions on the parameters that the value of $\widetilde{x}$ is greater than zero if $\phi^{\frac{1-\varphi}{\varphi}}[e^{-\mu(\tilde{t}+1)}] > \pi^{\frac{1}{\varphi}}$, implying that the patient will use T_2 only if the relative benefit associated with T_2 exceeds the present value of the relative cost of T_2 weighted geometrically by the elasticity of demand, and that is due to the uncertainty of treatment's cost. Moreover, since

$$\frac{\partial\widetilde{x}}{\partial\sigma^2} > 0 \text{ and } \lim_{\sigma\to\infty}\widetilde{x} = \infty,$$

the greater the volatility of the cost (i.e. the higher σ^2) the higher the critical value has to be to make it optimal for the patient to use the second-line treatment. The higher the expected trend of the treatment's cost, the less the option of using second-line treatment is worth, and thus the lower the value that triggers the use of T_2, i.e.

$$\frac{\partial\widetilde{x}}{\partial\alpha} < 0.$$

The reason for this is that the more expensive one expects technology to become, the lower the uncertainty that results from the switch from a situation of using T_1 to one where the patient uses the second-line treatment.

With regard to the discount rate, the greater the patient's time discount rate, the less she values the option, and thus the lower the value x that triggers optimal treatment switch; i.e.

$$\frac{\partial\widetilde{x}}{\partial\mu} < 0.$$

[5]The value above and beyond the direct value of the second-line treatment.

[6]QALY measures the cost of disease burden, including both the quality and the quantity of life lived.

This result stems from the fact that a higher time preference increases the patient's opportunity cost of not immediately using the second-line treatment. In the extreme case where the patient cares only about the present moment, so that $\mu \to \infty$, then $\lim_{\mu\to\infty} \frac{\beta_1}{\beta_1-\varphi} = 0$ and $\widetilde{x} = 0$, so that uncertainty is disregarded and the value of the second-line treatment option collapses to zero.

The greater the tax rate, the higher the patient's health capital, thus, the more patients value using the second-line treatment option, and thus the higher the value $\widetilde{x}$ that triggers optimal T_2 use; i.e.

$$\frac{\partial \widetilde{x}}{\partial \tau} > 0.$$

Lastly, the lower the relative cost of T_2, the higher the relative benefit of T_2 and the sooner T_2 will be used, the lower the threshold for using the second-line treatment.

$$\frac{\partial \widetilde{x}}{\partial \pi} < 0 \ \ , \frac{\partial \widetilde{x}}{\partial \phi} > 0 \text{ and } \frac{\partial \widetilde{x}}{\partial \widetilde{t}} < 0.$$

3 The Probability of using T_2

Before proceeding to the empirical application, it would be interesting to ascertain, from any point within the continuation region, the likelihood that using the second-line treatment will become optimal in the future. It is important for the patient to know the expected time that will transpire until the decision of using the second-line treatment becomes optimal.

Using standard properties of the Brownian motion and the lognormal distribution, (see Dixit (1993)) closed-form solutions for the probability $Q(x)$ and expected time $\mathrm{T}(x)$ for the process x to hit the barrier $\widetilde{x}$ from any point inside the continuation region, are given by:

$$Q(x) = \begin{cases} 1 & \text{if} \quad \alpha \leqslant 0 \\ e^{\left[\frac{2\alpha(\widetilde{x}-x)}{\sigma^2}\right]} & \text{if} \quad \alpha > 0 \end{cases} \tag{13}$$

$$\mathrm{T}(x) = \begin{cases} \infty & \text{if} \quad \alpha \geqslant 0 \\ \frac{\widetilde{x}-x}{\alpha} & \text{if} \quad \alpha < 0 \end{cases} \tag{14}$$

where (α) and $\left(\sigma^2\right)$ are respectively, the drift and variance parameters of the process (x).

Equations (13) and (14) indicate that the probability and expected time until using the second-line treatment to become optimal depend on the variability and trend of the patient´s particular characteristic. The greater is the variability, σ^2, the higher is the likelihood that x diverges away from the threshold that triggers the use of the second-line treatment, and so the lower the probability that using the second-line treatment will ever become optimal. Similarly, the higher the drift, α, the more likely long excursions of x away from the critical ratio become, and so, the more time the system is expected to take until hitting the threshold beyond which using the second-line treatment is optimal.

Using the second-line treatment will become optimal with certainty provided that $\alpha < 0$ and it is expected to occur sooner the higher x and the lower σ^2. For the limiting case where $\alpha = 0$, even though the probability that the patient will start using the second-line treatment in the future is one, the expected time for it to occur is infinite. The intuition behind this apparently contradictory result is that if the drift of x is zero, long diversions away from the barrier $\widetilde{x}$ might occur. Thus, since the probabilities for successfully longer hitting times do not fall sufficiently fast, the expectation, which is the average of the possible hitting times weighted by their respective probabilities, diverges[7].

[7]This argument is presented in Dixit (1993), p. 56.

For the set of parameters for which x has a positive drift, i.e. when $\alpha > 0$, there is still a positive probability that using the second-line treatment will become optimal sometime in the future, as given by Equation (13). This is because, in spite of x drifting away from the critical ratio, there is the possibility that a combination of positive shocks might just bring the system toward the threshold barrier. However, the expected time for this event is infinite, as given by Equation (14), given that there is a positive probability that x never reaches $\widetilde{x}$ that drives the expectation into diverging.

4 Empirical Application

The model presented gives clear indications regarding treatment switch decisions under cost uncertainty. It predicts that the higher the volatility of the patient's particular characteristic, the sooner T_2 is used, the higher the tax rate and the higher the relative cost of T_2, the more valuable the option of using the second-line treatment will be, and so the fewer switches of treatment will be observed. Conversely, the higher the trend of the patient's particular characteristics, the higher the discount factor and the higher the relative benefit of T_2, the more switches of treatment one would expect to observe. Thus, for empirical testing purposes, the reduced form of equation (12) can be written as follows:

$$\tilde{x} = f \begin{pmatrix} \sigma^2 & \alpha & \mu & \tilde{t} & \tau & \pi & \phi \\ + & - & - & + & + & + & - \end{pmatrix}.$$

These results are extended for Equation (12) using simulations. The simulations are performed against a benchmark case[8]. The data in the present application consist of the cost-effectiveness of imatinib for gastrointestinal stromal tumours (Wilson et al., 2005). Current guidelines at the time of the assessment recommended an initial dose of 400 mg/day, with the option of proceeding to a higher dose in the event of a poor response or disease progression, and withdrawal of treatment in the absence of benefit after 8 weeks. Nevertheless, because of a paucity of data, the best starting dose of imatinib and best treatment pattern were highly uncertain. The model had four health states: progressive disease, treatment with 400 mg imatinib, treatment with 600 mg imatinib, and death. Patients in the imatinib treatment group began with 400 mg/day. For those patients who failed to respond to 400 mg imatinib, a random number was generated to determine whether they would be moved to 600 mg, or straight to the progressive disease state. The probability of receiving 600 mg was based on the number of patients who had responded after crossing over from 400 mg to 600 mg imatinib in a clinical trial. We assume patients start with 400 mg/day and the second-line treatment consists of dose escalation of imatinib to 600 mg/day followed by sunitinib.

Figures 1–5 provide a sensitivity analysis[9] of the trigger value $\tilde{x}$ with respect to the following parameters of the model: σ, α, μ, τ, $\tilde{t}$, π, and ϕ. The simulations carried out on the critical values of cost shock confirm the results of the comparative statics discussed above. Figure (1) reveals that the trigger value is much more sensitive to σ than to α. This is due to the fact that the higher the uncertainty, the higher the risk of treatment switch, and thus the higher the threshold in order to trigger the use of the second-line treatment. Figure (2) illustrates that the dampening influence of higher μ on the critical value strengthens as σ increases.

(Insert Figure 1 here)
(Insert Figure 2 here)

[8] See Appendix B for data description.

[9] The parameters are calibrated with the values shown in Appendix B.

Figure (3) shows that the trigger value rises when both ϕ and π increase. It illustrates that as T_2 becomes less cost competitive, the uncertainty about using the second-line treatment decision becomes more dominant. With limited substitutability, higher quality of T_2 will increase the demand for T_2 disregarding the uncertainty of treatment switch. The accentuated curvature of the surface graphed in Figure (3), in which the critical value rises very quickly as both ϕ is high and π is low, indicates that the lower is the cost-quality relationship, the more the uncertainty of the use of the second-line treatment decision becomes dominant.

(Insert Figure 3 here)

Figure (4) illustrates that as soon as one is expected to use the new technology, the uncertainty about the use of the second-line treatment decision becomes more dominant. Figure (5) illustrates the impact of σ and τ on the critical value of treatment switch, showing that the lower is the patient's health capital, the higher is the uncertainty regarding the treatment switch decision.

(Insert Figure 4 here)
(Insert Figure 5 here)

Figures (6) and (7) illustrate, respectively, the impact of σ and x on the expected probability of treatment switch and on the probability of optimal treatment switch, as given by Equations (13) and (14)[10].

(Insert Figure 6 here)
(Insert Figure 7 here)

In Figure (6), when $\alpha > 0$, the probability that $\widetilde{x}$ will be hit in the future is increasing in x and decreasing in σ^2. This is because, first, the lower is σ^2, the less valuable is the option to switch treatment, and so the more treatment switches will be observed, and second, the higher is x, the more likely it is that the process will be "thrown off-course" by a sequence of positive shocks toward the optimality threshold.

In Figure (7), our simulation benchmark values are expressed in year terms so that the simulations for the expected time for optimal treatment switch can be read in years. Figure (7) shows that the lower is σ^2 and the higher is x, the sooner the treatment switch is expected to occur. However, the further away x is from the trigger value, the greater is the impact of an increase of σ^2 on the delay expected before treatment switch becomes optimal.

In summary, the lower the volatility of cost, the more likely treatment switch is to become optimal and the sooner it is expected to occur. Moreover, treatment switch becomes likelier and is expected sooner, the closer is the system to the critical threshold, that is, the closer x is to $\widetilde{x}$.

The empirical results suggest that as sunitinib becomes less cost competitive, the cost uncertainty becomes more dominant. With limited substitutability, higher quality of sunitinib will increase the demand for sunitinib disregarding the cost uncertainty. Several key insights emerge. The existence of an option value means that imatinib at 400 mg/day may be the better choice when considering lifetime welfare. Thus, under irreversibility, low-risk patients must begin the second-line treatment as soon as possible, which is precisely when the second-line treatment is least valuable. As the costs of reversing current treatment impacts fall, it becomes more feasible to provide the option-preserving treatment to these low-risk individuals later on.

[10]The parameters are calibrated with the values shown in Appendix B.

5 Conclusion

Growth in the availability of treatments for chronic diseases that require permanent intervention, along with general increases in life expectancy, suggests that the impact of omitting option values from evaluations will only become larger. In this paper we developed a formal valuation model that addresses to these intertemporal dependencies by explicitly modelling the cost uncertainty.

The optimal timing for treatment switch becomes increasingly important for patients who might switch treatments. This paper sets up a stochastic model that provides an optimal rule for the timing of treatment switch. In the real world, cost fluctuations amongst other factors, render the outcome of any treatment switch uncertain, so that patients might have to wait for more information before optimally switching treatments, especially when the incremental cost per QALY gained cannot be fully recovered later on. Since in most cases patients are not compelled to switch treatments at any specific moment, they hold an option to switch treatments that should only be exercised when it is optimal to do so.

Viewed from the perspective of real option theory, this paper sheds new light on some debates about switching treatments. Our theoretical model suggests that cost uncertainty reduces expected NB. In addition, cost volatility discourages switching treatments. The stochastic model also illustrates that as technologies become less cost competitive, the cost uncertainty becomes more dominant. With limited substitutability, higher quality of technologies will increase the demand for those technologies disregarding the cost uncertainty. Several key insights emerge. Irreversibility raises the value of the option-preserving treatment. The existence of an option value means that a seemingly poorer treatment may be the better choice when considering lifetime welfare. Optimal decision making requires a careful comparison of the "costs" of a less effective treatment for a condition today with the "benefits" of more effective treatments in the future.

The intuition for these results is deepened when we recognise that one of the principal features driving our results is that patients have particular characteristics that make technology's cost uncertain. Thus, under irreversibility, low-risk patients must begin the option-preserving treatment as soon as possible, which is precisely when the second-line treatment is least valuable. As the costs of reversing current treatment impacts fall, it becomes more feasible to provide the option-preserving treatment to these low-risk individuals later on.

Our basic framework of option valuation highlights potentially important macro-level strategies to improve social welfare through medical technologies. Research investments that focus on transforming irreversibility from complete to partial could generate large social benefits. Investments in the development of alternative treatments for future diseases are also important, but the return to such investments will depend on the degree of irreversibility.

The greatest obstacle to translating theory to practice is the intensive data requirement, which in some cases would require coordination across firms whose products might interact.

References

Aitchison, J. and Brown, J. (1957), *The Lognormal Distribution.*, Cambridge: Cambridge University Press.

Blanke, C. D., Demetri, G. D., *von* Mehren, M., Heinrich, M. C., Eisenberg, B., Fletcher, J. A. and *et* al. (2008), 'Long-term results from a randomized phase II trial of standard versus higher-dose imatinib mesylate for patients with unresectable or metastatic gastrointestinal stromal tumors expressing kit', *Journal of Clinical Oncology* **26**, 620–5.

Brito, J. B. and *de* Mello-Sampayo, F. (2005), 'The timing and probability of FDI: an application to US multinational enterprises.', *Applied Economics* .

Chakraborty, S. (2004), 'Endogenous lifetime and economic growth.', *Journal of Economic Theory* **116**, 119–37.

Claxton, K. (1999), 'The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies.', *Journal of Health Economics* **18**, 342–64.

Claxton, K., Eggington, S., Ginnelly, L., Griffin, S., McCabe, C., Philips, Z., Tappenden, P. and Wailoo, A. (2005), A pilot study of value of information analysis to support research recommendations for NICE., Technical report, Centre for Health Economics, CHE Research Paper 4, University of York.

Cohen, O. (2000), 'Antiretroviral therapy: time to think strategically.', *The Lancet* **132**, 320–22.

Contreras-Hernandez, I., Mould-Quevedo, J. F., Silva, A., Salinas-Escudero, G., Villasis-Keever, M. A., Granados-Garcia, V. and *et* al. (2008), 'A pharmaco-economic analysis of second-line treatment with imatinib or sunitinib in patients with advanced gastrointestinal stromal tumours.', *British Journal of Cancer* **98**, 1762–8.

de Mello-Sampayo, F., *de* Sousa-Vale, S. and Camões, F. (2011), 'Delaying the timing of offshoring low-skilled tasks.', *Economic Modelling* **27**, 951–58.

Dixit, A. (1989*a*), 'Entry and exit decisions under uncertainty', *Journal of Political Economy* **97**, 620–638.

Dixit, A. (1989*b*), 'Hysteresis, import penetration, and exchange rate pass-through', *Quarterly Journal of Economics* **104**, 205–228.

Dixit, A. (1993), 'The art of smooth pasting', *Fundamentals of Pure and Applied Economics* **55**.

Dixit, A. and Pindyck, R. (1994), *Investment under Uncertainty*, Princeton: Princeton University Press.

Dusansky, R. and Cagatay, K. (2010), 'Implications of the interaction between insurance choice and medical care demand.', *Journal of Risk and Insurance* **77**, 129–44.

Dusansky, R. and Ç. Koç (2007), 'The capital gains effect in the demand for housing.', *Journal of Urban Economics* **61**, 287–98.

Harrington, M. and Carpenter, C. (2000), 'Hit HIV-1 hard, but only when necessary.', *The Lancet* **355**, 2147–152.

Hau, A. (2008), 'When is a coinsurance-type insurance policy inferior or even giffen?', *Journal of Risk and Insurance* **75**, 343–64.

Hislop, J., Quayyum, Z., Elders, A., Fraser, C., Jenkinson, D., Mowatt, G., Sharma, P., Vale, L. and Petty, R. (2011), 'Clinical effectiveness and cost-effectiveness of imatinib dose escalation for the treatment of unresectable and/or metastatic gastrointestinal stromal tumours that have progressed on treatment at a dose of 400 mg/day: a systematic review and economic evaluation', *Health Technology Assessment* **15**.

Hoi, M. and Robson, A. J. (1981), 'Insurance as a Giffen good.', *Economics Letters* **8**, 47–51.

Huse, D. M., *von* Mehren, M., Lenhart, G., Joensuu, H., Blanke, C., Feng, W. and et al. (2007), 'Cost effectiveness of imatinib mesylate in the treatment of advanced gastrointestinal stromal tumours.', *Clinical Drug Investigation* **27**, 85–93.

Mabasa, V. H., Taylor, S. C., Chu, C. C., Moravan, V., Johnston, K., Peacock, S. and *et* al. (2008), 'Verification of imatinib cost-effectiveness in advanced gastrointestinal stromal tumor in British Columbia (VINCE-BC study).', *Journal of Oncology Pharmacy Practice* **14**, 105–12.

McDonald, R. and Siegel, D. (1985), 'Investment and the valuation of firms when there is an option to shut down.', *International Economic Review* **26**, 331–49.

NICE (2004), Gastrointestinal stromal tumours - imatinib (TA86), Technical report, National Institute for Health and Clinical Excellence (NICE), Available from http://guidance.nice.org.uk/TA86 (accessed May 2012).

Palmer, S. and Smith, P. (2011), 'Incorporating option values into the economic evaluation of health care technologies.', *Journal of Health Economics* **19**, 755–66.

Pennings, E. (2000), 'Taxes and stimuli of investment under uncertainty', *European Economic Review* **44**, 383–391.

Pindyck, R. (1988), 'Irreversible investment, capacity choice, and the value of the firm', *American Economic Review* **78**, 969–985.

Sculpher, M. and Claxton, K. (2005), 'Establishing the cost-effectiveness of new pharmaceuticals under conditions of uncertainty – when is there sufficient evidence?', *Value Health* **8**, 433–46.

Smith, R. (2007), 'Use, option and externality values: are contingent valuation studies in health care mis-specified?', *Health Economics* **16**, 861–69.

Teich, N., Hashizume, C. and Follador, W. (2009), 'Economic evaluation of sunitinib vs. imatinib in second line for gastrointestinal tumor (gist) in Brazil.', *Value Health* **12**(A496).

Wilson, J. S., Connock, M., Song, F., Yao, G., Fry-Smith, A., Raftery, J. and Peake, D. (2005), 'Imatinib mesylate for the treatment of patients with unresectable and/or metastatic gastrointestinal stromal tumours: systematic review and economic evaluation.', *Health Technology Assessment* **9**(25).

Yu, C.-F., Chang, T.-C. and Fan, C.-P. (2006), 'FDI timing: Entry cost subsidy versus tax rate reduction', *Economic Modelling* **24**, 262–271.

Zalcberg, J. R., Verweij, J., Casali, P. G., Cesne, A. L., Reichardt, P., Blay, J. Y. and *et* al. (2005), 'Outcome of patients with advanced gastro-intestinal stromal tumours crossing over to a daily imatinib dose of 800 mg after progression on 400 mg', *European Journal of Cancer* **41**, 1751–7.

Zivin, J. G. and Neideill, M. (2009), 'Medical technology adoption, uncertainty, and irreversibilities: Is a bird in the hand really worth more than two in the bush?', *Health Economics* **19**, 142–53.

Appendix

A Proofs

Proof: (*Proposition 1*)
We assume that patients initiated on T_1 will receive that treatment for the rest of their life and the same will be true for patients initiated on T_2. As such, the decision maker will calculate the expected present discounted value (PDV) of each treatment regime and select the one with the highest return. Thus, the NB of using T_1 is given by the intertemporal NB function when the patient is constantly optimising over time:

$$NB^*_{1t} = E_t \int_0^\infty NB_{1t} e^{-\mu(\tau - t)} d\tau. \tag{15}$$

Bearing in mind that $x(t)$ follows a geometric Brownian motion, the properties of the log-normal distribution[11] can be used to transform Equation (15) into:

$$NB^*_{1t} = \frac{NB_{1t}}{\mu - \alpha\delta - \frac{\delta(\varphi-2)}{2\varphi}\sigma^2}, \tag{16}$$

provided that $\mu - \alpha\delta - \frac{\delta(\varphi-2)}{2\varphi}\sigma^2 > 0$, which will be assumed here.

The same applies to the NB of using T_2, i.e.:

$$NB^*_{2t} = E_t \int_0^\infty NB_{2t} e^{-\mu(\tau - t)} d\tau = \frac{NB_{2t}}{\mu - \alpha\delta - \frac{\delta(\varphi-2)}{2\varphi}\sigma^2}, \tag{17}$$

provided that $\mu - \alpha\delta - \frac{\delta(\varphi-2)}{2\varphi}\sigma^2 > 0$, which will be assumed here.

T_2 will be the preferred option since $NB^*_{2t} > NB^*_{1t}$. For those with $NB^*_{2t} < NB^*_{1t}$, T_1 will be preferred.

Proof: (*Proposition 2*)
In the continuation region the Bellman equation is given by:

$$\mu V^1(x,t) = NB^*_{1t} + \frac{1}{t} E[dV^1], \tag{18}$$

where NB^*_1 is given by Equation (15).

Applying Ito's lemma to the RHS of Equation (18) yields the partial differential equation:

$$\frac{\partial V^1}{\partial x}\alpha x + \frac{1}{2}\frac{\partial^2 V^1}{\partial x^2}\sigma^2 x^2 - \mu V^1 + NB^*_{1t} = 0 \tag{19}$$

The general solution to the above-stated differential equation has the form:

$$V^1 = B_1 X^{\beta_1} + B_2 X^{\beta_2} + NB^*_{1t}, \tag{20}$$

where B_1, B_2 are constants that are yet to be determined, and β_1, β_2 are the roots of the characteristic equation $Q(\beta) = \frac{\sigma^2}{2}\beta^2 + (\alpha - \frac{\sigma^2}{2})\beta - \mu = 0$. The two roots are $\beta_1 = \frac{1}{2} - \frac{\alpha}{\sigma^2} + \sqrt{(\frac{\alpha}{\sigma^2} - \frac{1}{2})^2 + \frac{2\mu}{\sigma^2}}$ and $\beta_2 = \frac{1}{2} - \frac{\alpha}{\sigma^2} - \sqrt{(\frac{\alpha}{\sigma^2} - \frac{1}{2})^2 + \frac{2\mu}{\sigma^2}}$. Since the coefficient of β^2 is positive, $Q(\beta)$ is an upward pointing parabola. Moreover, since $Q(1) = \alpha - \mu$ and $Q(0) = -\mu$ are both negative by previous assumption, it follows that $\beta_1 > 1$ and $\beta_2 < 0$.

[11] See Aitchison and Brown (1957).

In order to satisfy the boundary condition[12], $V^1(0) = 0$, we must have $B_2 = 0$ and the general solution to Equation (19) simplifies to

$$V^1 = B_1 x^{\beta_1} + NB^*_{1t}, \tag{21}$$

where NB^*_{1t} is now given by Equation (16).

The set of boundary conditions that applies to this optimal stopping problem is composed of a value-matching condition,

$$V^1(\widetilde{x}) = V^2(\widetilde{x}), \tag{22}$$

and one smooth-pasting condition,

$$[V^1(\widetilde{x})]' = [V^2(\widetilde{x})]'. \tag{23}$$

Letting $\widetilde{t}$ denote the time at which T_2 will be used, the value of the optimal treatment switch, $V^2(\widetilde{x})$, is given by the NB after using T_2, see Equation (24), minus the incremental cost per quality-adjusted life-year (IC_{QALY}) gain of using a higher dosage of T_1, i.e. value of QALY wasted in T_1 dose escalation.

$$V^2 = E_t\{\int_{\widetilde{t}+1}^{\infty} NB_2 e^{-\mu(\tau-t)}d\tau - \int_0^{\widetilde{t}} IC_{QALY} e^{-\mu(\tau-t)}d\tau\} = \frac{NB_2 e^{-(\widetilde{t}+1)\mu}}{\mu - \alpha\delta - \frac{\delta(\varphi-2)}{2\varphi}\sigma^2} - (1-e^{-\widetilde{t}\mu})\frac{IC_{QALY}}{\mu}. \tag{24}$$

Making use of the value-matching and smooth-pasting conditions, the expression for the critical value ($\widetilde{x}$) is obtained and likewise for the constant B_1 as:

$$\widetilde{x} = [\frac{\beta_1}{\beta_1 - \delta} \times \frac{\mu - \alpha\delta - \frac{\delta(\varphi-2)}{2\varphi}\sigma^2}{\phi^{\frac{1-\varphi}{\varphi}}[e^{-\mu(\tilde{t}+1)}] - \pi^{\frac{1}{\varphi}}} \times \frac{\pi^{\frac{1}{\varphi}}\theta^{\frac{1-\theta}{\theta}}(1-\theta)\tau}{\varphi(1-\varphi)^{\frac{1-\varphi}{\varphi}}} \times \frac{IC_{QALY}}{\mu}(1-e^{-\mu\tilde{t}})]^{\frac{1}{\delta}}, \tag{25}$$

$$B = \frac{\delta}{\beta}\theta^{\frac{1-\varphi}{\varphi}}[\frac{\beta}{\beta-\delta}\theta^{\frac{1-\varphi}{\varphi}}\frac{IC_{QALY}}{\mu}(1-e^{-\mu\tilde{t}})]^{1-\frac{\beta}{\delta}}(\frac{1}{\mu - \alpha\delta - \frac{\delta(\varphi-2)}{2\varphi}\sigma^2}\frac{\phi^{\frac{1-\varphi}{\varphi}}\varphi(1-\varphi)^{\frac{1-\varphi}{\varphi}}}{(1-\theta)\tau}\frac{[e^{-(\widetilde{t}+1)\mu}] - \pi^{\frac{1}{\varphi}}}{\pi^{\frac{1}{\varphi}}})^{\frac{\beta}{\delta}}. \tag{26}$$

Equation (25) is the trigger value of demand separating the region in x space where the patient's option of using T_2 remains unexercised (i.e. for $x > \widetilde{x}$) from the one where immediate exercise of that option is perceived as optimal (i.e. for $x \leq \widetilde{x}$).

B Data

The simulations relate to the critical ratio obtained in Equations (12), (13) and (14). The time horizon for the model was limited to 10 years since the typical survival of GIST patients is relatively short. The input choices and the results of sensitivity analyses are presented in Table 1, below.

The parameters from the equations of the critical ratio are defined as:

σ: Volatility of cost, is the standard deviation of the natural logarithm of cost of imatinib at dose of 400 mg/day for the two years prior to dose escalation. The input values were based on the Novartis model by Wilson et al. (2005).

[12]If x reaches zero, the NB will stay at zero afterwards, in which case the value of patient's health will have no value.

α: Trend of the cost is the moving average of the natural logarithm of cost´s rate of growth of imatinib at dose of 400 mg/day for the two years prior to dose escalation. The input values were based on the Novartis model by Wilson et al. (2005).

μ: The costs and outcomes were discounted at 3.5% in accordance with the NICE reference case.

φ: Since the data on the parameter for the elasticity of demand are not available, the benchmark value and the range of variation were picked arbitrarily.

τ: Tax on income is the United Kingdom income tax, basic rate of 20%, and the range of variation were the taxable bands in accordance with the HM Revenue & Customs, http://www.hmrc.gov.uk/.

π: Since the data on the parameter for the relative benefit of T_2 are not available, the benchmark value and the range of variation were picked arbitrarily.

ϕ: Since the data on the parameter for the relative cost of T_2 are not available, the benchmark value and the range of variation were picked arbitrarily.

IC_{QALY}: The estimated incremental cost of quality-adjusted life-year (QALY) gained from the pathway of increasing imatinib dose from 400 mg to 600 mg/day estimated by (Hislop et al., 2011) at £27,304.

Table 1 presents the range as well as the mean of each parameter according to the data set specified above. The baseline values are used to define the benchmark case while the maximum and the minimum values bound the range used for the simulations of the critical ratio.

Table 1: Parameter Values

	Maximum	**Minimum**	**Baseline**
α	0.05	0.01	0.025
σ	0.12	0.15	0.135
μ	0.035	0.015	0.06
τ	0.5	0.1	0.2
θ	0.99	0.01	0.5
φ	0.99	0.01	0.5
ϕ	0.99	0.01	0.5
π	0.99	0.01	0.5
t	10	1	5
IC_{QALY}	-	-	27,304

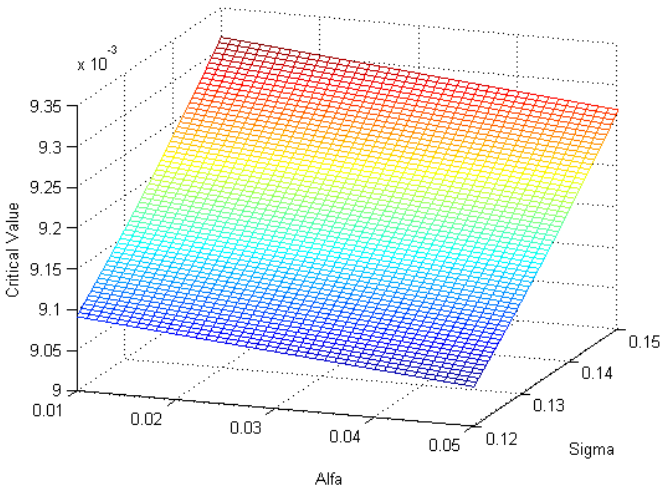

Figure 1

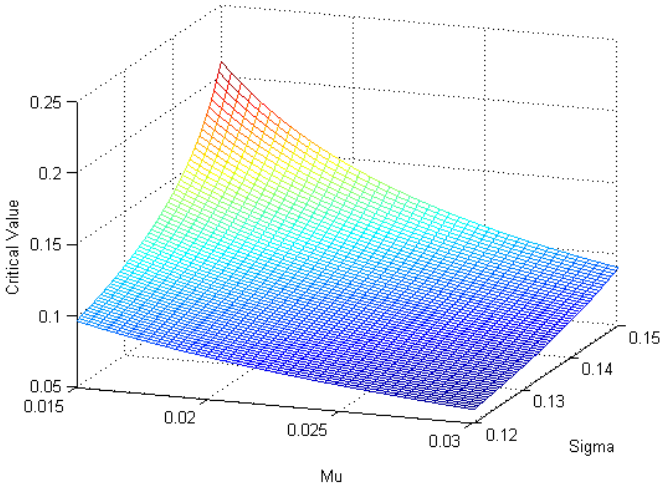

Figure 2

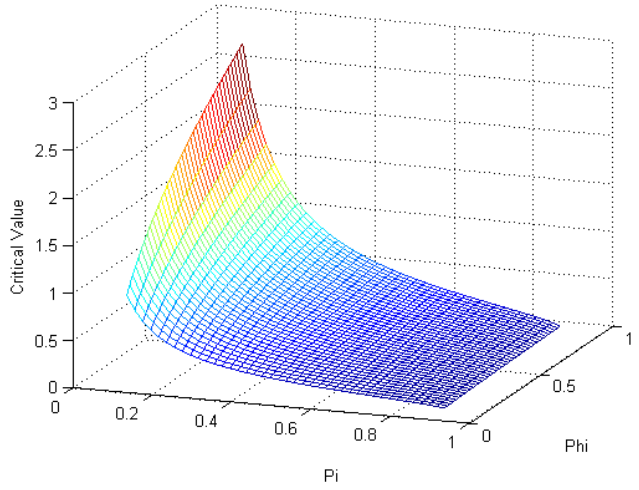

Figure 3

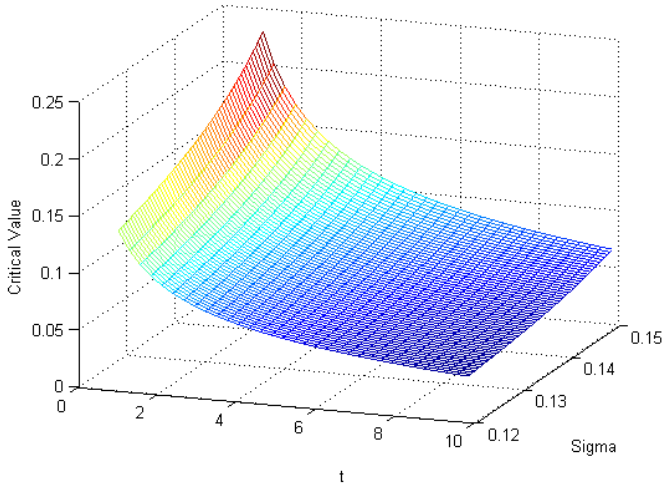

Figure 4

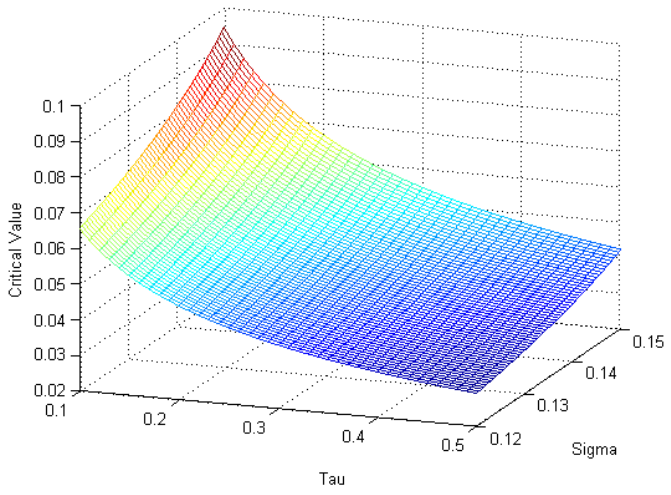

Figure 5

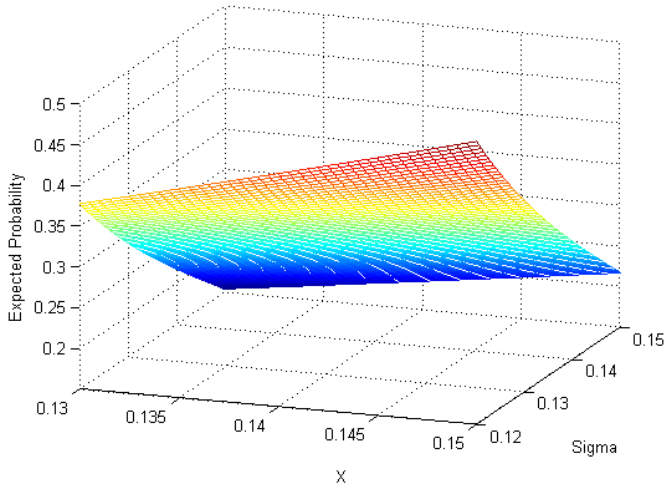

Figure 6

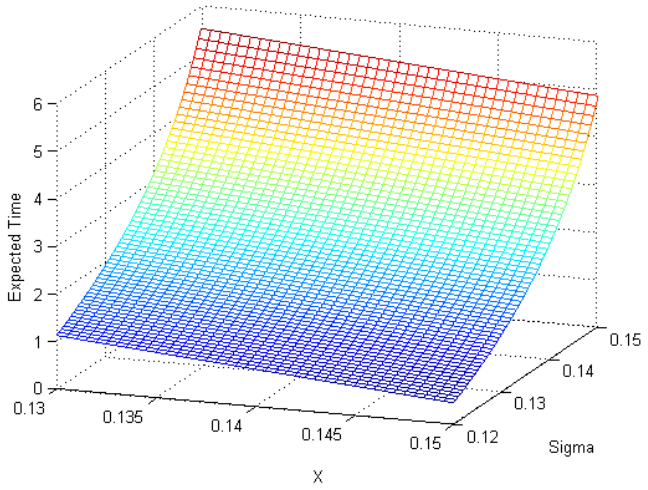
Expected Time
6
5
4
3
2
1
0
0.13
0.135
0.14
0.145
0.15
X
0.12
0.13
0.14
0.15
Sigma

Figure 7