Economic Theory **25**, 317–332 (2005)
DOI: 10.1007/s00199-003-0439-3

Economic Theory

Costly switching and investment volatility*

Tackseung Jun

School of Economics and International Trade, Kyung Hee University, 1 Hoegi-dong, Dongdaemoon-gu, Seoul 130-701, SOUTH KOREA (e-mail: Tj32k@hanmail.net)

Received: November 14, 2000; revised version: October 9, 2003

Summary. This paper extends the literature on the optimal switching rule between two investments by considering the case where switching between investments is costly. The model builds on the classic framework of the multi-armed Bandit problem by explicitly incorporating two key assumptions. First, switching investments is costly. Second, only the investment operated by the investor evolves as a random walk. The objective of the investor is to maximize the discounted sum of expected net profits over the infinite horizon. The main result is that when the volatility of profits from investments increases, so does the minimum profit gain needed for an investor to switch investments.

Keywords and Phrases: Switching cost, Optimal switching rule, Random walk, Volatility.

JEL Classification Numbers: C44, C61.

1 Introduction

Decisions to switch investments are largely driven by the prospects of higher profits.[1] As we commonly observe, however, firms often continue to operate in markets that they had entered under more favorable circumstances even after conditions deteriorate. Conversely, they are loath to re-enter seemingly profitable markets that

* I am indebted to Prajit K. Dutta for his guidance throughout this research. I am grateful for President's Summer 2000 Research Fellowship of Columbia University. I appreciate the comments from the anonymous referee. I also thank Lalith Munasinghe and Rajiv Sethi. I also thank Dr. Jong Myeon Kim for editing this version of the paper.

[1] The term "investment" in the paper refers to the profit-generating opportunity in general.

they had been forced to abandon previously. In either case, firms often do not react immediately to changes in market conditions. The question is then what causes this delay in switching and how long do such delays last?

If an alternative investment is sufficiently better than the current investment *and* the gain in profits from switching investments is expected to continue for a sufficiently long period, switching investments may indeed be profitable. However, even if the alternative investment promises greater profits, the firm may decide not to switch away from its current investment if the relative profitability is likely to be reversed shortly. This consideration suggests that the value of delaying a switch is affected by the volatility of profits from investments: greater volatility in the profits from investments induces the investor to assign a higher value on the option to delay the decision to switch investments.[2]

The model presented here builds on the classic framework of the multi-armed Bandit problem. The basic structure of the problem involves two statistically independent processes and a decision maker, who at each time t, selects one process. The selected process yields an immediate reward and changes its state, while the other processes do not change state and yield no reward at t. The objective of the decision maker is to select a process at each time so as to maximize the stream of the expected profits.

Two assumptions are explicitly incorporated into this classic framework to set up my model. First, I posit two investments, the profits from which evolve as random walk processes with drift. Second, switching investments entails irrevocable switching costs.[3]

Without switching costs, it is well-known that the Gittins Index rule is optimal,[4] where the index of an investment is defined as the maximal attainable discounted profit rate of the investment. According to the Gittins Index rule, an investor will choose whichever investment with the highest index. This rule is relatively simple to implement because the index of an investment, which depends only on the properties of the corresponding investment, enables the n-dimensional problem to be reduced to a one-dimensional problem, where n is the number of investments.

Banks and Sundaram (1994) show that this index rule is not optimal when switching incurs costs. According to their argument, there are two requirements

[2] This implication is based on the results from the financial option pricing theory (Dixit and Pindyck 1994). When an investor makes an irreversible investment, he basically exercises his option to avoid the downside risk and to realize the upside potential. This irreversible decision must consider this loss in the option for delaying. The volatility of the future return is simply the spread in the density of the option, and so the higher volatility increases the value of the option.

[3] Other Bandit models with switching costs include Agrawal et al. (1990) and Asawa and Teneketzis (1996). Agrawal et al. (1990) consider the two cases for the distribution of the profits obtained from each investment: the i.i.d. process and the Markovian process. In either case, profits are assumed to follow a joint distribution with unknown parameters. Besides the assumptions of unknown parameters, their model differs from mine in that their switching rule is optimal only asymptotically in the sense that achieves the same asymptotic performance as the optimal rules for the Bandit problem without switching cost. Asawa and Teneketzis (1996) consider the case where each reward process is a deterministic sequence.

[4] Gittins and Jones (1974) show the optimality of this rule. See also Gittins (1989) and Banks and Sundaram (1992).

for an index rule to be optimal: if there is a possibility of switching back to the current investment, the index of the current investment must depend on the switching cost, while if there is no such possibility, the index should be independent of the switching cost. As a result, the index of the current investment will differ, depending on whether we compare it to the investment whose profit is fixed and known for sure, or compare it with the investment whose profit is evolving and thus uncertain. They show that these two requirements for the optimality of an index rule are not *mutually* consistent.[5]

This paper contributes to the literature by characterizing the optimal switching rule in the presence of switching costs. Without switching costs, we revert to the standard Bandit result that, holding everything else constant, investors will switch investments whenever an alternative investment offers higher profits than their current investment. In contrast, the presence of switching costs demands that the gain in profits from switching must be sufficiently large, i.e., above a certain positive threshold, before investors will find the decisions to switch justifiable.

A more interesting result is that when the volatility of investment profits increases, so does the minimum profit gain needed for an investor to switch investments. The intuition for this result is as follows. With greater volatility of profits, profits from investments will be spread higher and lower than they would have been at lower profit volatility. With higher *levels* of investment profits, current and alternative investments will both deliver higher profits and thus leave the incentive to switch unaffected. In contrast, greater *volatility* of investment profits implies a greater chance that the ranking of profits can be reversed. This creates a greater downside risk for both current and alternative investments. However, by staying with her current investment, the investor can avoid exposing herself to the downside risk from switching investments. This raises the value of the option for delaying, thus requiring a higher gain in profits that can justify switching.

The balance of the paper is organized as follows. Section 2 describes the model, for which the optimal switching rule is derived in Section 3. Section 4 covers the comparative statics and a discussion that compares my results with those from the learning models of the Bandit problem. Section 5 closes the paper by considering possible extensions to my model. All proofs are provided in the Appendix.

2 The model

Assume that the investor is currently holding an investment, denoted by B. Assume also that the investor is aware of an alternative, "outside" investment, denoted by A, from her prior search of an outside distribution of investment opportunities. Let us refer to the profits from the current and outside investments as the current and outside profit, respectively. For analytical simplicity, assume that all investments evolve as symmetric random walk processes with drift. Formally, we define the evolution of profits from investments as follows.

[5] Bergemann and Valimaaki (2001) show that it is possible to define optimal index polices in the Bandit problems with switching costs if one restricts attention to the stationary Bandit setting (see Theorem 2 in Bergemann and Valimaaki, 2001).

Definition 1. *For each investment $i \in \{A, B\}$, let the incremental process $\{x_t\}$ denote an i.i.d. sequence with the following transition property:*

$$\Pr(x_t = h) = p, \;\; \Pr(x_t = -h) = 1 - p$$

for $p \in [0, 1]$ and $h > 0$. The profit from investment i at time $t + 1$ is given by

$$s^i_{t+1} = \begin{cases} s^i_t + x_t & \textit{if investment } i \textit{ is chosen at time } t \\ s^i_t & \textit{otherwise} \end{cases} \tag{1}$$

At any given time t, the investor observes the current and outside profit and then decides whether to switch to the outside investment or stay with the current investment. If the investor decides to switch investments then she must incur a fixed and irrecoverable cost C. The investor is assumed to be risk neutral – i.e., utility is a linear function of profits – and infinitely lived. All the underlying primitives of the model are known to the investor and so there is no learning on the side of the investor. An investor chooses $\theta_t \in \Theta = \{A, B\}$ at every period.[6] The switching rule is $\pi : \{s^i_t\}_{i \in \Theta} \times \theta_{t-1} \to \theta_t$ for all t. So a rule generates the sequence of the choices $\{\theta_t\}_{t=0}^{\infty}$. The objective of an investor is to maximize the sum of the expected present-value of net profits over the infinite horizon, where the net profit at any given time is the profit from the chosen investment minus the switching cost (conditional on an investment switch):

$$\max_{\{\theta_t\}} E \sum_{t=0}^{\infty} \beta^t \left(s^{\theta_t}_t - 1_{\{\theta_t \neq \theta_{t-1}\}} C \right), \tag{2}$$

where $\beta \in (0, 1)$ is the discount factor and 1_φ is equal to one iff φ holds.

The following example provides some insights into the source of randomness.[7] Suppose there are two ore mines and an exploration company can explore either one of the mines at any given time. As mining proceeds at a mine, the properties of the vein of ore may change. This can affect the quality of the ore positively or negatively, which in turn will affect its market value. Therefore, ultimately, profits from a mine are subject to random changes in the quality of the ore.[8]

The Bellman representation of the model presented here is defined as follows:

$$v_B\left(s^A, s^B\right) = \max\left\{s^A - C + \beta W_A, \; s^B + \beta W_B\right\}, \tag{3}$$

[6] θ_0 is given exogenously.

[7] I thank an anonymous referee for suggesting this example.

[8] In the literature, most investment models deal with stochastic output prices. Dixit (1989) considers a model of investment by a firm where uncertainty comes from the market price of the output. Brennan and Schwartz (1985) consider a model of investment into mine where the spot price of ore follows a stochastic process. Other models of price uncertainty include Pindyck (1988, 1991), and Lippman and Rumelt (1985). These models are concerned with the evaluation of a single project where price uncertainty can be easily treated exogenously. However, with more than one market (or project), the effect of a market operation on output prices is much more sensitive to the overall structure of markets, such as what happens to the output price in a market when an investor is involved in another market? Investigating such complications would require a model with a richer structural construct than my model. There are, however, models of multiple projects with price uncertainty. For example, He and Pindyck (1992) consider a multi-output firm whose demands in two markets evolve simultaneously.

where $v_B\left(s^A, s^B\right)$ is the value function for an investor whose current investment B has the profit of s^B and the outside investment A has the profit of s^A. In general, let $v_i\left(s^A, s^B\right)$ denote the value function for an investor who currently holds investment i, given s^A and s^B. The continuation value of each investment is:

$$W_A \equiv pv_A\left(s^A+h, s^B\right)+(1-p)\, v_A\left(s^A-h, s^B\right),$$
$$W_B \equiv pv_B\left(s^A, s^B+h\right)+(1-p)\, v_B\left(s^A, s^B-h\right).$$

Since all investments evolve as symmetric random walk processes, conditional on initial profits, all investments are identical to the investor. Therefore, given switching cost, the investor only needs to evaluate the difference in profits between the outside and current investment to decide whether or not to switch. This implies that there is a *reservation* value in the gain in profits from switching investments such that whenever the gain in profits from switching investments exceeds this threshold value the investor will switch investments. This reservation value will depend on switching cost and, more importantly, the *volatility* of profits from investments. The key point is that the optimal switching rule does not depend on the *level* of profits from investments. This intuition is formally developed in the next section.

3 Optimal switching rule

Although the optimal switching rule may not be as simple as the index rule, one can get a sense for it from the following observation. Fix $q \in \mathbb{R}$, and consider the difference $v_i\left(s^A+q, s^B+q\right) - v_i\left(s^A, s^B\right)$. Let π_0 be the optimal rule for $v_i\left(s^A, s^B\right)$ and π_q be the optimal rule for $v_i\left(s^A+q, s^B+q\right)$. Consider applying the rule π_0 to both pairs of $\left(s^A, s^B\right)$ and $\left(s^A+q, s^B+q\right)$. Clearly, by following π_0, the sum of the present-discounted net profits from any sequence of realizations beginning with $\left(s^A+q, s^B+q\right)$ is higher than that beginning with $\left(s^A, s^B\right)$ by at least $q/\left(1-\beta\right)$. Since π_0 is not necessarily optimal for $\left(s^A+q, s^B+q\right)$, we get:

$$v_i\left(s^A+q, s^B+q\right) \geq v_i\left(s^A, s^B\right)+\frac{q}{1-\beta}. \tag{4}$$

Repeating the above exercise, the application of π_q to both $\left(s^A, s^B\right)$ and $\left(s^A+q, s^B+q\right)$ yields:

$$v_i\left(s^A+q, s^B+q\right)-\frac{q}{1-\beta} \leq v_i\left(s^A, s^B\right). \tag{5}$$

This is because, by following π_q, the present value of the net profit stream from $v_i\left(s^A+q, s^B+q\right)$ will be greater than that from $v_i\left(s^A, s^B\right)$, but by no more than $q/\left(1-\beta\right)$. Inequalities (4) and (5) imply:

$$v_i\left(s^A+q, s^B+q\right)-v_i\left(s^A, s^B\right)=q/\left(1-\beta\right) \text{ for all } q \text{ and } i.$$

This implies that if an investor is indifferent over switching now and waiting at a particular difference of profits, he will remain indifferent as long as the difference in profits from investments remains the same. Hence *the optimal switching rule is* *a uniform threshold type*.[9] Here on, we search for the minimum profit gain needed for an investor to switch investments (or equivalently the optimal switching rule).[10] Let s denote the minimum profit gain needed for an investor to switch investments. The following example provides an heuristic explanation of the minimum profit gain in the general case.

Definition 2. *Investment $i \in \Theta$ is monotonically increasing iff*

$$s_{t+1}^i = \begin{cases} s_t^i + h & \text{if investment } i \text{ is chosen at time } t \\ s_t^i & \text{otherwise} \end{cases}$$

Example 1. If both investments are *monotonically increasing*, then the minimum profit gain needed for an investor to switch investments is equal to $C(1-\beta)$.

Example 1 can be understood as follows. Suppose that, given the current investment B, the profit from investment A is higher than the profit from investment B by the minimum profit gain s. Normalize the profit from investment B to zero, and further assume that the investor had chosen investment B in the previous period. If an investor switches to investment A, it is optimal to continue with it forever because the profit from investment A continues to increase by h each period (thus increasing the gap with investment B), earning $\sum_{i=0}^{\infty}(s+ih)\beta^i$. By the same reason, continuing with investment B forever is also optimal, earning $\sum_{i=0}^{\infty} ih\beta^i$. Hence the present-discounted value of the difference in profits from investments is $s/(1-\beta)$. By the definition of indifference, this difference should be balanced by the cost of switching C.

Now consider the general case where evolution of states $\{s_t^i\}$ is defined in Definition 1. As argued earlier, the process Δ_t, which is defined as $s_t^A - s_t^B$, can fully characterize the optimal switching rule.

Theorem 1 (Optimal switching rule). *The solution to the minimum profit gain s is given by:*

(1.1) $s = C(1-\beta) + h\lambda \frac{1-\lambda^{2s/h}}{\left(1+\lambda^{2s/h+1}\right)(1-\lambda)}$ *if $2s \geq h$ and $2s/h$ is an integer.*

[9] The fact that the optimal switching rule is fully characterized by the profit difference is in contrast with the results from learning models (Rothschild, 1974; McLennan, 1984; Easley and Kiefer, 1988; Rustichini and Wolinsky, 1995; Keller and Rady, 1999) where profitabilities of investments are unknown a priori. In learning models, there is a trade-off between the informational benefits and the short-run profits. This trade-off will divide the strategy space into two distinctive regions. In one region, it is optimal to do extensive experimentation, largely deviating from the myopic strategy, to resolve the uncertainty about profitabilities of investments. In the other region, experimentations are rare and myopic strategies are chosen mostly. The first strategy sacrifices the short-run reward, while the second one loses the valuable information that may be used to improve future profits. When the information acquisition is more beneficial than the short-run profits, the degree of experimentation is non-negligible, and so the optimal switching rule is not fully explained by the difference in the expected rewards from alternative investments. In this paper, there is no concern about learning. Hence only the gain in profits matters in the decisions to switch investments.

[10] The threshold rule itself is not new in the literature. See, for example, Balvers and Cosimano (1993) and Keller and Rady (1999).

(1.2) $s = C(1-\beta)+((s-lh)+C(1-\beta))\frac{\lambda^{l}(1+\lambda)}{1+\lambda^{2l+1}}+\frac{h(1-\lambda^{2l})\lambda}{(1+\lambda^{2l+1})(1-\lambda)}$ *if* $2s \geq h$ *and* $2s/h$ *is not an integer and where* $2s/h < l < 2s/h + 1$.

$$\lambda = \frac{2\beta(1-p)}{1+\sqrt{1-4\beta^{2}p(1-p)}} \tag{6}$$

Here a sketch of the proof strategy is outlined. First, consider an investor who switches investments whenever the gain in profits is at least as large as some minimum profit gain. Second, we can compute the sum of the expected present-discounted net profits from this threshold rule. This enables us to calculate the sequences of switching times, and the sum of the present-discounted profits for each completed spell on an investment. Third, for any arbitrary threshold difference in profits, we compute: (I) the expected profits from *switching* now and then following some threshold rule, and (II) the expected profits from *staying* now and then following the same threshold rule. Finally, the minimum profit gain s is defined and solved for the case where (I) and (II) are set to be equal.

Note that in Theorem 1 if $p = 1$ then $s = C(1-\beta)$, which confirms Example 1. $C(1-\beta)$ is actually the minimum among all the minimum profit gains needed for an investor to switch investments. Intuitively, in Example 1, the number of times an investor needs to switch investments is minimized when $p = 1$ and so the continuation profit from switching investments is largest. Hence an investor will switch investments as long as the present-discounted cost of switching, $C(1-\beta)$, is covered.

4 Comparative statics

Below, I show how the minimum profit gain s depends on the primitives of the model. It seems to be intuitive that the minimum profit gain should be larger as switching cost is larger, since an investor should be compensated with a better future stream of profits when switching costs more. However, it is not unambiguous regarding how p and h affect the optimal switching rule.

Theorem 2 (Comparative statics). *The minimum profit gain s in Theorem 1 (1) increases in C, (2) increases in h, and (3) decreases in p.*[11]

As argued earlier, a higher h means that the ranking of profits from investments is more likely to alternate, which increases the value of the option for delaying decisions to switch investments. Hence an investor will only switch investments if he gets a higher gain in profits by doing so. This implies that the minimum profit gain needed for an investor to switch investments increases in h.

Deriving the dependency of the minimum profit gain on β is not a trivial exercise since β appears multiplicatively with switching cost, the size of which can be arbitrary. Therefore simulations are shown in Figures 1 and 2, which are drawn by

[11] The proof of Theorem 2 is available from the author, and it is omitted because it requires only differentiations.

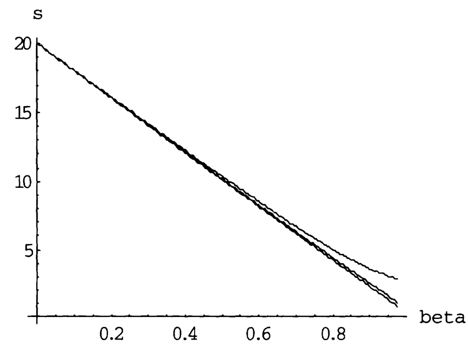

Figure 1. s as a function of β, shown (from above) for $p = 0.8, 0.7$ and 0.5

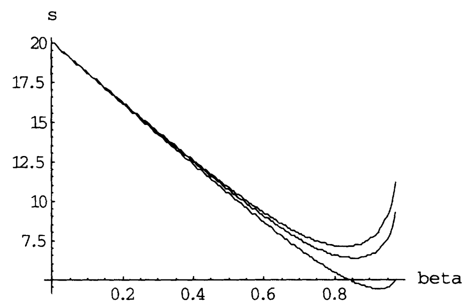

Figure 2. s as a function of β, shown (from above) for $p = 0.1, 0.2$ and 0.4

parameterizing the solution found in Theorem 1.[12] Figure 1 shows that the minimum profit gain decreases in β for $p \geq 1/2$. In contrast, when $p < 1/2$, Figure 2 shows that the minimum profit gain follows a J-shaped curve – as before, s declines in β but turns back and increases for sufficiently high values of β. One explanation for this observation is that a really patient investor values the continuation profit so much that he will wait for a really good realization of the profit from the alternative investment when investments are likely to deteriorate.

In the rest of this section, I examine the frequency of switching expected to be observed. The absolute size of the threshold is not sufficient to tell us how frequently an investor who follows the optimal switching rule switches investments. One measure for the frequency is the eventual probability of the next switch if a switch is made today, or equivalently, the *first passage time* to reach the difference in profits of $2s$ from the difference in profits of zero.[13] Let $\mathcal{T}_{2s}$ denote this stopping time. To derive $\mathcal{T}_{2s}$, I construct the moment generating function of the probability of the first passage time as follows:

$$\mathcal{F}_{2s}(t) = \sum_{i=1}^{\infty} t^i \Pr(\mathcal{T}_{2s} = i).$$

[12] Figures 1 and 2 assume $C = 20$ and $h = 1$. The results are robust to other parametric specifications.

[13] Note that if an investor just started to operate an investment today, the profit from current investment needs to decline by $2s$ before the investor will switch to the alternative investment.

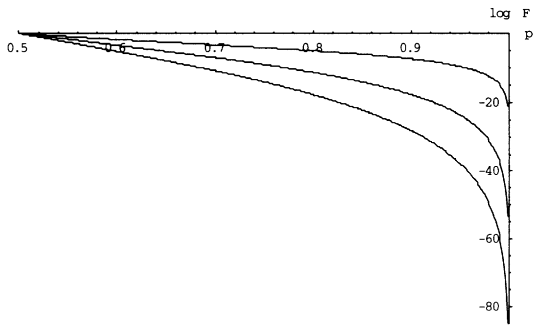

Figure 3. Frequency of switching as a function of p

Hence the probability that a switching occurs at all is $\mathcal{F}_{2s}(1)$:[14]

$$\mathcal{F}_{2s}(1) = \Pr(\mathcal{T}_{2s} < \infty) = \begin{cases} \left(\frac{1-p}{p}\right)^{\frac{2s}{h}} & \text{for } p > 1/2 \\ 1 & \text{for } p \leq 1/2. \end{cases} \tag{7}$$

Since switching is a sure event for $p \leq 1/2$,[15] the interesting case is when $p > 1/2$. As p increases, both $(1-p)/p$ and $2s/h$ decrease. Hence the sign of $\mathcal{F}_{2s}(1)$ is not unambiguous. I use simulations to see how $\mathcal{F}_{2s}(1)$ changes as p increases in $(1/2, 1]$. Since $\mathcal{F}_{2s}(1)$ can be very small, I take the natural logarithm of $\mathcal{F}_{2s}(1)$. A simulation result is shown in Figure 3:[16] as p increases from $1/2$, this eventual probability of switching decreases, namely, the effect of $(1-p)/p$ dominates that of $2s/h$. This implies that as p increases in $(1/2, 1]$, although an investor is *more* willing to switch (according to Theorem 2), the investor will switch *less* in the probabilistic sense.

Notice from (7) that when $p > 1/2$ the duration of the random walk is indefinite with a positive probability, and the walk may never reach its absorbing barrier, where the barrier in my paper is simply the minimum profit gain needed for an investor to switch investments. In other words, it is probable for an investor not to have a chance to switch and thus stay with the current investment forever. This is because, when investments drift toward improvement of the current profits, there exists, with a positive probability, a sequence of realizations that leads the profit from current investment to stay continuously above the threshold, below which it is optimal to switch investments.[17] This resonates with the popular result in the optimal learning literature that a decision maker can be locked-in to one alternative forever.[18] These

[14] The details of the derivation can be found in Cox and Miller (1964, p. 37). The underlying idea is to calculate the first hitting time to reach the target distance by setting the absorbing barrier at the target profit.

[15] For $p \leq 1/2$, the optimal switching will occur in a finite time with probability one.

[16] In Figure 3, I assume that $C = 8$ and $h = 1$. Curves are, from top to bottom, for β equal to 0.8, 0.5 and 0.2.

[17] Consult pp. 190–191 in Rothschild (1974) for an heuristic argument on the possibility of ending up with one choice forever in the Bandit problem.

[18] See, for example, Banks and Sundaram (1992) and Keller and Rady (1999) for such results in the learning model context.

two results share the Martingale property: it is possible to expect the current state will prevail in the future when all the possible realization of states are considered. This leads the belief of the decision maker to stay in a particular region of the belief space forever with a positive probability, consequently choosing a particular strategy over and over again. The difference is that, in this optimal switching under complete information, this lock-in is driven by a pair of random walk processes, which are given *exogenously* to the decision maker, while in learning models, it is the *endogenous* learning that can induce a decision maker to experiment with only one alternative under incomplete information.

5 Conclusions

In this paper, I derive the optimal switching rule in the Bandit problem when switching between alternatives is costly. The main result is that when the volatility of profits from investments increases, so does the minimum profit gain needed for an investor to switch investments.

The following discusses some possible extensions to my model. An interesting variation would be to examine the optimal switching rule in a finite horizon. Intuitively, the minimum profit gain needed for an investor to switch investments is expected to increase as the horizon gets shorter, since the expected gain from switching investments gets smaller due to a shorter horizon, and the investor therefore requires a higher gain in profits from switching investments. Another extension concerns the assumption that the investments are assumed to be symmetric in their transition dynamics. But, relaxing this assumption, I was not able to derive the optimal switching rule. With any type of asymmetry in investments, the optimal switching rule will differ according to whether the currently-held investment is A or B. An even more ambitious extension would be to allow an investor to operate two investments *simultaneously*. Switching there means changing portfolios of resources over two investments.

The model has wider applicability such as explaining job mobility decisions by a worker whose cost of switching jobs is non-negligible.[19]

Appendix

I present the proof of Theorem 1 for the case where $2s/h$ is an integer. The case where $2s/h$ is not an integer employs a similar argument and hence is omitted (the complete proof is available from the author). To find the minimum profit gain, I need to derive the differences $v_A(s+h,0) - v_B(s,h)$ and $v_A(s-h,0) - v_B(s,-h)$ (See Lemmas 1 and 2). The proof is done as follows. I will construct a sequence of optimal choices of investments, from which the difference between value functions are derived.

[19] Workers switching jobs entail a variety of costs: costs of learning new skills needed at the new job, or having their children adapt at a new school when the new job requires that the family re-locates geographically.

Let s^A and s^B be the initial profit from investment A and B, respectively. Suppose the initial difference in profits from investments is equal to Δ_0 at time 0. I assume that $\Delta_0 = s - h$. Hence an investor who currently holds investment B will stay with it until the difference in profits increases to at least $\Delta_\tau = \Delta_0 + h$ for the first time, where the random time τ is defined as follows:

$$\tau \in \{T > 0 | \Delta_j < \Delta_0 + h \text{ for all } 0 \leq j < T \text{ and } \Delta_T = \Delta_0 + h\}.$$

I am interested in the discounted sum of profits from investment B until time τ, starting with s^B. This sum of the profits can be decomposed into two components: the "base profit" and the remainder term. I define the base profit as the present value of the hypothetical profit stream that remains constant at the initial level of s^B until the random time τ. The remainder term is simply the weighted present-discounted sum of the difference of the actual profit streams from the hypothetical constant profit stream of s^B, using the probabilities of the different streams as the weights. By definition, the remainder term does not depend on the initial profit. For each $\tau \geq 1$, there are multiple sample paths that lead to the target profit difference $\Delta_0 + h$ at the random time τ. Let N_j denote the set of such sample paths where a random time j is the first hitting time to the target profit difference $\Delta_0 + h$, and n_j be the cardinality of this set. Fixing j, the base profit corresponding to a specific sample path is given by $s^B \sum_{i=1}^{j} \beta^i$ (there are n_j of such sums). Summing for all sample paths with initial profit s^B, we obtain the base profit, denoted by $G(s^B)$:

$$G\left(s^B\right) = s^B \left[1 + \sum_{j=0}^{\infty} p^j (1-p)^{j+1} n_{2j+1} \sum_{i=1}^{2j+1} \beta^i\right].$$

Now consider the remainder term, denoted by $F_{j,h}$, where j is the first hitting time to the target difference in profits of $\Delta_0 + h$ from Δ_0. I do not have an explicit form for this term. But, we do not need the explicit formula for this term. As it will be shown below, what we need to know about $F_{j,h}$ is that it does not depend on the initial state (by definition) nor a particular investment (by symmetry).

The value function $v_B\left(s^A, s^B\right)$ can be written as the sum of two components: first, the sum of the base profit and the remainder term before the profit from investment B declines to $s^B - h$ for the first time and second, the value function at the end of this "travel", namely $v_B\left(s^A, s^B - h\right)$:

$$\begin{aligned} & v_B\left(s^A, s^B\right) \\ &= G\left(s^B\right) + \sum_{j=0}^{\infty} p^j (1-p)^{j+1} F_{j,h} \\ &\quad + \sum_{j=0}^{\infty} p^j (1-p)^{j+1} n_{2j+1} \beta^{2j+1} v_B\left(s^A, s^B - h\right) \end{aligned} \tag{A1}$$

Since the optimal switching rule depends only on the difference in profits, we can translate the profits from each investment up by h, and still satisfy the optimality.

By doing so, we get:

$$v_B\left(s^A+h,s^B+h\right) \tag{A2}$$
$$=G\left(s^B+h\right)+\sum_{j=0}^{\infty}p^j(1-p)^{j+1}\left[F_{j,h}+n_{2j+1}\beta^{2j+1}v_B\left(s^A+h,s^B\right)\right].$$

The term $F_{j,h}$ is identical whether the initial profit is s^B or s^B+h. Hence subtracting $(A1)$ from $(A2)$, we get

$$v_B\left(s^A+h,s^B+h\right)-v_B\left(s^A,s^B\right) \tag{A3}$$
$$=h\left[1+\sum_{j=0}^{\infty}p^j(1-p)^{j+1}n_{2j+1}\sum_{i=1}^{2j+1}\beta^i\right]$$
$$+\lambda\left(v_B\left(s^A+h,s^B\right)-v_B\left(s^A,s^B-h\right)\right),$$

where

$$\lambda=\sum_{j=0}^{\infty}p^j(1-p)^{j+1}n_{2j+1}\beta^{2j+1}=\frac{2\beta(1-p)}{1+\sqrt{1-4\beta^2p(1-p)}}.$$

Notice that λ is nothing but the "discounted" moment generating function for the first hitting time that the profit from an investment declines by h from the profit at the previous hitting time. Since the optimal switching rule is fully characterized by the difference in profits, equation $(A3)$ can be reduced as follows:

$$\frac{h}{1-\beta}=h\left[1+\sum_{j=0}^{\infty}p^j(1-p)^{j+1}n_{2j+1}\sum_{i=1}^{2j+1}\beta^i\right]+\lambda\frac{h}{1-\beta},$$

which can be rearranged as follows:

$$1+\sum_{j=0}^{\infty}p^j(1-p)^{j+1}n_{2j+1}\sum_{i=1}^{2j+1}\beta^i=\frac{1-\lambda}{1-\beta}. \tag{A4}$$

Equation $(A4)$ is critical in deriving the optimal switching rule later. To find the minimum profit gain, the differences $v_A(s+h,0)-v_B(s,h)$ and $v_A(s-h,0)-v_B(s,-h)$ are derived in Lemmas 1 and 2 below.

Lemma 1. *Let $k=2s/h$. Then*

$$v_A(s+h,0)-v_B(s,h)=\frac{h\lambda\left(1-\lambda^k\right)}{(1-\beta)\left(1+\lambda^{k+1}\right)}+C.$$

Claim 1. With $k=2s/h$,

$$v_A(s+h,0)-v_A(s,0)=\frac{h}{(1-\beta)\left(1+\lambda^{k+1}\right)}.$$

Proof of Claim 1. The proof is done by constructing a sequence of optimal choices of investments, and then "decompose" the difference of $v_A(s+h,0) - v_A(s,0)$ according to this optimal sequence. For $v_A(s+h,0)$, suppose an investor starts with investment A and continues as follows:

$$\begin{aligned}
&\theta_t = A \text{ for } 0 \le t < \tau_1\\
&\quad \text{where } \left\{\tau_1 > 0 | s_t^A > -s \,\forall t < \tau_1, s_{\tau_1}^A = -s\right\},\\
&\theta_t = B \text{ for } \tau_1 \le t < \tau_1' \\
&\quad \text{where } \left\{\tau_1' > \tau_1 | s_t^B > -2s - h \forall \tau_1 \le t < \tau_1', s_{\tau_1'}^B = -2s - h\right\},\\
&\theta_t = A \text{ for } t = \tau_1'.
\end{aligned} \tag{A5}$$

Clearly, the sequence of choices in $(A5)$ satisfies optimality. For $v_A(s,0)$, suppose an investor starts with investment A and continues as follows:

$$\begin{aligned}
&\theta_t = A \text{ for } 0 \le t \le \tau_2\\
&\quad \text{where } \left\{\tau_2 > 0 | s_t^A > -s - h \,\forall t < \tau_2, s_{\tau_2}^A = -s - h\right\},\\
&\theta_t = B \text{ for } \tau_2 \le t < \tau_2' \\
&\quad \text{where } \left\{\tau_2' > \tau_2 | s_t^B > -2s - h \,\forall \tau_2 \le t < \tau_2', s_{\tau_2'}^B = -2s - h\right\},\\
&\theta_t = A \text{ for } t = \tau_2'.
\end{aligned} \tag{A6}$$

Note that the sequence of choices in $(A6)$ also satisfies optimality. By the end of the sequence of choices in $(A5)$ and $(A6)$, the value functions are $v_A(-s, -2s-h)$ and $v_A(-s-h, -2s-h)$, respectively. Since both $v_A(s+h,0) - v_A(-s,-2s-h)$ and $v_A(s,0) - v_A(-s-h,-2s-h)$ are equal to $(2s+h)/(1-\beta)$, which is a constant *and* the optimal switching rule does not depend on the *level* of profits from investments, the sequence of choices that starts with investment A and then follows choices in $(A5)$ infinitely is optimal. Denote the rule that generates this sequence of choices as π_1^*. The same argument applies to the sequence of choices that starts with investment A and follows choices in $(A6)$ ad infinitum and so denote π_2^* as the optimal switching rule for $v_A(s,0)$. Further define $\{\tau_i^*\}$ as the sequence of switching times generated from π_i^* for $i \in \Theta$. Since each τ_i^* is a stopping time, $\{\tau_i^*\}$ is simply a sequence of the independent stopping times such that at each stopping time the profit from an investment declines by h for the first time from the profit at the previous stopping time. From the way that π_1^* and π_2^* are constructed, it follows that $\tau_1 = \tau_2, \tau_1' = \tau_2'$ and so $\{\tau_1^*\} = \{\tau_2^*\}$. Therefore the timing of switching are matched under π_1^* and π_2^*. So the switching cost will disappear in the difference $v_A(s+h,0) - v_A(s,0)$.

Now we are ready to derive the difference of $v_A(s+h,0) - v_A(s,0)$. Consider $v_A(s+h,0) - v_A(s,0)$ until the first switch of investments under optimal switching rules π_1^* and π_2^*. Similar to the way in which the sequence of $\{\tau_i^*\}$ is decomposed, $v_A(s+h,0) - v_A(s,0)$ can be also decomposed by the $k+1$ stopping times such that at each stopping time the profit from an investment declines by h from the profit at the previous stopping time for the first time. Therefore each of the $k+1$ terms in $v_A(s+h,0) - v_A(s,0)$ is exactly equal to h times the L.H.S.

of $(A4)$, which is $(1-\lambda)/(1-\beta)$. This is because, for each sample path of π_1^* at t from zero to τ_1, there exists an identical sample path of π_2^* except it is always h below. Since $\pi_i^*, i \in \Theta$, is merely a repetition of the sequence of choices defined in $(A5)$ and $(A6)$, this difference occurs between every n-th and $(n+1)$-th switching time for all n equal to zero and even integers. Therefore, we get:

$$\begin{aligned} & v_A(s+h,0) - v_A(s,0) \\ & = \left[h\frac{1-\lambda}{1-\beta}\frac{1-\lambda^{k+1}}{1-\lambda}\right]\sum_{j=0}^{\infty}\lambda^{2(k+1)j} = h\frac{1}{(1-\beta)\left(1+\lambda^{k+1}\right)}. \end{aligned}$$

This completes the proof. □

Claim 2. With $k = 2s/h$,

$$v_B(s,h) - v_B(s,0) = h\frac{1+\lambda^{k+1}-\lambda}{(1-\beta)\left(1+\lambda^{k+1}\right)}.$$

Proof of Claim 2. Similar to Claim 1, the proof is done by constructing a sequence of optimal choices. For $v_B(s,h)$, suppose an investor starts with investment B as follows:

$$\theta_t = B \text{ for } 0 \le t < \tau_1 \text{ where } \left\{\tau_1 > 0 | s_t^B > 0 \,\forall 0 \le t < \tau_1, s_{\tau_1}^B = 0\right\}, \qquad \text{(A7)}$$

and, after time τ_1, the investor chooses as follows.

$$\begin{aligned} & \theta_t = A \text{ for } \tau_1 \le t < \tau_1' \\ & \text{where} \left\{\tau_1' > \tau_1 | s_t^A > -s-h \,\forall \tau_1 \le t < \tau_1', s_{\tau_1'}^A = -s-h\right\}, \\ & \theta_t = B \text{ for } \tau_1' \le t < \tau_1'' \\ & \text{where} \left\{\tau_1'' > \tau_1' | s_t^B > -2s-h \,\forall \tau_1' \le t < \tau_1'', s_{\tau_1''}^B = -2s-h\right\}. \end{aligned} \qquad \text{(A8)}$$

Clearly, the sequence of choices that starts with $(A7)$ and then follows $(A8)$ satisfies optimality. For $v_B(s,0)$, suppose an investor starts with investment B as follows:

$$\begin{aligned} & \theta_t = B \text{ for } 0 \le t < \tau_2 \\ & \qquad \text{where} \left\{\tau_2 > 0 | s_t^B > -h \,\forall 0 \le t < \tau_2, s_{\tau_2}^B = -h\right\}, \end{aligned} \qquad \text{(A8)}$$

and, after time τ_1, the investor chooses as follows.

$$\begin{aligned} & \theta_t = A \text{ for } \tau_2 \le t < \tau_2' \\ & \text{where} \left\{\tau_2' > \tau_2 | s_t^A > -s-h \,\forall \tau_2 \le t < \tau_2', s_{\tau_2'}^A = -s-h\right\}, \\ & \theta_t = B \text{ for } \tau_2' \le t < \tau_2'' \\ & \text{where} \left\{\tau_2'' > \tau_2' | s_t^B > -2(s+h) \,\forall \tau_2' \le t < \tau_2'', s_{\tau_2''}^B = -2(s+h)\right\}. \end{aligned} \qquad \text{(A10)}$$

The sequence of choices that starts with $(A8)$ and then follows $(A10)$ also satisfies optimality.

By the end of the sequence of choices – following $(A7)$ and then $(A8)$ – the value function is $v_A(-s-h, -2s-h)$. The value function at τ_2'' is $v_A(-s-h, -2s-2h)$. These two value functions are the same as the value functions at time $t = \tau_1$ and τ_2 as far as the optimal switching rule is concerned, since the difference in the value functions at time $t = \tau_1''$ and τ_2'' and the difference in the value functions at time $t = \tau_1$ and τ_2 are the same and are equal to $(2s+h)/(1-\beta)$ which is constant. This implies that the sequence of choices – following $(A7)$ and then repeating $(A8)$ over and over – is optimal. Denote π_1^* as the optimal switching rule that generates this sequence of choices. The sequence of choices – following $(A8)$ and then repeating $(A10)$ ad infinitum – is also optimal. Denote π_2^* as this optimal switching rule. Further define $\{\tau_i^*\}$ as the sequence of stopping times from π_i^* for $i \in \Theta$. As in Claim 1, $\{\tau_i^*\}$ is simply a sequence of the independent stopping times such that at each stopping time the profit from an investment declines by h for the first time from the profit at the previous stopping time. From the way that π_1^* and π_1^* are constructed, it follows that $\tau_1 = \tau_2, \tau_1' = \tau_2'$, and $\tau_1'' = \tau_2''$ and so $\{\tau_1^*\} = \{\tau_2^*\}$. Hence switching costs will disappear from the difference $v_B(s,h) - v_B(s,0)$.

Now we prove Claim 2. As in Claim 1, until the first switch of investments under π_1^* and π_2^*, each of the $k+1$ terms in $v_B(s,h) - v_B(s,0)$ is exactly equal to $h(1-\lambda)/(1-\beta)$. Since optimal switching rules π_1^* and π_2^* are mere repetitions of the sequence of choices defined in $(A7)$ to $(A10)$, the difference of $h(1-\lambda)/(1-\beta)$ only happens between every n-th and $(n+1)$-th switching time for all n equal to zero and even integers. In other words, whenever investment B is chosen, $v_B(s,h)$ leads $v_B(s,0)$ by h and there is no difference between $v_B(s,h)$ and $v_B(s,0)$ otherwise. Hence the following can be derived:

$$v_B(s,h) - v_B(s,0) = h\frac{1-\lambda}{1-\beta} + \lambda^{k+2}\left[h\frac{1-\lambda}{1-\beta}\frac{1-\lambda^{k+1}}{1-\lambda}\right]\sum_{j=0}^{\infty}\lambda^{2(k+1)j}.$$

Rearranging terms completes the proof. □

Proof of Lemma 1. The difference $v_A(s+h,0) - v_B(s,h)$ can be rewritten as

$$\begin{aligned} & v_A(s+h,0) - v_B(s,h) \\ &= v_A(s+h,0) - v_A(s,0) + v_B(s,0) - v_B(s,h) + C \\ &= -h\frac{-\lambda^{2k+2} + 1 - \lambda^k + \lambda^{2k+1}}{(1-\beta)(1-\lambda^{2k+2})} + h\frac{\lambda^{k+1}}{(1-\beta)(1+\lambda^{k+1})} + C. \end{aligned}$$

The last equality is from Claims 1 and 2. Rearrangement completes the proof. □

Lemma 2. *Let* $k = 2s/h$.

$$v_A(s-h,0) - v_B(s,-h) = \frac{h(\lambda^k - 1)}{(1-\beta)(1+\lambda^{k+1})} + C\,.$$

I omit the proof of Lemma 2 since it uses a similar argument to Lemma 1.

Proof of Theorem 1. The proof is a direct consequence of Lemmas 1 and 2. Solving for s when an investor is indifferent between switching now and delaying, we get:

$$s - C(1-\beta) + \beta ph \frac{\lambda(1-\lambda^k)}{(1-\beta)(1+\lambda^{k+1})} + \beta(1-p)\, h \frac{\lambda^k - 1}{(1-\beta)(1+\lambda^{k+1})} = 0.$$

Rearranging terms, I complete the proof. □

References

Agrawal, R., Hegde, M., Teneketzis, D.: Multi-armed bandit problems with multiple plays and switching costs. Stochastics and Stochastic Reports **29**, 437–439 (1990)

Asawa, M., Teneketzis, D.: Multi-armed bandits with switching penalties. IEEE Transactions on Automatic Control **41**, 328–348 (1996)

Balvers, R. J., Cosimano, T. F.: Periodic learning about a hidden state variable. Journal of Economic Dynamics and Control **17**, 805–827 (1993)

Banks, J. S., Sundaram, R. K.: Denumerable-armed bandits. Econometrica **60**, 1071–1096 (1992)

Banks, J. S., Sundaram, R. K.: Switching costs and the Gittins index. Econometrica **62**, 687–694 (1994)

Bergemann, D., Valimaaki, J.: Stationary multi choice bandit problems. Journal of Economic Dynamics and Control **25**, 1585–1594 (2001)

Brennan, M. J., Schwartz, E. S.: Evaluating national resource investments. Journal of Business **58**, 135–157 (1985)

Cox, D. R., Miller, H. D.: The theory of stochastic processes. New York: Chapman & Hall 1964

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

Dixit, A. K., Pindyck, R. S.: Investment under uncertainty. Princeton: Princeton University Press 1994

Easley, D., Kiefer, N. M.: Controlling a stochastic process with unknown parameters. Econometrica **56**, 1045–1064 (1988)

Gittins, J. C.: Multi-armed bandit allocation indices. New York: Wiley 1989

Gittins, J. C., Jones, D. M.: A dynamic allocation index for the sequential design of experimentation. In: Progress in statistics, pp. 241–266. Amsterdam: North-Holland 1974

He, H., Pindyck, R. S.: Investments in flexible production capacity. Journal of Economic Dynamics and Control **16**, 575–599 (1992)

Keller, G., Rady, S.: Optimal experimentation in a changing environment. Review of Economic Studies **66**, 475–507 (1999)

Lippman, S. A., Rumelt, R. P.: Industry-specific capital and uncertainty. Mimeo, University of California Los Angeles (1985)

McLennan, A.: Price dispersion and incomplete learning in the long run. Journal of Economic Dynamics and Control **7**, 549–564 (1984)

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

Pindyck, R. S.: Irreversibility, uncertainty, and investment. Journal of Economic Literature **29**, 1110–1148 (1991)

Rothschild, M.: A two-armed bandit theory of market pricing. Journal of Economic Theory **9**, 185–202 (1974)

Rustichini, A., Wolinsky, A.: Learning about variable demand in the long run. Journal of Economic Dynamics and Control **19**, 1283–1292 (1995)