Tortoises and Hares: Choice Among Technologies of Unknown Merit

Robin Cowan

The Economic Journal, Vol. 101, No. 407 (Jul., 1991), 801-814.

Stable URL:
http://links.jstor.org/sici?sici=0013-0133%28199107%29101%3A407%3C801%3ATAHCAT%3E2.0.CO%3B2-S

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

The Economic Journal is published by Royal Economic Society. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/res.html.

The Economic Journal
©1991 Royal Economic Society

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact jstor-info@umich.edu.

©2003 JSTOR

http://www.jstor.org/
Fri Aug 29 05:04:30 2003

The Economic Journal, **101** (*July* 1991), 801–814
Printed in Great Britain

TORTOISES AND HARES: CHOICE AMONG TECHNOLOGIES OF UNKNOWN MERIT*

Robin Cowan

Of the recent work in the area of competing technologies and increasing returns, virtually all has been concerned with the effectiveness of the market in delivering optimal outcomes.[1] Arthur (1989), Farrell and Saloner (1985), and Katz and Shapiro (1985) develop models in which every agent acts rationally, yet an inferior technology can dominate the market. Increasing returns to adoption – learning by using, technical co-ordination, or network externalities – lie at the source of these potential sub-optimal outcomes. It may appear that in the face of increasing returns to adoption intervention by a central authority would prevent these inefficiencies from appearing. Certainly many interventions of this sort exist. Managers of research and development laboratories play this role. In every country that has had a nuclear energy programme, the branch of government overseeing the programme distributed research funds among competing technologies. Currently, as budgets shrink, the co-ordinators of Star Wars research are faced with choices among many strategies. In these and other cases (AIDS, and acid rain reduction research for example) intervention is seen as appropriate, not only due to the size and risk of the project but also as a way of co-ordinating research so that the best technology emerges. Could it be that in such instances of centralised decision making, the eventual choice of an inefficient technology can be avoided?

One might think that a central authority could simply subsidise all of the available technologies until it knows which is superior and then force its adoption.[2] This is only partly so. This paper will argue that even with complete central control, when the technologies are improving and the central authority cannot precisely predict the paths of development, it is possible to lock in to an inferior technology. Discounting, coupled with concern for social profitability, can cause the central authority to pursue a technology that it believes to be best, at the cost of further learning about other technologies. This allows erroneous estimates of the relative 'goodness' of technologies to be perpetuated.[3]

I. TECHNOLOGY CHOICE WITHOUT UNCERTAINTY

Two technologies, A and B, are available for performing the same task and both are subject to increasing returns to adoption. The source of increasing

* I thank Brian Arthur, Tim Bresnahan, Boyan Jovanovic and two anonymous referees for helpful comments. Financial assistance from the Social Sciences and Humanities Research Council of Canada, and the C. V. Starr Center for Applied Economics is gratefully acknowledged.

[1] See Arthur (1988*a*) or Foray (1989) for surveys of this literature.

[2] 'Central authority' should not be read necessarily to mean 'government'. What is important here is co-ordination of actions, not the agent performing the co-ordination.

[3] For an insightful discussion of problems faced by government in this area, see David (1987).

returns is learning which, following Arrow (1962), cannot be applied retroactively, so for each technology the net benefit to the next adopter increases with the number of previous adopters of that technology. Technologies are *unsponsored* and are supplied competitively.[4] Each period one consumer adopts one of the technologies. His decision is irrevocable. He receives a payoff defined as the present value of the net benefits of his use of that technology. Finally, a central authority can steer the adoption process through selective subsidies. Its objective is to maximise the present value of the entire stream of adoptions.

In this model the dynamics of the market are trivial. Let P^i_j be the payoff to the ith adopter of technology j. If $P^1_A > P^1_B$ then the first adopter in the market will adopt technology A. The second adopter will either be the second adopter of technology A or the first of technology B. Increasing returns to adoption imply that $P^{i+1}_j > P^i_j$, so the second adopter in the market will make the same choice as the first.[5] All adopters chose the same technology, that with the higher initial payoffs, regardless of the relative magnitude of the present value of this adoption stream.

Here, a central authority can be effective. At the beginning of the process it could simply do a present value calculation for two potential paths – one in which only A is adopted and one in which only B is adopted. It can then pursue the superior path forcing the adoption only of A or only of B. Further, it follows from Arthur's Doctor/Lawyer Theorem that this is an optimal policy. If two technologies are competing and adopters act sequentially, the policy that maximises the present value of the stream of adoption will specify that only one of the technologies is ever used.[6]

The simplicity of the results stems from the fact that the development paths of the two technologies are known with certainty. If the paths are not well-known we would expect to see different behaviour on the part of the central authority. He would be likely to make projections of the two paths and base his early choices on these projections. With the early adoptions, though, comes information with which the projections can be updated. This updating is likely to cause the authority to switch back and forth between technologies before settling on one of them. This is the essence of the model in the next section.

II. TECHNOLOGY CHOICE WITH UNCERTAINTY

Assume two competing technologies, A and B, both subject to increasing returns to adoption. The technologies are unsponsored and are competitively supplied. Each period one of the technologies is adopted, but a central authority is using selective subsidies to guide the adoption process. I assume that the technologies are large relative to the budget of the authority so lumpiness ensures that only a single technology is adopted each period. Each adoption generates a payoff, again defined as the discounted net benefits accruing to the particular adoption. The source of increasing returns, learning-by-using, does not apply retroactively.

[4] See Arthur (1983) for definition and discussion of *sponsored* and *unsponsored* technologies.

[5] This abstracts from differences in tastes. This issue has been addressed by Arthur (1989).

[6] Arthur states this theorem in Arthur (1988*b*), but does not provide a proof. Though intuitively obvious, this theorem is difficult to prove from first principles. A proof is provided in the Appendix.

The objective of the central authority is to maximise the expected present value of the net benefits of the strêam of adoptions. The policy instrument is a subsidy, so its problem is to decide, each period, which technology to subsidise.

This model differs from the previous model in that the learning curves of the two technologies are unknown. Assume that the learning curves are known to belong to a family of functions parameterised by a single variable. For ease of exposition, assume both technologies belong to the same family. Thus the net benefit to the n_Ath adopter of technology A is $f(a, n_A)$, similarly the net benefit of the n_Bth adopter of technology B is $f(b, n_B)$. The parameters a and b are random variables, the randomness arising from measurement error and from variation in operating environments of the technologies. Observed payoffs thus will be distributed around the learning curves.

At the beginning of the adoption process the central authority forms beliefs about the densities of a and b. The true densities are estimated by priors $g_A(a)$ and $g_B(b)$, which have corresponding distributions $G_A(a)$ and $G_B(b)$. Every adoption is like a random experiment. Each period, based on its current beliefs $g_A(a)$ and $g_B(b)$, the central authority decides which technology (A say) should be adopted that period. That technology is adopted and the payoff to that adoption is observed. This observation is then used to update $g_A(.)$, and the new beliefs are used in deciding about the next adoption. For notational convenience, the new, posterior beliefs are also denoted $g_A(a)$. When confusion is possible time superscripts are added. The updating process is such that $g_A(a)$ and $g_B(b)$ are consistent estimators.

Finally, assume that the central authority guides the process indefinitely, having an infinite horizon and a positive discount rate.

Define $x_A(t)$ to be the state of technology A at time t. $x_A(t)$ is an ordered m-tuple of information: $\{n_A(t); \alpha^n, \alpha^{n-1}, \dots, \alpha^1\}$ where $n_A(t)$ is the number of adoptions of technology A prior to time t, and α^j is the observation on a at the jth adoption of technology A. $x_B(t)$ is defined similarly. The central authority sees the process as moving from state to state according to the following transition rule:

If $x_A(t)$ is defined as above then

$\Pr\{x_A(t+1) = [n_A(t)+1; h, \alpha^n, \alpha^{n-1}, \dots, \alpha^1]$ given that A is used at $t\} = g_A^t(h)$

$\Pr\{x_A(t+1) = [n_A(t); \alpha^n, \alpha^{n-1}, \dots, \alpha^1]$ given that A is not used at $t\} = 1$.

The transition for $x_B(t)$ is described similarly.

We can now state the problem of the central authority as a dynamic programming problem. Defining $V[x_A(t), x_B(t)]$ as the value function, and γ as the discount factor, we can write the recurrence relation as

$$V[x_A(t), x_B(t)] = \mathrm{Max}\left\{\begin{array}{l}\int f[a, n_A(t)+1]\, dG_A^t + \gamma \int V[x_A(t+1), x_B(t)]\, dG_A^t \\ \int f[b, n_B(t)+1]\, dG_B^t + \gamma \int V[x_A(t), x_B(t+1)]\, dG_B^t\end{array}\right\}.$$

Solving this equation is the task of the decision maker.

28-2

III. THE TWO-ARMED BANDIT

The dynamic programming problem can be finessed by considering the two-armed bandit from probability theory. The two-armed bandit is characterised as a slot machine, or fruit machine, that has two arms. The jackpot probabilities of each of the arms are unknown but are independent. Each play generates either success – a jackpot of fixed, known size, or failure – zero payoff. The objective is to play the arms, one at a time in any order, so as to maximise the expected discounted winnings. The model of technology choice differs from the two-armed bandit in two respects. First, there is a payoff to every adoption, drawn from a distribution of payoffs. Second, with each adoption there is a movement along the learning curve, so there is a tendency for payoffs to increase in size as a technology is used. That is, the decision maker knows that the true distribution of payoffs shifts to the right when a technology is used, though he does not know how big this shift is.

The work of Gittins and Jones (1974) has produced an optimal policy for a large class of problems that share the objective of maximising the expected present value of the selection process. This is known as the Gittins index policy. The problems addressed by Gittins and Jones are adaptive control problems consisting of several simultaneously operating Markov processes. Benefits are time separable and each period are determined by some function of the results of past and present states. Of the two controls, *continue* and *wait*, *continue* can be applied to only one process each period; *wait* must be applied to all others.[7] Each arm of the two-armed bandit (or more generally the multi-armed bandit) represents one of the Markov processes.

The Gittins index, sometimes known as the Dynamic Allocation Index, is a scalar assigned to each individual arm, which is independent of the other arms and which ranks the value of a single try on that arm. It can be conceived of using the following thought experiment: Call the arm for which the index value is desired 'the unknown arm'. Postulate a hypothetical arm 'the known arm' for which the probability of getting a jackpot is known to be p. Now offer the player a choice between playing the known arm forever, or playing the unknown arm at least once, possibly more times, with the option of switching at some future time to the known arm, which must then be played forever. The value of p for which the player is indifferent between these two options is the value of the Gittins index for the unknown arm. This thought experiment forms the basis of the dynamic programming procedure used to calculate the index.

The Gittins index policy consists in playing, each period, that arm which has the highest index value.

PROPOSITION 1. *The expected present value of the adoption process controlled by a central authority is greater than the expected value of an uncontrolled process.*

This follows directly from the optimality of the Gittins index policy.[8] The

[7] See Gittins and Jones (1974) for more details of this class of problems.

[8] For a proof of this result see Gittins and Jones (1974).

technology choice model described above is a member of the class of problems for which the Gittins policy is optimal. While a central authority maximising the expected value of the process will follow this policy, a market will not. Intuitively, the divergence between the policies of a central authority and a market arises from the potential trade-off between immediate payoff and knowledge. A central authority, in deciding which technology to subsidise, takes into account what might be learned about the merits of each technology. In an unfettered market an adopter will be unwilling to sacrifice his immediate payoffs so that those who follow will be able to make better informed decisions. To maximise the expected benefit from his action an adopter will ignore (or set to zero) the second terms in the recurrence relation. The second term includes potential benefits from learning about payoffs, and to ignore it will generate a sub-optimal policy.[9]

The Gittins result implies that a central authority can maximise the expected present value of the adoption process and so will be more successful than an unfettered market.[10] This is not to say that a central authority would be able to prevent the dominance of an inferior technology, however. Whether or not this occurs has to do with the long run dynamics of the system, with which Gittins and Jones were not concerned. The following propositions indicate that the two common results of the literature on markets – that the market will lock in to one technology, and that it might not be the right one – are found in this model.

PROPOSITION II. *The optimal policy will, with probability one, lead to only one technology being used an infinite number of times*

This is a well-known result in the bandit literature – that within finite time all but one of the arms cease to be used.[11] The intuition is straightforward. Suppose both technologies are used indefinitely often. The estimators of the true distributions of a and b are consistent, so within finite time the estimates get very close to the true distributions. Further, there is a time after which the estimates stay close to the true distributions. The variances of the estimates also approach zero, and small variances imply that there is little to be gained from experimenting. At this point the present value of exclusive use of one of the technologies will be greater than the other. Given that there is little to be gained from experimenting, the optimal policy will be to use the technology with the higher payoff. Consistency of the estimator implies that the authority

[9] If the central authority is seen as government, intervention may create opportunities for rent-seeking or other inefficiencies. Policy makers have many tools – co-ordination and support of research; standards setting; procurement – with which to implement policies in this area. The particular circumstances of the technologies will determine which tools are appropriate and the degree to which policy generates its own inefficiencies. See Cowan (1990*b*) for a discussion of government standards policy for information technology.

[10] The Index policy is not robust to informational spillovers. That is, if using technology A yields information about how well technology B will perform, the Gittins index policy is no longer optimal. It can be modified, however, to accommodate learning spillovers. If using technology A allows improvements to be made to technology B, by altering the way payoffs enter the calculation of the index the Gittins policy will retain its optimality.

[11] Rothschild (1974) proves this proposition for the simple two-armed bandit. Cowan (1988) provides a proof which takes advantage of the increasing returns.

is never proved wrong – it never gets news bad enough to force it to change its beliefs.[12]

This proposition states that eventually, but in finite time, one of the two technologies will be dropped, never to be used again. If the worse one is dropped, this might be seen as an argument for the efficacy of intervention. In the argument for the proposition, though, there was no assumption that the technology favoured in the end is better. This is not merely an artifact of the proof. As the next proposition states, it is possible that the superior technology is dropped, and that an inferior one dominates.

Before stating the proposition, a caveat is needed to prevent the problem from being trivial. Clearly, there are situations where one cannot lock in to the wrong technology, since it looks so bad at the start that it will never be tried. (Similarly if for some reason the good one looks bad at the start it will never be tried.) To eliminate this possibility assume that the supports of the initial priors are not disjoint, so both technologies will be used with positive probability. In addition, there are increasing returns functions where the increasing returns are very strong very early, such that the first technology tried will be the only one ever tried. Assuming, therefore, that the priors and increasing returns functions are such that there is positive probability of trying both technologies in finite time, we can state:

PROPOSITION III. *The optimal policy will, with positive probability, lead to the inferior technology being adopted infinitely often and the superior technology being adopted only a finite number of times.*

This is, in effect, the two-armed bandit result that the player can lock in to the arm less likely to yield jackpots.[13] A word must be said about 'inferior'. In the bandit model, 'inferior' simply means 'lower jackpot probability'. In this model, because the learning curves of technologies might cross, this definition will not do. Here, an inferior technology is one which, as the dominant technology, yields less than maximal net benefits.[14]

The argument for this proposition is very simple. By assumption, it is possible to use every technology at least once. Suppose the bad technology is used, but that it produces good results.[15] Not getting a bad reading, the central authority would be inclined to use it again. This can continue for many adoptions. But each time a technology is used it advances along its learning curve, so the expected value of its true benefits increases. If this technology continues to be

[12] There are two interpretations of this argument. The first is simply that the Gittens index approaches the mean of the posterior distribution of discounted payoffs. The second is that eventually we obtain close to perfect knowledge, at which point the arguments for the Doctor/Lawyer theorem will hold.

[13] Again, proofs exist in Rothschild (1974) and Cowan (1988). The presence of increasing returns facilitates both the process and the proof.

[14] If the system locks into A at t^* (B is last used at t^*-1) then A is the inferior technology if

$$\mathbf{E}\sum_{i=0}^{\infty}\gamma^i f[a, n_A(t^*)+i] < \mathbf{E}\sum_{i=0}^{\infty}\gamma^i f[b, n_B(t^*)+i],$$

where $\mathbf{E}$ is the expectation operator taken over the true distributions of a and b.

[15] Technically, a good result is a payoff large enough that the Gittins index does not decrease.

used, it will eventually have advanced sufficiently along its learning curve that the other one, no matter how good, will have been left behind.

IV. DISCUSSION

Propositions two and three echo the results of the literature on market processes. Even with complete centralised control of decision making, the market share of one of the technologies goes to one, the other goes to zero. Further, if the goal of the central authority is to maximise the expected present value of the process, it cannot prevent an inferior technology from capturing the entire market.

Two Sources of Lock-In

In the model described above we can identify two distinct forces driving the system to lock in to one technology. The first is the reduction of uncertainty. When the adoption process begins, the merits of neither technology are well-known. An important aspect of early use of the technologies is simply the reduction of this uncertainty. This reduction will, by itself, cause lock-in to occur, and is the sole cause of it in the simple two-armed bandit. As technologies are used more and more, beliefs about the parameters a and b converge to the true distributions. If we were to use both technologies many times we would eventually 'know' how good they were and would then pick the better. This reasoning generates the lock-in result. The 'possibly inferior technology' result needs first that the estimates of the value of the superior technology be erroneously low. This can arise from the initial prior, or from bad luck with early implementations of it. The inferior technology can then be used. If the inferior technology does not have bad luck it will continue to be used, and beliefs about it will converge to the truth. We eventually 'know' how good this technology is and, further, know that it is better than what we believe the other to be. (Notice that in the absence of increasing returns our beliefs about the superior technology must be that it is worse than the *true* value of the inferior one.) Because the superior technology is not being used, our beliefs about it do not change – it has no way to demonstrate its superiority, and we are locked in to the other.

That the reduction of uncertainty is enough to generate lock-in implies that the results will hold for most regimes of increasing or decreasing returns. The increasing returns regime was described above. Constant returns resemble the two-armed bandit and has just been described. Bounded increasing or decreasing returns will eventually approach a constant returns regime and are treated that way. The exception is a regime of unbounded decreasing returns in both technologies. It seems reasonable to assume, though, that the amount that can be lost from the adoption of a technology has some upper bound.

The second mechanism driving lock-in is increasing returns. This is the driving force behind Arthur's (1989) results and was isolated in the policy context in the first simple model above. More generally, in the early stages of the competition, as technologies are tried out, inevitably one of them will be used more frequently than the other. As this occurs, the more-used technology advances along its learning curve more quickly, and the other technology gets left behind. It becomes more and more costly to shift away from the more-used technology if, for example, inherent problems with it are discovered. In switching, one would be forced to suffer through the early, low-benefit segment of the learning curve of the less-used technology before payoffs rise to match those of the more-used technology.

An important effect of the snowballing nature of adoption under increasing returns implies that the central authority can believe it has locked in to an inferior technology. This is not so under constant or decreasing returns.

Under constant returns, early in the adoption process the optimal policy can dictate the use of the technology believed to have the lower mean net benefit. This occurs if its prior distribution has relatively large higher order moments. While using this technology causes losses in terms of immediate payoffs it provides gains by reducing uncertainty. If there are constant returns ($\partial f/\partial n = 0$) then as the process continues and the higher order moments fall, a technology with lower expected payoffs will not be used. If there is little to be gained in terms of reduced uncertainty and a technology is thought to yield lower payoffs, why use it? Thus late in the process, the central authority always believes that the superior technology is being used. This is not necessarily the case if payoffs systematically increase. As the higher moments of the priors on the much-used technology (A, say) fall, it may become apparent that it is not a good technology. The central authority may believe, in fact, that $f(a, n)$ lies everywhere below $f(b, n)$. In this case, though, to switch away from the well-used technology would involve suffering through the low learning curve portion of the other. Discounting can prevent this. Thus with constant or decreasing returns a central authority will always believe, sometimes wrongly, that it has steered the process to the best outcome. With increasing returns by contrast, the authority (and its constituents) can believe that the system has locked in to an inferior technology.

The Efficacy of Intervention

The work of Gittins and Jones (1974) shows that with regard to the goal of maximising the expected present value of the adoption process, a central authority will do better than a market. The third proposition states that pursuit of this goal, which is often presumed to be an appropriate action for government, will not necessarily lead to the right technology dominating the market. This analogue of Arthur's (1989) inefficiency result may cast doubt on the efficacy of intervention.

The degree to which this result presents a problem of significant magnitude depends on several factors, all of which interact.[16] The first is the degree to which one technology is superior to the others. The greater the difference in quality the less likely the dominance of a bad technology. This is clear, and

[16] Simulating this system with $f(a, n) = ag(n)$, suggests that when the true values of the technologies differ by 10%, $b = 1{\cdot}1a$, with a small degree of increasing returns, the inferior technology dominates about 30% of the time.

works through the sampling mechanism – the better the technology, the higher the chances of getting a 'good' reading, and so the higher the chances of raising the estimates of its value. If two technologies have very different true values, then this is likely to be clear from a small number of samples. There are confounding factors though.

There is a bias in favour of the better known technology. The bias is stronger in the market than it is under the Gittins policy but it exists there as well. If the variance of the original prior is small, each sample adds only a small amount of information, and so the estimate of the true distribution changes very slowly. If the technology with the small-variance prior ever gets used, then there is a strong tendency to continue using it, as the estimate of its value (which indicated that it was better than the others) changes very little. The fact that the Gittins index gives positive weight to the variance of the prior mitigates this effect, but does not eliminate it. This bias in favour of a well-known technologies has implications for competitions between old and new technologies. Old technologies tend to be well-understood – the estimates of their values have low variances. This is not typically true of new technologies, which are often accompanied by considerable uncertainty. Thus these competitions are biased towards older technologies and against new ones. To win these competitions, new technologies must both appear better at the start, and produce early results good enough that the older ones are quickly left behind. The apparent conservatism favouring old technologies has a source not only in the risk aversion of the agents but also in the optimal policy itself.

The degree of increasing returns can also have an effect on the probability of locking in to an inferior technology. Consider a constant returns regime. Eventually the true value of one of the technologies comes to be learned. It will continue to be used only if the estimates of the other technologies suggest that they are worse. If the former is an inferior technology, the others will only look worse if, early in the process, their estimates have been driven below their true values. If the estimates of the best technology are accurate, however, eventually it will be used. The same is not true in an increasing returns regime. All that need be true is that the inferior technology was used until it had a large lead in learning. The stronger are the increasing returns (i.e. the faster learning is accumulated) the easier it is to gain this lead. Even if the estimates of the better technology are accurate, an erroneously high estimate of the inferior technology will allow it to be used, and increasing returns will allow it to continue being used. Whether or not an inferior technology is ever used will be determined by the results of early usage of the various technologies and, crucially, on the initial estimates of their relative merits, but strong increasing returns facilitates the continued use of an inferior technology.

The discussion suggests a further result: if increasing returns are very strong, lock-in will occur very rapidly.[17] If learning takes place rapidly enough, the first technology used will be the only one ever used. Rosenberg (1982) points out that complex technologies representing large advances over the current

[17] See Cowan (1988) for a proof.

state of technology very often exhibit enormous learning advances as they are first used. This implies that the choice of technology for use in the first prototype is of crucial importance. It can determine which of the technologies will dominate the market. Cowan (1990*a*) suggests that this may have happened with nuclear power technology in the United States. The first prototype was a light water reactor at Shippingport Pennsylvania. The experience gained there, combined with its work for the US Navy, gave Westinghouse (and later General Electric) the headstart it needed to dominate the reactor market with the light water technology.

The final factor of importance in determining the probability of locking in to an inferior technology is the initial estimates of their relative merits. The importance of the variances and higher order moments was discussed above; the importance of the means is self-evident.

Other Objectives

The problem of the central authority was formulated with the objective of maximising the expected present value of the adoption process. Another conceivable objective would be to minimise the probability of getting the wrong technology in the end. Without constraints on the possible actions of the central authority, for example a total budget constraint or perhaps a time constraint, the problem may not have a solution, since at any finite time, there are always actions one can take that will generate more accurate estimates of the relative merits of the technologies, and so lower the probability of getting the wrong one. With the current formulation of the problem, though, we can see that relative to the market a central authority will lower the probability of locking in to an inferior technology. The optimal policy generates a higher expected value for the entire process. This implies that the technology with higher payoffs must be used relatively more often. But if it is used more frequently than the other technologies, it will advance more quickly along its learning curve. This will tend to reinforce the desirability of using it and so further its use. Thus a policy cannot both increase the present value of the process and reduce the probability of locking in to the superior technology.

Research and Development

The model as developed contained the assumption that every adoption yields an immediate payoff. It is the case, however, that much expenditure on technology, and often crucial decision-making, takes place before the first prototype is ever built. This was certainly true of the development of nuclear power. It is true now with AIDS research, Star Wars research, and research on seed varieties. These situations can be captured by the model with specific modifications.

If only a single technology survives past the research and development stage, we define the 'payoff stream' as follows:[18]

If at time t the product has already been completed, the net benefit is zero.

[18] This description is due to Gittins (1979) and matches the case where a constant research expenditure is being made every period. It can be modified to fit the case where expenditures vary.

If at t the product has not been brought to market and both technologies are still in the research and development stage, the net benefit is -1.

The goal is to maximise the discounted sum of these net benefits. Again, in this system we will lock in to one of the technologies, and possibly the wrong one. Here, the wrong technology is the one that would take longer to bring to market.

If more than one technology might survive past the research and development stage, the payoff function has two parts. Until the technology is brought to market, the (negative) payoff is the cost of one period of research and development. After the product is brought to market the payoff is the immediate payoff to the adopter as described above. The dynamics here are identical to those of the model above with the exception that during the research and development phase, beliefs are not updated by observations on payoffs to adoption, rather they are updated directly by reference to the research and development work. The propositions apply as before.

V. CONCLUSION

When competing technologies are being adopted sequentially, if there is uncertainty about the relative merits of the competitors, the market will under-supply experimentation. There is no incentive for an adopter to experiment with what appears to be an inferior technology in the hope of improving the estimate if its merit. A central authority can internalise this externality, and so raise the expected discounted value of a stream of adoptions. This suggests that there is a place for intervention. The goal of any intervention must be moderate, however, as even with a strong policy maker results found in models of the market obtain – one technology, and possibly not the best, will dominate.

Further, if the technologies are improving rapidly lock-in occurs quickly, and the first technology used gains a large advantage. In the development of any new, complex technology there is a relatively long period during which a substantial amount of learning takes place. This implies that the forces driving lock-in and facilitating the dominance of inferior technologies are very strong. Unfortunately from the point of view of government, large, complex technologies tend to be the ones about which it contributes to the decision making.

Finally, if it is important that the best technology be the one to which we lock in (as might be presumed to be the case with technologies like Star Wars for example) fiscal conservatism will have to be forsaken. In order to determine absolutely that one of the technologies is best, it would take an infinite expenditure, and an infinite amount of time. To be reasonably certain that we have the best technology will involve a very large expenditure. Any attempt to rationalise this expenditure will introduce a bias into the technology selection process – we are more likely to get the hare than the tortoise.

New York University and University of Western Ontario

Date of receipt of final typescript: October 1990

Appendix

Consider a control problem consisting of two processes A and B. The objective is to maximise the discounted present value of the payoff stream. There are two controls {*continue*, *wait*}, but *continue* is applied to exactly one process each period. The one-period payoffs to the processes are defined as $R_A(n_A)$ and $R_B(n_B)$ where n_i is the number of times *continue* has been applied to process i.

THEOREM. (Doctor/Lawyer) *If $R_A(n_A)$ and $R_B(n_B)$ are non-decreasing then the optimal policy is to apply* continue *to only one of the processes and* wait *to the other.*

The proof proceeds by mathematical induction on the number of times *continue* is switched from one process to the other.[19] First we prove a lemma:

LEMMA. *When $R_A(n_A)$ and $R_B(n_B)$ are non-decreasing, then from any time t^* the present value of the remainder of the process is greater if the control is never switched than if it is switched once.*

Proof. Normalise so that $n_A(t^*) = n_B(t^*) = 0$. Define γ as the discount factor. Without loss of generality, let $\sum_{i=0}^{\infty} \gamma^i R_B(i) > \sum_{i=0}^{\infty} \gamma^i R_A(i)$. There are two cases: one in which A waits first, and one in which B waits first.

Case I: *Continue* is applied first to B after t^*.

Suppose that the optimal time for a single switch is t^*+j.

Then if switching once is preferred,

$$\sum_{i=0}^{j-1} \gamma^i R_B(i) + \sum_{i=0}^{\infty} \gamma^{j+i} R_A(i) > \sum_{i=0}^{\infty} \gamma^i R_B(i).$$

Equivalently

$$\sum_{i=0}^{j-1} \gamma^i R_B(i) + \sum_{i=0}^{\infty} \gamma^{j+i} R_A(i) > \sum_{i=0}^{j-1} \gamma^i R_B(i) + \sum_{i=0}^{\infty} \gamma^{j+i} R_B(i+j)$$

or

$$\sum_{i=0}^{\infty} \gamma^i R_A(i) > \sum_{i=0}^{\infty} \gamma^i R_B(i+j) \geqslant \sum_{i=0}^{\infty} \gamma^i R_B(i)$$

since R_B is non-decreasing. This is a contradiction, so if B is used first it is preferred never to switch than to switch once.

Case II: *Continue* is applied first to A after t^*.

The optimal time for switching given a single switch is j. If starting with A and switching once is preferred to never switching then

$$\sum_{i=0}^{j-1} \gamma^i R_A(i) + \sum_{i=0}^{\infty} \gamma^{j+i} R_B(i) > \sum_{i=0}^{\infty} \gamma^i R_B(i) > \sum_{i=0}^{\infty} \gamma^i R_A(i). \tag{1}$$

Further, since switching at j is preferred to switching at $2j$, then

$$\sum_{i=0}^{j-1} \gamma^i R_A(i) + \sum_{i=0}^{\infty} \gamma^{j+i} R_B(i) > \sum_{i=0}^{2j-1} \gamma^i R_A(i) + \sum_{i=0}^{\infty} \gamma^{2j+i} R_B(i).$$

Equivalently,

$$\sum_{i=0}^{j-1} \gamma^i R_A(i) + \gamma^j \sum_{i=0}^{\infty} \gamma^i R_B(i) > \sum_{i=0}^{j-1} \gamma^i R_A(i) + \gamma^j \left[\sum_{i=0}^{j-1} \gamma^i R_A(j+i) + \sum_{i=0}^{\infty} \gamma^{j+i} R_B(i)\right].$$

[19] This proof is a revised version of the proof that appears in Cowan (1988).

But with (1), this implies that the value from switching at $2j$ (if $j \neq 0$) is higher than that from switching at j.

So for no finite j (> 0) is j an optimal time to switch. ($2j$ is always preferred.) Thus never switching yields higher present value than does switching once. ∎

Proof of the theorem (by mathematical induction).

Basis case. It is better never to switch than it is to switch once.
This follows directly from the Lemma.

Inductive hypothesis. For all $m < n$, it is better never to switch than it is to switch m times.

Claim. This implies that it is better never to switch than to switch n times.

Proof. Let the $n-1$th switch take place at time k. The claim then follows directly from the Lemma with $t^* = k$. (If switch $n-1$ is a switch to process A, then allow the nth switch to take place during the same period, so A is not used.)

Thus there is no finite $n > 0$ which is the optimal number of switches. ∎

To show that 0 switches is preferred to an infinite number, consider the optimal pattern of switching if there are an infinite number of switches. Call the value of this policy P_∞. Consider a policy that follows this same switching pattern (i.e. switches at the same times) to time t and never switches thereafter. Call its value P_t.

It follows from the Lemma that for all t there is a finite k such that $\forall s > t+k$, $P_t > P_s$. Discounting ensures that $P_t \rightarrow P_\infty$. But for all t $P_0 > P_t$ and so $P_0 > P_\infty$. ∎ ∎

References

Arrow, Kenneth J. (1962). 'The economic implications of learning by doing.' *Review of Economic Studies*, vol. 29, pp. 155–73.

Arthur, W. Brian (1983). 'On competing technologies and historical small events: the dynamics of choice under increasing returns.' International Institute for Applied Systems Analysis, Working Paper WP-83-90.

—— (1988*a*). 'Competing technologies: an overview.' In *Technical Change and Economic Theory* (ed. G. Dosi, C. Freeman, R. Nelson, G. Silverberg and L. Soete). London: Pinter.

—— (1988*b*). 'Self-reinforcing mechanisms in economics.' In *The Economy as an Evolving Complex System* (ed. P. W. Anderson, K. J. Arrow and D. Pines). Santa Fe Institute Studies in the Sciences of Complexity, Redwood City: Addison-Wesley.

—— (1989). 'Competing technologies, increasing returns, and lock-in by historical events.' Economic Journal, vol. 99 (March), pp. 116–31.

Cowan, R. (1988). 'Backing the wrong horse: sequential technology choice under increasing returns.' Mimeo, New York University.

—— (1990*a*). 'Nuclear power reactors: a study in technological lock in.' *Journal of Economic History*, vol. 50 (September), pp. 541–67.

—— (1990*b*). 'Standards policy and information technology.' OECD Committee for Information, Computer and Communications Policy.

David, Paul A. (1975). *Technical Choice, Innovation, and Economic Growth*. Cambridge: Cambridge University Press.

—— (1985). CLIO and the economics of QWERTY.' *American Economic Review*, vol. 75, pp. 332–7.

—— (1987). 'Some new standards for the economics of standardization in the information age.' In *Economic Policy and Technological Performance* (ed. P. Dasgupta and P. Stoneman). Cambridge: Cambridge University Press.

Farrell, J. R. and Saloner, G. (1985). 'Standardization, compatibility and innovation.' *Rand Journal of Economics*, vol. 16, pp. 70–83.

Foray, Dominique (1989). 'Les modèles de compétition technologique: une revue de la littérature.' In *Revue d'Économie Industrielle*, vol. 48, pp. 16–34.

Gittins, J. C. (1979). 'Bandit processes and dynamic allocation indices.' *Journal of the Royal Statistical Society* B, vol. 41, pp. 148–77.

Gittins, J. C. and Jones, D. M. (1974). 'A dynamic allocation index for the sequential design of experiments.' In *Progress in Statistics* (ed. J. Gani, K. Sarkadi and I. Vincze). Amsterdam: North-Holland.

—— (1979). 'A dynamic allocation index for the discounted multiarmed bandit problem.' *Biometrika*, vol. 66, pp. 561–5.

Katz, Michael L. and Shapiro, Carl (1985). 'Network externalities, competition and compatibility.' *American Economic Review*, vol. 75 (June), pp. 442–4.

Rosenberg, Nathan (1982). *Inside the Black Box: Technology and Economics*. Cambridge: Cambridge University Press.

Rothschild, Michael (1974). 'A two-armed bandit theory of market pricing.' *Journal of Economic Theory*, vol. 9, pp. 185–202.