Working Paper Series

DEPARTMENT OF AGRICULTURAL AND RESOURCE ECONOMICS AND POLICY

BERKELEY

California Agricultural Experiment Station

University of California

DEPARTMENT OF AGRICULTURAL AND RESOURCE ECONOMICS AND POLICY
DIVISION OF AGRICULTURE AND NATURAL RESOURCES
UNIVERSITY OF CALIFORNIA AT BERKELEY

WORKING PAPER NO. 935

UNCERTAINTY, LEARNING, AND THE IRREVERSIBILITY EFFECT

by

Urvashi Narain, Michael Hanemann, and Anthony Fisher

California Agricultural Experiment Station
Giannini Foundation of Agricultural Economics
September 2002

Uncertainty, Learning, and the Irreversibility Effect

Urvashi Narain
Resources for the Future, Washington, DC

Michael Hanemann
University of California, Berkeley

Anthony Fisher
University of California, Berkeley

Abstract

The well-known investment irreversibility effect states that investment is optimally delayed if future benefits are uncertain, the investment decision is irreversible, and there is no possibility of learning about future benefits. An unresolved question is whether this effect holds if the benefit function is nonlinear and investment is a continuous choice variable. Contrary to some earlier results which suggest that the effect does not hold widely under these conditions, we show that it does. We show, first, that necessary and sufficient conditions in the literature are only sufficient and not necessary; second, that the irreversibility effect holds for a case in which it is apparently violated; and third, that two cases in which the effect is violated are somewhat special and probably not empirically important.

This research received financial support from the National Science Foundation grant No. SES-9818642.

1. Introduction

Arrow and Fisher (1974) and Henry (1974) have established that faced with a linear benefit function or with an all-or-nothing investment choice, a decisionmaker will optimally delay investment if the future benefits are uncertain, investment decisions are irreversible, and there is a possibility of learning about future benefits, even when there is a cost to waiting. The question is whether the decision to delay investment to maintain flexibility in the face of uncertainty, irreversibility, and learning, which we will refer to as the irreversibility effect, still holds if investment is a continuous choice variable and the benefit function is nonlinear.

The existing literature suggests that the effect *does not* hold widely under these conditions. The seminal paper is by Epstein (1980), who provides a sufficient condition for the irreversibility effect to hold and uses the condition to argue that the effect is violated whenever the benefit function is intertemporally nonseparable. Similarly, Freixas and Laffont (1984) develop a sufficient condition for the irreversibility effect and use it, along with a numerical example, to argue that the sufficient condition is, in fact, necessary. Finally, Gollier, Jullien and Treich (2000) develop a different necessary and sufficient condition for the irreversibility effect.

In this paper we come to a somewhat different conclusion, namely that the irreversibility effect holds widely with nonlinear benefits and continuous choices. First, using one of Epstein's examples and his sufficient condition, we show that the irreversibility effect holds even when the benefit function is intertemporally nonseparable. Second, we show that the necessary and sufficient conditions developed by Freixas and Laffont and by Gollier et al. are only sufficient and not necessary. Consequently, we establish that the irreversibility effect is not necessarily violated when their conditions fail to hold. Third, we argue that two cases in which the effect is legitimately violated are somewhat special and probbaly not empirically important.

In the process of defending the irreversibility effect we also distinguish the concept of flexibility from that of irreversibility, and develop new, less restrictive definitions for each. We further show that the irreversibility effect is separate from risk aversion.

The rest of the paper is organized as follows. The irreversibility effect is defined in section 2. Section 3 reviews the existing necessary and sufficient conditions that lead to the conclusion that the effect does not hold widely. Section 4 shows that the necessary and sufficient conditions are only sufficient and section 5 establishes that the irreversibility effect holds for a case in which it appears to be violated. Section 6 discusses the two cases where the effect is in fact violated and also establishes that risk aversion can be distinguished from the irreversibility effect.

2. Irreversibility Effect: A Definition

Consider the following dynamic optimization problem:

$$\max_{x_1 \in C_1} \left(B_1(x_1) + \sum_j q_j \max_{x_2 \in C_2(x_1)} \left[\sum_i \pi_{ij} B_2(x_1, x_2, z_i) \right] \right) \tag{1}$$

where x_1 is the choice variable in the first period, x_2 is the choice variable in the second period, and z_i, for $i = 1...M$, is a discrete random variable that reflects the underlying uncertainty about the nature of net benefits. A decisionmaker chooses x_1 and x_2 to maximize expected net benefits. Net benefits in the first period, denoted by $B_1(x_1)$, are deterministic and depend only on x_1, but net benefits in the second period, denoted by $B_2(x_1, x_2, z_i)$, are stochastic and are a function of x_1, x_2, and z_i. We assume that B_1 is concave and twice continuously differentiable in x_1, and B_2 is concave and twice continuously differentiable in x_1 and x_2. Note that since B_2 is a function of x_1, the benefit function is nonseparable. If, on the other hand, B_2 were only a function of x_2 and z and not of x_1, then the benefit function would be separable. Before the second period decision is made, the decision maker receives a signal, denoted by y_j, that reveals some information about z. The amount of information contained in y depends on how closely related z and y are. Let y and y' denote two potential signals where the correlation between y and z is greater than the correlation between y' and z. y is said to be more informative about z and leads to greater learning about the true nature of z than y'. After the signal is received, the decision maker updates her prior expectations about z by formulating a posterior distribution denoted by $\pi_{ij} = p(z = z_i / y = y_j)$ and then chooses x_2 for each signal to maximize the expected benefit over the different states. Let q_j denote the probability distribution for y, C_1 the constraint function for x_1, and $C_2(x_1)$ the constraint function for x_2. Because x_1 constrains the choice of x_2, its choice in turn implies a certain loss of flexibility, and is thus the source of irreversibility. For example, $C_2(x_1) = x_1 > x_2$ implies that x_2 is constrained to be less than x_1. If, on the other hand, x_1 did not constrain the choice of x_2, then there would be no irreversibility.

Finally, assume that a unique solution exists, and lies in the interior of C_1. Let x_1^* denote the maximum corresponding to the more informative signal y, and x_1^{**} the maximum corresponding to the less informative signal y'. Define $\hat{x}_1$ as the value of x_1 that gives maximum decision making flexibility in the future. For example, if x_2 is constrained to be greater than (less than) x_1, $x_1 \in [0,1]$ and $x_2 \in [0,1]$, then $\hat{x}_1 = 0$ $(\hat{x}_1 = 1)$. This is because with $x_1 = 0$ $(x_1 = 1)$ there is no constraint on the choice of x_2, and so there is maximum decision making flexibility. An irreversibility effect exists if

$$(2) \qquad |x_1^* - \hat{x}_1| \leq |x_1^{**} - \hat{x}_1|,$$

that is, if the optimum corresponding to the more informative signal is at least as close to the point of maximum flexibility as the optimum corresponding to the less informative signal. Note that if x_2 is constrained to be less than x_1, $x_1 \in [0,1]$ and $x_2 \in [0,1]$, then since $\hat{x}_1$ is 0, equation (2) is equivalent to $x_1^* \leq x_1^{**}$. In other words, the irreversibility effect holds if $x_1^* \leq x_1^{**}$. Alternatively, if x_2 is constrained to be greater than x_1, $x_1 \in [0,1]$ and $x_2 \in [0,1]$, then since $\hat{x}_1$ is 1, equation (2) implies that the irreversibility effect holds if $|x_1^* - 1| \leq |x_1^{**} - 1|$. Since x_1 lies between 0 and 1, this simplifies to $x_1^* \geq x_1^{**}$. Our definition for the irreversibility effect allows for both these possibilities and is thus more general than the simple inequality between x_1^* and x_1^{**}. We do, however, sometimes use, rather than our definition, the simple inequality to define the irreversibility effect.

In some economic models $\hat{x}_1$ may be a constant, while in others it may be a function of the model parameters. We explore this further in section 5. Moreover, equation (2) is *not* equivalent to, and in fact is more general than, the condition that is usually used to define the irreversibility effect as well as flexibility in decision making—namely, $C_2(x_1^*) \supseteq C_2(x_1^{**})$ (Freixas and Laffont 1984). We show in section 5 that this definition for both flexibility and the irreversibility effect is too restrictive, and in some economic problems it leads to the conclusion that the irreversibility effect is violated when in fact it is not.

3. Necessary and Sufficient Conditions: A Literature Review

We now turn to a review of the relevant literature on necessary and sufficient conditions for the irreversibility effect to hold with nonlinear benefit functions and continuous choices. Epstein (1980) establishes a sufficient condition under which the initial decision with uncertainty and the possibility of future learning (x_1^*) is greater or less than the initial decision with uncertainty and less learning (x_1^{**}).[1] Irreversibility per se does not affect the sufficient condition because the constraint that defines the irreversibility effect does not enter the sufficient condition.

Using the model described in section 2, we can state Epstein's sufficient condition. Let $J(x_1, \xi)$ denote the value function, which is defined as

[1]The literature often refers to this as the precautionary effect.

$$(3) \qquad J(x_1, \xi) \equiv \max_{x_2 \in C_2(x_1)} \sum_i \pi_{ij} B_2(x_1, x_2, z_i)$$

$$(4) \qquad \equiv \max_{x_2 \in C_2(x_1)} \left[\xi_j B_2(x_1, x_2, z_i)\right]$$

where $\xi = [\xi_1, \xi_2, ...\xi_j...\xi_N]$ and $\xi_j = [\pi_{1j}, \pi_{2j}, ..\pi_{ij}.., \pi_{Mj}]$ and is a vector of the posterior probability distribution corresponding to the signal y_j. Assume that $J(x_1, \xi_j)$ is concave and differentiable with respect to x_1.[2] The sufficient condition relating x_1^* to x_1^{**} is given in theorem 1.

Theorem 1. *If $J_{x_1}(x_1^*, \xi)$ is a concave (convex) function of ξ, then $x_1^* \leq (\geq) x_1^{**}$,*

where $J_{x_1}(x_1, \xi)$ is the slope of the value function with respect to its first argument. In words, the sufficient condition states that if the slope of the value function with respect to x_1 is concave (convex) in the posterior probability distribution, then the optimal choice of x_1 associated with the more informative signal is no greater (no less) than the optimal choice associated with the less informative signal, or that $x_1^* \leq (\geq) x_1^{**}$. Consequently, even though irreversibility per se does not affect the sufficient condition, the sufficient condition can nonetheless be used to establish the irreversibility effect. Furthermore, Epstein uses this condition to establish that the irreversibility effect is violated whenever the benefit function is nonlinear and intertemporally nonseparable. He does so by developing a series of five models, three with intertemporally separable, and two with intertemporally nonseparable, benefit functions, and establishing that the irreversibility effect holds whenever the benefit function is separable and is violated whenever the benefit function is nonseparable.

Similarly, Freixas and Laffont (1984) develop a sufficient condition under which the initial decision with the possibility of learning is greater or less than the initial decision with no possibility of learning. Specifically, the irreversibility effect is said to hold if the value function is strictly quasi-concave. After establishing sufficiency analytically, Freixas and Laffont develop a numerical example to show that their sufficient condition is also necessary. The implication is that the irreversibility effect is violated whenever the value function is not strictly quasi-concave.

More recently, Gollier et al. (2000) have developed necessary and sufficient conditions for two classes of economic models to sign the second derivative of the slope of the value function in the random variable and have used Epstein's theorem to argue that these conditions are necessary and

[2] This assumption holds if $B_2(x_1, x_2, z)$ is concave in x_1 and x_2 and if for $C_2(x_1) = \{x_2 | f(x_1, x_2) \geq 0\}$, the function f is concave.

sufficient for the irreversibility effect. Firstly, within the class of models characterized by hyperbolic absolute risk aversion preferences—that is, with utility functions

$$B(x) = \frac{\gamma}{1-\gamma}\left[\eta + \frac{x}{\gamma}\right]^{1-\gamma},$$

where x is a function of x_1 and x_2 and the coefficient of absolute risk aversion is $\eta + \frac{x}{\gamma}$—they show that the slope of the value function is concave (convex) in the random variable if and only if $\gamma < 1$ ($\gamma > 1$ or $\gamma < 0$). And secondly, in models with small risks or models in which the random variable has a two-atom support, they establish that the slope of the value function is concave (convex) if and only if absolute prudence is larger (smaller) than twice the absolute aversion to risk. Furthermore, they argue that $x_1 * \leq (\geq) x_1^{**}$ if and only if the slope of the value function is concave (convex) in the random variable. These conditions imply that the irreversibility effect is violated whenever they fail to hold. Consider, for example, a model where the choice variable in the first period, x_1, and that in the second period, x_2, are constrained to lie between 0 and 1 and x_2 is constrained to be less than x_1. Note that these conditions imply that $\hat{x}_1 = 0$. Further, assume that the agent's preferences are characterized by constant relative risk aversion where the coefficient of risk aversion is greater than one. Since preferences characterized by constant relative risk aversion are equivalent to those characterized by hyperbolic absolute risk aversion with the coefficient $\eta = 0$, according to Gollier et al, the slope of the value function for this example is convex in the random variable and $x_1^* \geq x_1^{**}$. Since $\hat{x}_1 = 0$, $|x_1^* - \hat{x}_1| \not\leq |x_1^{**} - \hat{x}_1|$, and the irreversibility effect is violated.

4. Necessary versus Sufficient Conditions

We want to suggest that the irreversibility effect in fact holds more widely than one might conclude on the basis of this literature. We first establish that the two necessary and sufficient conditions, due to Gollier et al. (2000) and Freixas and Laffont (1984), are only sufficient and not necessary for the irreversibility effect to hold. While a violation of a necessary condition leads to certain rejection of the irreversibility effect, a violation of a sufficient condition does not. Consequently, by weakening the conditions, we expand the scope of the irreversibility effect.

4.1. **Gollier et al.** To see that the conditions developed by Gollier et al are not necessary, note that their conditions establish the irreversibility effect indirectly. Their conditions sign the second derivative of the slope of the value function (defined by equation 3), but say nothing directly about the relationship between the initial decision with greater learning and that with less learning, that is, whether $x_1^* \lesseqgtr x_1^{**}$ (as these are defined in section 2). Once they establish the sign of the third derivative of the value function, Gollier et al state that the sign of the third derivative is necessary and sufficient to determine whether $x_1^* \lesseqgtr x_1^{**}$. Since the sign of the third derivative is, however, only a sufficient condition for whether or not the initial decision with more learning is greater or less than the initial decision with less learning (see theorem 1), their conditions, in turn, are only sufficient for the irreversibility effect.

4.2. **Freixas and Laffont.** To see that the condition developed by Freixas and Laffont is only sufficient, consider the following dynamic optimization problem, with an additively separable benefit function, as proposed in their paper:

$$\max_{x_1}\left(B_1(x_1) + \sum_j q_j \max_{x_2 \le x_1} \sum_i \pi_{ij} B_2(x_2, z_i)\right)$$

where x_1 is the choice variable in the first period, x_2 is the choice variable in the second period, z is a random variable, q_j is the prior probability on the signal, and π_{ij} is the posterior probability distribution. As in section 2, let x_1^* be the optimal choice of x_1 under the more informative signal, x_1^{**} the optimal choice under the less informative signal, and $J(x_1, \xi)$ be the value function. Theorem 2 specifies the condition for the irreversibility effect to hold given that a unique solution exists.

Theorem 2. $x_1^* \ge x_1^{**}$ *if* $B_1(x_1) + J(x_1, \xi)$ *is strictly quasi-concave.*

Freixas and Laffont provide a numerical example to show that strict quasi-concavity is also necessary for the irreversibility effect to hold. They do so by establishing that if strict quasi-concavity is violated then, in fact, the irreversibility effect is violated. The irreversibility effect (that is, $x_1^* \ge x_1^{**}$ given that $x_2 \le x_1$) is also shown to be equivalent to $J(x_1, \xi') - J(x_1, \xi)$ being locally increasing in x_1^{**} where ξ' is the more informative signal, ξ is the less informative signal, and $J(x_1, \xi') - J(x_1, \xi)$ is the value of information.

A slight modification of the numerical example however re-establishes the irreversibility effect even when strict quasi-concavity is violated. It is also worth noting that strict quasi-concavity is not very restrictive. Most pausible benefit functions would pass this test. In the original example the random variable z is assumed to take two possible values, z_1 and z_2, each with probability 0.5. Furthermore, there are two levels of learning, perfect or none at all. The functional forms of the benefit functions are:

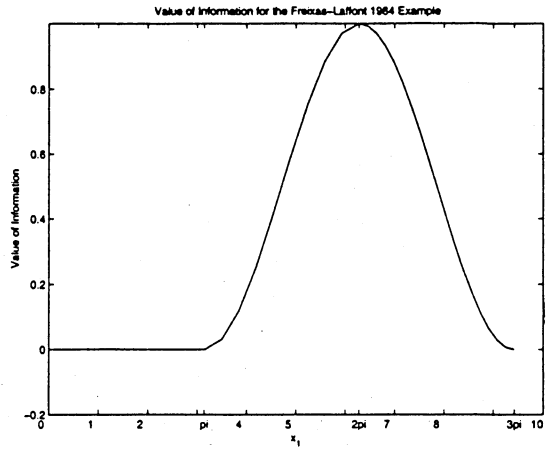

FIGURE 1. Value of Information for the Example in Freixas-Laffont 1984

$$B_1(x_1, 0 \leq x_1 \leq 2.5\pi) = \pi$$

$$B_1(x_1, x_1 \geq 2.5\pi) = -1.25(x_1 - 2.5\pi) + \pi$$

$$B_2(x_2, z_1) = 2x_2$$

$$B_2(x_2, z_2) = -\cos x_2 + 1$$

With these benefit functions strict quasi-concavity is violated, and contrary to expectations, $x_1^* \leq x_1^{**}$. $J(x_1, \xi') - J(x_1, \xi)$, the value of information, is not increasing in x_1. This is shown in figure 1 where the choice variable in the first period, x_1, is drawn on the x-axis and the value of information on the y-axis. Note that in the range $[2\pi, 3\pi]$ the value of information decreases in x_1. So long as the optima lie in this range the irreversibility effect is violated.

Now consider a slight modification where $B_1(x_1)$ and $B_2(x_2, z_1)$ remain unchanged and $B_2(x_2, z_2)$ is given by

$$B_2(x_2, z_2, 0 \leq x_2 \leq 2\pi) = -\cos x_2 + 1$$

$$B_2(x_2, z_2, x_2 \geq 2\pi) = x_2 - 2\pi$$

The original and modified benefit functions are illustrated in figure 2. Both functions are identical in the range $[0, 2\pi]$ and thereafter the original function is represented by the dotted line and the modified function by the broken line.

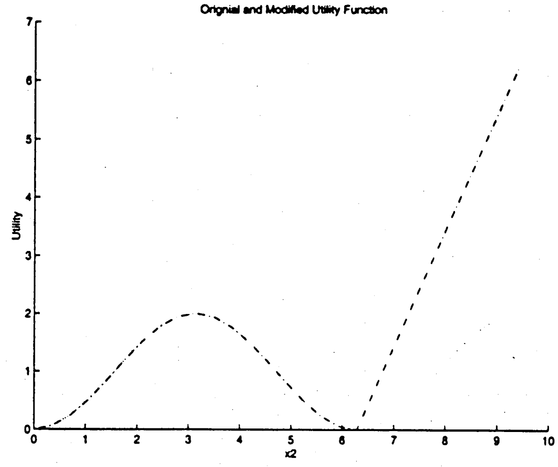

FIGURE 2. Original and Modified Utility Functions

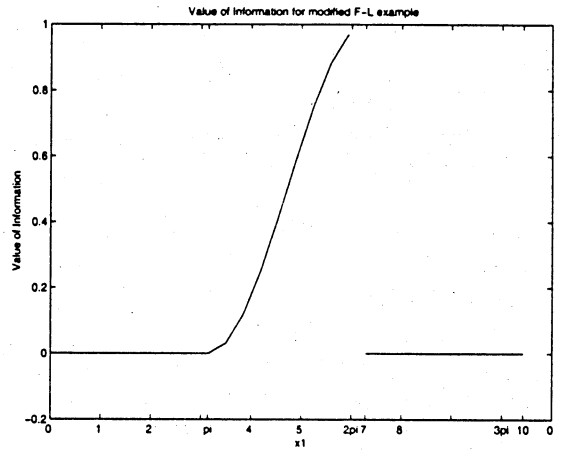

FIGURE 3. Value of Information for the Modified Example in Freixas-Laffont 1984

With the modified value function strict quasi-concavity is still violated, but now $x_1^* \geq x_1^{**}$ (strictly greater if the optimum lies between π and 2π and equal otherwise) and the value of information is a monotonically increasing function of x_1. These results are shown in figure 3. Note that with the modified benefit function the value of information no longer decreases in the range $[2\pi, 3\pi]$. The irreversibility effect holds though strict quasi-concavity is violated.

5. INTERTEMPORAL SEPARABILITY AND THE IRREVERSIBILITY EFFECT

In this section we show that the irreversibility effect holds for a class of benefit functions discussed in Epstein's analysis of a firm's demand for capital, namely for intertemporally non separable as well as separable benefit functions. In the process, we separate the definition of flexibility from that of the irreversibility effect, and develop a new, less restrictive, definition for the latter.

Consider the following problem faced by a profit maximizing firm:

$$\max_{K\geq 0}\left(-cK+\sum_j q_j \max_{L\geq 0}(\sum_i \pi_{ij}p_i F(K,L)-wL)\right)$$

where K is capital, L is labor, c is the cost of capital, w is the wage rate, F is a strictly concave production function, and p_i is the unknown output price. The firm determines its demand for capital in the first period and its demand for labor in the second period, after it receives some information about output prices. Capital is thus quasi fixed while labor is variable. In the second period the firm can neither invest nor disinvest in capital.[3] The question is how does the possibility of future learning affect the firm's demand for capital in the first period, specifically, is there an irreversibility effect?

From Epstein's sufficient condition the answer to whether the irreversibility effect holds depends on the second derivative of the slope of the value function in the random variable. For the following constant elasticity of substitution production function

$$F(K,L)=[aK^{-\beta}+bL^{-\beta}]^{\frac{-\mu}{\beta}}$$

where $a>0$, $b>0$, $\beta>-1$, $\beta\neq 0$, $0<\mu<1$ (where μ is a measure of returns to scale) and the elasticity of substitution, σ, is equal to $\frac{1}{(1+\beta)}$, Hartman (1976) has established that the third derivative of the value function depends on the relationship between the elasticity of substitution and returns to scale. Specifically, Hartman has shown that if $\sigma>(<)\frac{1}{(1-\mu)}$ then $J_K(K,p_i)$ is concave (convex) in p_i. This combined with theorem 1 implies that if $\sigma>(<)\frac{1}{(1-\mu)}$ then the

[3]Note that if the firm was allowed to invest or disinvest in capital in the second period then the problem faced by the firm would become intertemporally separable and the irreversibility effect would hold (see Narain, Hanemann and Fisher (2002) for proof that intertemporal separability is sufficient for the irreversibility effect). Consider the case where the firm is allowed to disinvest in the capital stock, at a cost, in the second period. The problem described by equation (5) would change to

(5) $$\max_{K_1\geq 0}\left(-c_1K_1+\sum_j q_j \max_{L\geq 0, K_2\leq K_1}(\sum_i \pi_{ij}p_i F(K_2,L)-wL+c_2(K_1-K_2))\right)$$

where K_1 denotes capital in the first period, K_2 denotes capital in the second period, c_1 is the cost of capital in the first period and c_2 is the cost of capital in the second period. Since there is a cost associated with disinvestment $c_2>c_1$. Equation (5) can be re-written as

$$\max_{K_1\geq 0}\left((c_2-c_1)K_1+\sum_j q_j \max_{L\geq 0, K_2\leq K_1}(\sum_i \pi_{ij}p_i F(K_2,L)-wL-c_2K_2)\right)$$

Since K_1 does not affect the benefit function in the second period, the problem is intertemporally separable. A similar case can be made for when the firm is allowed to invest in the second period.

demand for capital is lower (higher) when there is a possibility of learning than when there is no possibility of learning. Since the demand for capital does not unambiguously increase or decrease with learning Epstein concludes that this is evidence that the irreversibility effect is violated and furthermore, that the irreversibility effect does not hold for intertemporally nonseparable benefit functions.

We shall argue instead that even if the slope of the value function is concave for some parameter values and convex for other values there may still be an irreversibility effect, because the flexible value of capital, defined below, also changes with the parameters. Note that the firm can neither increase nor decrease its capital stock in the second period. Consequently, *a priori* one cannot tell whether a high or a low demand for capital in the first period constitutes a flexibility-enhancing decision. When σ is high so that capital and labor can be easily substituted then a lower capital stock today may very well give the firm greater flexibility tomorrow. If it turns out that the firm has underestimated production needs, it can compensate for the low stock of capital by hiring more labor. On the other hand if σ is low so that capital and labor cannot be substituted then a higher capital stock today may maintain greater flexibility tomorrow. If this is so then when $\sigma > (<) \frac{1}{(1-\mu)}$ a decrease (increase) in the demand for capital when there is a possibility of learning constitutes an irreversibility effect. We show that this is in fact the case and that the model does give rise to an irreversibility effect. We first define what is meant by flexibility in this context and then show that the level of capital that gives the greatest amount of flexibility is lower (higher) when $\sigma > (<) \frac{1}{1-\mu}$.

5.1. **Definition of Flexibility.** There are two definitions of flexibility in the literature. One is expressed in terms of the set of choice variables and the other, made precise by Jones and Ostroy (1984), is in terms of the set of second period positions that can be attained from the first period position at a given cost and for a particular state of the world. By the first definition x_1^* is said to be more flexible than x_1^{**} if $C_2(x_1^*) \supseteq C_2(x_1^{**})$, that is, if the choice set associated with x_1^* is larger than the choice set associated with x_1^{**}. For the second definition, let $c(x_1, x_2, z_i)$ denote the cost of moving from x_1 to x_2 given that the state of the world is z_i. Then $G(x_1, z_i, \alpha)$, where

$$G(x_1, z_i, \alpha) \equiv \{x_2 : c(x_1, x_2, z_i) \leq \alpha\},$$

is the set of second period positions attainable from x_1 at a cost that does not exceed α in state s. In general x_1^* is said to be more flexible than x_1^{**} when for all $\alpha \geq 0$ and for all z_i, $G(x_1^*, z_i, \alpha) \supseteq G(x_1^{**}, z_i, \alpha)$.

If we attempt to apply the first definition of flexibility to Epstein's model, since capital can neither be increased nor decreased in the second period, and letting x_1 and x_2 represent the levels of capital chosen in the first and the second period respectively, the set $C_2(x_1)$ is empty. Defining x_2 as the level of labor chosen in the second period does not help to determine the level of capital that gives more or less flexibility in the second period either as the first period's choice of capital in no way restricts the choice of labor in the second period. Under this definition, flexibility is independent of the choice in the first period, a fact that makes the definition too restrictive. Since the set C is always either empty (if capital is the choice variable in the second period), or equal to the entire set of labor choices (if labor is the choice variable in the second period), and thus does not vary with the level of learning, the set C cannot be used to define the irreversibility effect. Our definition for the irreversibility effect (see equation 2), on the other hand, with $\hat{x}_1$ appropriately defined, can still be used.

Under the definition of flexibility due to Jones and Ostroy, so long as x_1 and x_2 are defined in terms of the choice variables, it similarly appears that the set $G(x_1, z_i, \alpha)$ is either empty, if the choice variable in the second period is capital, or equal to the entire set of choices, if the choice variable in the second period is labor. Note that up to this point we have measured flexibility in terms of the choice variables, that is, in terms of the choices of capital or labor in the second period that are feasible given the choice of capital in the first period. One could instead measure flexibility in terms of the level of output that can be attained in the second period given the choice of capital in the first. Presumably the firm cares about the level of capital, or any other input, only in so far as it allows the firm to produce output in the second period. With this alternative measure of flexibility, if we define x_1 as the level of capital chosen in the first period, x_2 as the level of output attained in the second period, z_i as the price of output in the second period and α as the wage rate (which translates into units of labor), then $G(x_1, z_i, \alpha)$ can be defined as the set of outputs that can be attained for given levels of capital and labor and for a particular price of output. With this interpretation the set G is no longer independent of the choice variable in the first period.

The question still remains as to how one compares the set G for different levels of z_i. One possibility is to define the set in terms of the *range of output*[4] that can be attained for a given level of capital and to define x_1^* as being more flexible than x_1^{**} if the range of output attained by x_1^* is greater than the range attained by x_1^{**}. This seems reasonable as flexibility for the firm does

[4]Note that this is consistent with Hirshleifer and Riley (1992) who point out that flexibility is different from the range of actions which in our example would mean the range of capital or labor. We instead equate flexibility to the range of outputs.

manifest itself in terms of the range of output that the firm can produce. If the firm learns that the price of output is likely to be high tomorrow, it would like to produce a high output and if it learns that the price is likely to be low, it would prefer a low output.

5.2. **Proof of the Irreversibility Effect.** With this understanding of flexibility, we now prove that the most flexible level of capital changes with a change in the parameters and furthermore, that the level of capital that gives the greatest amount of flexibility is lower (higher) when $\sigma > (<)\frac{1}{1-\mu}$.

Proposition 1. *If* $\sigma > (<)\frac{1}{1-\mu}$ *then* $\hat{K} = \underline{K}(\overline{K})$,

where $\hat{K}$ is the level of capital that implies the greatest amount of flexibility, $\underline{K}$ is the minimum capital stock and $\overline{K}$ is the maximum capital stock.

Proof. Let $\bar{y}(K)$ denote the range of output that can be achieved for a given level of capital and let $\gamma = \frac{-\mu}{\beta}$. Note that when $\sigma > (<)\frac{1}{1-\mu}$, $\gamma < (>)1$ since $\sigma = \frac{1}{1+\beta} > (<)\frac{1}{1-\mu}$ implies that $\mu < (>) - \beta$.

$$\bar{y}(K) = (aK^{-\beta} + b\overline{L}^{-\beta})^{\gamma} - (aK^{-\beta} + b\underline{L}^{-\beta})^{\gamma}$$

where $\underline{L}$ is the minimum labor and $\overline{L}$ is the maximum labor. The derivative of the range of output with respect to the capital stock is given by

$$\frac{\partial \bar{y}}{\partial K} = -a\gamma\beta K^{-(\beta+1)}\left((aK^{-\beta} + b\overline{L}^{-\beta})^{\gamma-1} - (aK^{-\beta} + b\underline{L}^{-\beta})^{\gamma-1}\right)$$

When $\gamma < (>)1$, $\frac{\partial \bar{y}}{\partial K} < (>)0$. This in turn implies that when $\gamma < (>)1$ then the level of capital that gives the maximum range of output, $\hat{K}$, is equal to the minimum (maximum) stock of capital. □

By our definition of the irreversibility effect (see equation 2), the effect holds for $\sigma > (<)\frac{1}{1-\mu}$ if $|K_1^* - \hat{K}_1| \le |K_1^{**} - \hat{K}_1|$, or $|K_1^* - \underline{K}| \le |K_1^{**} - \underline{K}|\left(|K_1^* - \overline{K}| \le |K_1^{**} - \overline{K}|\right)$, or $K_1^* \le K_1^{**}$ $(K_1^* \ge K_1^{**})$, where K_1^* is the choice with greater learning and K_1^{**} is that with less learning. Since this is in fact the case, we have shown that the irreversibility effect holds for this example. Moreover, this proves that the irreversibility effect is not violated in models with intertemporally nonseparable benefit functions.[5]

[5]Note that if the elasticity of substitution is equal to 1, that is, $\sigma = 1$, so that the production function is a Cobb-Douglas, then $\gamma = 0$ and a higher level of capital, unambiguously, gives a greater range of output. If $\sigma = 0$, so that the production function is a Leontief, then it is difficult to determine what level of capital gives greater flexibility tomorrow.

6. Two cases of Violation

We now come to the two cases where the irreversibility effect is violated to argue that these are somewhat special, or perhaps empirically unimportant. The first case is where the benefit or utility function is not strictly quasi-concave, as in the numerical example discussed by Freixas and Laffont (1984). Utility or profit functions are however normally specified as concave, or at least quasi-concave, in accordance with intuition and some evidence. Further, as we show in section 4, strict quasi-concavity is not necessary for the irreversibility effect to hold, merely sufficient. That is, the effect can hold even when strict quasi-concavity is violated.

The second case arises in the consumption and savings problem discussed by Epstein. In this problem an individual allocates an initial amount of wealth between consumption and savings over three periods. Investment in the first period yields a fixed return while investment in the second period yields a random return. Some information is gained about the random rate of return at the beginning of the second period. In symbols, the problem is:

$$\max_{0\leq x_1\leq w} B_1(w-x_1)+\beta\sum_j q_j \max_{0\leq x_2\leq rx_1}\left(B_2(rx_1-x_2)+\beta\sum_i \pi_{ij}B_3(x_2 z_i)\right)$$

where x_1 and x_2 denote savings in periods 1 and 2 respectively, w is the initial wealth, r is the sure gross rate of return to the first period savings, β is the discount factor, z_i is the random gross return to second period savings, B_1 is the utility function in the first period, B_2 is the utility function in the second period, and B_3 is the utility function in the third period.

With the following constant relative risk aversion utility function,

$$B(c)=\begin{cases}\frac{c^{1-\alpha}}{(1-\alpha)} & \text{if } \alpha\neq 1,\\ \log c & \text{if } \alpha=1,\end{cases}$$

where α is the coefficient of relative risk aversion, Epstein shows that the effect of learning on the optimal level of savings in the first period depends on the elasticity of intertemporal substitution, that is, on $\sigma=\frac{1}{\alpha}$. When $\sigma>1$ the slope of the value function is convex and the possibility of learning about the future rate of return leads to an *increase* in savings in the first period. On the other hand, when $\sigma<1$, the slope of the value function is concave and the possibility of learning leads to a *decrease* in the level of first period savings. With $\sigma>1$ consumption between periods is highly substitutable and the decisionmaker cares more about the sum of expected utility across the three periods and less about consumption smoothing between periods. This in turn implies that faced with uncertainty about the rate of return to second period savings and the possibility of learning about the return, the decisionmaker postpones consumption until she has better information. This in turn helps avoid a situation where the rate of return is revealed to be higher than expected and savings are too low to exploit the situation. If, on the other hand, gross return to second period savings is lower than expected then she can simply increase consumption in the second period without any loss of utility. Consequently, first period savings are increased. With $\sigma < 1$ consumption between periods is not highly substitutable and the decisionmaker cares more about consumption smoothing than about total expected utility. Furthermore, learning enables her to better allocate consumption between the second and the third periods thus increasing the sum of expected utility in those periods. Consequently, first period savings are reduced to balance consumption between the different periods.

Since savings in the second period are constrained to be no greater than the gross rate of return times savings in the first period and since the level of savings does *not* unambiguously increase with learning, this is evidence that the irreversibility effect is violated in this problem. Specifically, the irreversibility effect is violated when $\sigma < 1$ as $x_1^* \leq x_1^{**}$ while $\hat{x}_1 = w$ and thus $|x_1^* - w| \not\leq |x_1^{**} - w|$. Note that the violation is not due to the fact that the net benefit function is intertemporally nonseparable but rather due to the fact that the benefits are *intertemporally non-substitutable* or the *coefficient of relative risk aversion is large.* Under these conditions the decisionmaker cares more about benefits in each period than about total benefits. Note that this is somewhat of a special case as it rules out what we may call the standard specification of the benefits of an investment, in which they are simply summed over time, discounted as appropriate, with no restrictions on the relationship of one period's benefits to those of another. Under the standard specification the irreversibility effect continues to hold.

The consumption and savings example also enables us to shed light on another issue in the literature on the irreversibility effect, namely whether risk aversion can be separated from the irreversibility effect. Kolstad (1996) has argued that it can, while Gollier et al suggest not. To analyze this issue we need to consider non-expected utility preferences, as in these, unlike in Von-Neumann-Morgenstern preferences, the coefficient of relative risk aversion is not constrained to be the reciprocal of the elasticity of intertemporal substitution. The effect of risk aversion can thus be separated from that of intertemporally substitutability.

Consider the following generalized isoelastic preferences instead of the constant relative risk aversion utility function (otherwise the problem is the same as Epstein's consumption-savings example):

$$J_t = B(c_t, E_t J_{t+1})$$
$$= \frac{\left((1-\beta)c_t^{1-\rho} + \beta[1 + (1-\beta)(1-\alpha)E_t J_{t+1}]^{\frac{1-\rho}{1-\alpha}}\right)^{\frac{1-\alpha}{1-\rho}} - 1}{(1-\beta)(1-\alpha)}$$

where $\beta \in (0,1)$, $\alpha > 0$ and is the coefficient of relative risk aversion and $\frac{1}{\rho} = \sigma$ is the elasticity of intertemporal substitution. Note that σ is no longer constrained to be equal to $\frac{1}{\alpha}$. Unfortunately, with isoelastic preferences it is hard to establish the relationship between the convexity or concavity of the slope of the value function and α and ρ analytically and so we turn to numerical simulations to separate the effects of risk aversion and intertemporal substitution. Our simulations, which compare the optimal choice of savings in the first period with perfect learning and with no learning for a wide range of parameter values for α and ρ, give the results in table 1.

Table 1: Experiments with Generalized Isoelastic Preferences

	$\alpha < 1$	$\alpha > 1$
$\sigma < 1$	$x_1^* \lesseqgtr x_1^{**}$	$x_1^* \lesseqgtr x_1^{**}$
$\sigma > 1$	$x_1^* \geq x_1^{**}$	$x_1^* \geq x_1^{**}$

When $\sigma < 1$ it is feasible for $x_1^* < x_1^{**}$, that is, for savings to decrease with learning for both $\alpha < 1$ and $\alpha > 1$. However, when $\sigma > 1$ x_1^* is always at least as large as x_1^{**} irrespective of the coefficient of relative risk aversion. This implies that though the irreversibility effect is violated even with non-expected utility preferences, the violation is caused by a low elasticity of intertemporal substitution and not by a high coefficient of relative risk aversion. This numerical result provides some support for an assertion by Epstein that violation of the irreversibility effect can be attributed to intertemporal substitution rather than to risk aversion. Furthermore, it supports Kolstad's assertion that the irreversibility effect can be separated from risk aversion.

7. Concluding Remarks

The central point of the paper is that the irreversibility effect holds more widely than one might conclude from the existing literature. Conditions suggested as necessary and sufficient appear to be only sufficient. Further, for an important case in which the effect appears not to hold, intertemporally nonseparable benefit functions, we show that it does, based on a careful analysis of the concept of flexibility. Finally, of the two cases where the effect is violated, the numerical example discussed by Freixas and Laffont and the consumption-savings problem discussed by Epstein, the former appears, to say the least, to be an unusual problem, one that does not correspond to an ordinary economic model, as the violation of the irreversibility effect involves a violation of quasi-concavity. The consumption-savings problem generates a violation of the effect if in addition to being intertemporally nonseparable, preferences are intertemporally non-substitutable. Consequently, even for this problem, if the decisionmaker cares about total benefits over time and not too much about benefits in each period, the irreversibility effect holds.

References

Arrow, K. J. and Fisher, A. C.: 1974, Environmental preservation, uncertainty, and irreversibility, *Quarterly Journal of Economics* **88**(2), 312–319.

Epstein, L. G.: 1980, Decision making and the temporal resolution of uncertainty, *International Economic Review* **21**, 269–283.

Freixas, X. and Laffont, J.-J.: 1984, On the irreversibility effect, *in* M. Boyer and R. Kihlstrom (eds), *Bayesian Models in Economic Theory*, NHPC, pp. 105–114.

Gollier, C., Jullien, B. and Treich, N.: 2000, Scientific progress and irreversibility: an economic interpretation of the precautionary principle, *Journal of Public Economics* **75**, 229–253.

Hartman, R.: 1976, Factor demand with output price uncertainty, *American Economic Review* **66**, 675–681.

Henry, C.: 1974, Investment decisions under uncertainty: The irreversibility effect, *American Economic Review* **64**, 1006–12.

Hirshleifer, J. and Riley, J.: 1992, *The Analytics of Uncertainty and Information*, Cambridge University Press.

Jones, R. A. and Ostroy, J. M.: 1984, Flexibility and uncertainty, *Review of Economic Studies* **51**, 13–32.

Kolstad, C. D.: 1996, Fundamental irreversibilities in stock externalities, *Journal of Public Economics* **60**, 221–233.

Narain, U., Hanemann, W. and Fisher, A.: 2002, The irreversibility effect: Necessary versus sufficient conditions, Resources for the Future Discussion Paper.