

THE VIRTUES OF HESITATION[‡]

URMEE KHAN AND MAXWELL B. STINCHCOMBE

ABSTRACT. In many economic, political and social situations, circumstances change at random points in time, reacting is costly, and changes appropriate to present circumstances may be inappropriate after later changes, requiring further costly change. Waiting is informative if the hazard rate for the arrival of the next change is non-constant. In a broad range of contexts we show that hesitation or delayed reaction is often optimal when the hazard rate is decreasing, and that it might be optimal never to change at times when the hazard rate is increasing. The first result corresponds to having waited long enough to know that future changes in circumstances are comfortably in the future, the second corresponds to the counsel of patience in unsettled circumstances. These results provide a new set of motivations for building delay into legislative and other decision systems. These results in non-stationary dynamic optimization arise from extensions to semi-Markovian decision theory.

Key Words and Phrases. Hazard Rates, Semi-Markov Decision Theory, Optimal Timing

Date: January, 2011

[‡]Many thanks to Martin Dumav and Zahid Arefin Chowdhury for their timely and much appreciated help, and especially to participants at the IHS Summer Research Conference for many helpful comments. All remaining errors are ours. The first author gratefully acknowledges the financial support of a Summer Research Fellowship from the Institute of Humane Studies.

A handful of patience is worth more than a bushel of brains. - Dutch Proverb

Patience has its limits. Take it too far, and it's cowardice. - George Jackson

The essential ingredient of politics is timing. - Pierre Trudeau

Timing is everything - Attributed to various authors

1. INTRODUCTION

In many social, economic and political situations, there is a stochastic environment that changes at random points in time and changes in action are costly. Actions are costly, and we know the current state may give way to another new state at some random time in the future, potentially making today's optimal action again obsolete. The question is whether to take an action in response to a change in the environment or to delay any change (or changes).

Variants of this problem have been extensively analyzed in economics (for example [Boyarchenko and Levendorski \[2007\]](#), [Stokey \[2009\]](#) and the references therein). However, a crucial aspect of most existing analyses is that the passage of time by itself does not reveal any information. By contrast, we study problems in which the passage of time without a change contains information about the arrival time of the next change. In such problems, there may be value to delaying decisions beyond the usual option value of waiting. We begin with examples where the time which a change has survived may be of crucial importance to its future longevity. We begin with some examples.

1.1. Political Change. Political process in a democratic system are driven by 'political issues' and the configuration of opinions and attitudes of the polity on these issues. Such configurations are hardly, if ever, static. There are slow and gradual changes that take place side by side with rapid and explosive changes. Some changes are long-lasting, some short-lived. As [Carmines and Stimson \[1990\]](#) say:

... we shall see that issues, like species, can evolve to fit new niches as old ones disappear. But, unless they evolve to new forms, all issues are temporary. Most vanish at their birth. Some have the same duration as the wars, recessions, and scandals that created them. Some become highly associated with other

similar issues or with the part system and thereby lose their independent impact. And some last so long as to reconstruct the political system that produced them

Vietnam War and the Watergate scandal seem to have very little traces left today either in public attitude or legislative response to the issues of war and executive power respectively. But they were the biggest issues of their day. On the other hand, the Civil Rights Movement and its aftermath marked a fundamental realignment in US politics. In general, some ideas and opinions “wear out their welcome” after a time, perhaps through changes in the conditions that gave rise to them, perhaps by the accumulation of counterarguments to their veracity. Hence, the likelihood that such an idea would become irrelevant increases with time. By contrast, some types of issues or opinions tend to get more entrenched the longer they live. Political actors in various capacities try to cope and make decisions in the face of such ‘issue evolution’ [Carmines and Stimson, 1990]. Legislatures choose whether or not to change a law, Supreme Courts decides whether or not to re-interpret or overturn past precedents, political parties decide whether or not to realign politically and redefine the agenda. Oftentimes, the most crucial ingredient in such decisions is the aspect of timing.

Each of these decisions entail some fixed cost either to the society at large or the actor herself.¹ One would have to trade off the immediate gains with substantial future losses if the initial change that triggered the costly action turns to be rather short-lived.

1.2. Constitutional Amendments. Constitutions establish the fundamental legal structures of a society. They are meta-institutions through which institutions are introduced, reformed and interpreted [Ostrom, 1990]. A constitution and the legal order it creates must have the support of, or at least tacit approval of, the governed to have legitimacy. Maintaining the legitimacy and relevance of a constitution require a certain degree of adaptability or flexibility to change because technology, environment and public opinion are forever changing. On the other hand, the basic value of a constitution lies in its stability because

¹In case of legislative changes and court decisions, the citizens have to re-adjust and re-optimize with respect to the new rules. In case of a political party, realignment may mean losing a traditional support base.

it coordinates the actions and expectations of people and reduces the uncertainty in the environment [Hardin, 2003]. Hence the basic tradeoff between ‘commitment’ and ‘flexibility’ lies at the heart of the constitution design problem, as encapsulated in the famous exchange between Thomas Jefferson and James Madison (Smith [1995]; Madison [1961]).² It is also costly to change the constitution because it acts as a coordination device for peoples’ behavior, and changes are likely to impose large adjustment costs on significant parts of the population [Hardin, 2003] and disrupt ancillary institutions that grow around the constitution.

From these perspectives, it is reasonable to presume that an optimal rule for constitutional change should be more sensitive to long-lasting changes than to transitory changes. It is clear that waiting longer will help answer whether a change will have a longer or shorter total life, but what matters for decisions is the longer or shorter *future* life of the change. One tradeoff is between costly unneeded or ultimately unwanted changes (e.g. Prohibition) and undermining the legitimacy of the constitutional regime by ignoring new realities. It is from this perspective that we study the general question of why some changes in laws should be more difficult to implement, and what this should depend on. Under study is a class of explanations that we regard as complementary to the many previously offered ones, a class of explanations based on the observation that the persistence of changes in sentiment have predictive power for the future length of time the changes will last. For us the question becomes “How much longer should one wait before acting?”

The US constitution has had four different amendments that have extended voting rights to different parts of the population: Amendment XV (1870), which was passed at the end of the Civil War, extended suffrage to men independent of race or previous condition of servitude; XIX (1920) extended suffrage to women; XXIV (1964) made poll taxes illegal; and XXVI (1971) extended suffrage to those eighteen years of age or older. These formalized long-lived widely-shared changes in sentiment, but Amendment XVIII, Prohibition in 1919, was an expensive and short-lived failure, being repealed fourteen years later by Amendment XXI (1933).³

²for interesting empirical evidence to the effect that flexibility actually helps the sustainability of constitutions, see Elkins et al. [2009]

³For a detailed history of all amendments, see Amar [2006].

If one dates the beginning of the women's suffrage movement to the 1848 Seneca Falls Convention,⁴ it took 72 years, until 1920, for the 19'th Amendment to pass. At various points in the political process, there was evidence that the recognition of women's rights to vote would be long-lasting: the passage of suffrage at the state level in western states by the early 20'th century;⁵ the nation's westward expansion and the Civil War led to an expanded need for women both in industrial settings and as teachers; the slow increase in the numbers of college educated and professional women; unionization movements among female professions in the late 1800's and early 1900's. Even after one could perhaps clearly see that general sentiment had shifted in favor of the Nineteenth Amendment, there was (much) further delay in implementing what turns out to have been a long-lasting change in sentiment, perhaps consistent with unwillingness to believe that so drastic a change could be long-lasting.

By contrast, Amendment XVIII (Prohibition, 1919) proved to be very costly to society, and was short-lived, repealed fourteen years later.⁶ The Temperance Movement had as long a history as the women's suffrage movement, and was even used by some women's suffrage organizers as an occasion to teach women the necessity of having a voice in politics in order to achieve changes (Flexner and Fitzpatrick [1996]). From our point of view, this is a change of action that led to a change in the distribution of the time until general sentiment was reversed. This is an example of a more complicated scenario where one is not only wondering about how long the current state would last, but also has to consider the fact that her choice of action might actually affect the timing and nature of the next change.

1.3. Marketing Strategy. Research in consumer behavior has shown that when and how consumers switch brands depend on the last purchased brand and time since the last purchase. The inter-purchase time may exhibit increasing or decreasing hazard rates depending on the consumer characteristics named

⁴Flexner and Fitzpatrick [1996] emphasize the experience of female abolitionists and fighters for women's education in the early 19'th century as the roots of the suffrage movement.

⁵By 1915, Arizona, California, Colorado, Idaho, Illinois, Kansas, Montana, Nevada, Oregon, Utah, Washington, and Wyoming had granted full women's suffrage, and several other states or municipalities had granted suffrage in primary elections.

⁶Prohibition was repealed by the only Amendment to be passed by state ratifying conventions rather than by state votes.

“inertia” or “variety seeking,” and these change over time since the last purchase [Chintagunta, 1998]. It has been suggested that optimal timing of targeting consumers for marketing should depend on such considerations instead of the traditional demographic variables (Chintagunta [1998], see also Gonul and Ter Hofstede [2006] for an empirical approach to optimal timing for catalog mailing). The class of optimization models under study here are directly applicable to such situations.

1.4. Outline. The next section contains two simple examples that give a sense of what is involved in the more general analyses that follow. The essential aspects of the model include: a starting state and action, s_0 and a_0 ; random times Y_k at which the state changes from S_{k-1} to S_k according to an imbedded Markov process; and the option to engage in costly actions changes during the interarrival times $W_k = Y_{k+1} - Y_k$.

The within interval problems, from Y_k to Y_{k+1} will be central to the analysis, and the first example, on optimal search duration, highlights the role of the hazard rates for the W_k . The second example demonstrates how the value functions for the problem interact with the within interval problems.

The general model, existence of optima, and their recursive characterization through the value function follow. The following section develops the corresponding first order conditions (Euler equations) for a broad range of problems. The last section concludes.

2. TWO EXAMPLES

A pair of non-stationary problems demonstrate the essential features of both the optimization problems under study and of their solutions. The first problem is about the determinants of optimal search duration and highlights the role of changing hazard rates in both first and second order conditions for an optimum. The second problem is about optimal adaptation to changing circumstances and highlights the role of stochastic intervals.

2.1. Optimal Search Duration. At a flow cost of $c > 0$, one can keep searching for a source of higher profits (a low cost source of a crucial input, a process breakthrough, a new product). If found, expected net profits of $\bar{\pi}$ result. If one abandons the search, the alternative yields expected net profits of $\underline{\pi}$,

$\bar{\pi} > \underline{\pi} > 0$. Let W denote the waiting time till the source is found. We will assume throughout that waiting times have densities on $[0, \infty)$, hence having no atoms, except perhaps at ∞ . If W has an atom at ∞ , it is called an **incomplete** distribution, which corresponds, in the present search problem, to the object of search not existing or not being findable.

Since one optimally searches in the more likely locations or ideas first, we expect the arrival rate of W to be decreasing over time. The non-constancy of the hazard rate makes the problem non-stationary. The non-stationary choice problem is at what time, t_1 , does one stop searching and accept the lower $\underline{\pi}$? The results are special cases of Theorem 4 (below), but we give both an intuitive and a more formal development of the first order and the second order conditions for an optimal $0 < t_1^* < \infty$ for this problem.

- First order conditions: the expected benefits of waiting an extra instant dt at t_1 are $(\bar{\pi} - \underline{\pi})h_W(t_1)$, the expected costs are $(c + r\underline{\pi})$ because $r\underline{\pi}$ is the perpetual annuity flow value of $\underline{\pi}$. At an interior optimum, $0 < t_1^* < \infty$, the necessary first order conditions are $(\bar{\pi} - \underline{\pi})h_W(t_1^*) = (c + r\underline{\pi})$.
- Second order conditions: in order for the solution just given to be a local maximum rather than a local minimum, the benefits of waiting must be positive before t_1^* and negative after t_1^* . For this to be true, the hazard rate must be decreasing, $h'_W(t_1^*) < 0$.

In order to give a more formal analysis, note the following:

- (1) if $1_{[0, t_1)}(W) = 1$, i.e. if $W < t_1$, one incurs the search cost $\int_0^W (-c)e^{-rt} dt$ and receives the discounted profits of $\bar{\pi}e^{-rW}$; and
- (2) if $1_{[t_1, \infty)}(W) = 1$, one incurs the search cost $\int_0^{t_1} (-c)e^{-rt} dt$ and receives the discounted profits of $\underline{\pi}e^{-rt_1}$.

Thus the problem is

$$(1) \quad \max_{t_1 \in [0, \infty]} \mathbb{E} \left[1_{[0, t_1)}(W) \left(\int_0^W (-c)e^{-rt} dt + \bar{\pi}e^{-rW} \right) + 1_{[t_1, \infty)}(W) \left(\int_0^{t_1} (-c)e^{-rt} dt + \underline{\pi}e^{-rt_1} \right) \right].$$

Evaluating the terms in rounded brackets and rewriting yields

$$(2) \quad \psi(t_1) = \int_0^{t_1} \left(-c \frac{1}{r} (1 - e^{-rw}) + \bar{\pi} e^{-rw} \right) f_W(w) + \left(-c \frac{1}{r} (1 - e^{-rt_1}) + \underline{\pi} e^{-rt_1} \right) G_W(t_1).$$

Taking derivatives with respect to t_1 , using $G'_W = -f_W$, and rearranging yields

$$(3) \quad \psi'(t_1) = \left(e^{-rt_1} G_W(t_1) \right) [(\bar{\pi} - \underline{\pi}) h_W(t_1) - (c + r\underline{\pi})].$$

As $e^{-rt_1} G_W(t_1) > 0$, $\psi'(t_1^\circ) = 0$ only if $(\bar{\pi} - \underline{\pi}) h_W(t_1^\circ) - (c + r\underline{\pi}) = 0$, yielding

$$(4) \quad \psi''(t_1^\circ) = \left(e^{-rt_1^\circ} G_W(t_1^\circ) \right)' [0] + (\bar{\pi} - \underline{\pi}) h'_W(t_1^\circ),$$

which can only be strictly negative if $h'_W(t_1^\circ) < 0$. Interior strict optima, t_1^* , are indicated by (3) being satisfied and $h'_W(t_1^*) < 0$, which makes the comparative statics of t_1^* immediate: decreasing in c , r , and $\underline{\pi}$, increasing in $\bar{\pi}$, and increasing in uniform upward shifts of the hazard rate.

If W has a negative exponential distribution with parameter λ , the hazard rate is constant at λ and $\psi'(t_1) \lesseqgtr 0$ as $\lambda \lesseqgtr \frac{c+r\underline{\pi}}{\bar{\pi}-\underline{\pi}}$, so that $t_1^* = 0, \infty$ or $[0, \infty]$ depending on the sign of ψ' , which does not vary with t_1 . Of particular interest are the cases of monotonically increasing and decreasing hazard rates.

A Weibull distribution with parameters λ and γ is of the form $W = X^\gamma$ where X has a negative exponential(λ) distribution. The associated hazard rate is $h_W(t) = \frac{\lambda}{\gamma} t^{\frac{1-\gamma}{\gamma}}$.

- (1) If $\gamma > 1$ in the Weibull case, then $h_W(0+) = \infty$ and the hazard rate is strictly decreasing to 0, which means that there is always a unique optimal strictly positive delay before ending search.
- (2) If $\gamma < 1$ in the Weibull case, then the hazard rate starts at 0 and increases without bound.⁷ Depending on r , λ and γ , the optimal strategy at $t = 0$ may be to end search immediately, $t_1^* = 0$, or to wait until success, $t_1^* = \infty$. Even if $t_1^* = 0$ at $t = 0$, because $\gamma < 1$, there will always be a time T with the property that if one has already waited until T , then the conditionally optimal choice is $t_1^*(T) = \infty$.

⁷An interpretation of the increasing hazard rate is that learning-by-doing in process of search makes search more and more effective over time.

2.2. Optimal Adaptation to Circumstances. We now study a simple model of the optimal timing of adaptations to a stochastic dynamic state. The first set of random variables used to describe the problem are a Markov process, $\{X_k : k = 0, 1, \dots\}$ taking values in a two state set, $S = \{s', s''\}$, with the transition matrix $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. The second set of random variables are arrival times, Y_0, Y_1, \dots , represent the switching times for the states. These are non-negative random variables that satisfy $Y_0 \equiv 0$ with $W_k := Y_k - Y_{k-1}$ being i.i.d. non-negative random variables with densities on $[0, \infty)$, hence having no atoms, except perhaps at ∞ when they are incomplete.

Our first use of stochastic intervals is to define the continuous time process that the decision maker is reactively adapting to. A **stochastic interval** is a subset of $\Omega \times [0, \infty)$ of the form $\llbracket Y_k, Y_{k+1} \llbracket = \{(\omega, t) : Y_k(\omega) \leq t < Y_{k+1}(\omega)\}$. The Markov chain and the waiting times combine to form the continuous time stochastic process $(\omega, t) \mapsto X(\omega, t)$ defined by

$$(5) \quad X(\omega, t) = \sum_k X_k 1_{\llbracket Y_k, Y_{k+1} \llbracket.$$

Thus, if $X_0 = s''$, then the state remains s'' until Y_1 , at which point it switches to s' , where it stays until Y_2 , when it switches back, and so on.

There are two possible actions, $A = \{a', a''\}$. The flow payoffs to being in state

s and taking action a are given by

	s'	s''
a'	1	0
a''	0	2

. This means that one always

wants the action to match the state, matching action to state when $X_t = s'$ earns a flow of 1, matching action to state when $X_t = s''$ earns a flow of 2, and mis-matching earns a flow of 0. What may make instantaneous adjustments suboptimal is the cost of switching actions, $C > 0$. If another change in the state is expected soon, it may not be worth incurring C to enjoy the extra flow. The question is which changes in state to react to? and after what amount of delay?

The **value function**, $V_*(a, s)$, gives the maximal expected discounted utility to starting at $t = 0$ in state $s \in S$ with the present action being a . From Theorem 2 (below), $V_*(a, s)$ is well-defined, and further, it is achievable by sequentially solving an optimization problem within each stochastic interval $\llbracket Y_k, Y_{k+1} \llbracket$.

For the present model, the decision problem within an interval $\llbracket Y_k, Y_{k+1} \rrbracket$ is to pick a time $t_1 \in [0, \infty]$ at which to change actions and incur the cost C . There will be two cases: $Y_{k+1} < Y_k + t_1$, i.e. $1_{[Y_k, Y_k+t_1)}(Y_{k+1}) = 1$, corresponding to the new change in state arriving before the planned change in action; and $Y_{k+1} > Y_k + t_1$, i.e. $1_{[Y_k+t_1, \infty)}(Y_{k+1}) = 1$, corresponding to the new change in state arriving after the planned change in action. It is clear that if $(a, s) = (a', s')$ or $(a, s) = (a'', s'')$, then $t_1^* = \infty$ is optimal because any change both incurs C unnecessarily and loses flow payoff. Subtracting Y_k , setting $k = 0$ and $W = Y_1 - Y_0$, and letting $\kappa = \mathbb{E} e^{-rW}$, the value function satisfies

(6)

$$\begin{aligned}
V_*(a', s') &= \frac{1}{r}(1 - \kappa) + \kappa V_*(a', s''), \\
V_*(a', s'') &= \max_{t_1 \in [0, \infty]} \mathbb{E} \left[1_{[0, t_1)}(W) \left(\int_0^W 0e^{-rt} dt + e^{-rW} V_*(a', s') \right) + \right. \\
&\quad \left. 1_{[t_1, \infty)}(W) \left(\int_0^{t_1} 0e^{-rt} dt + \int_{t_1}^W 2e^{-rt} dt - e^{-rY_k+t_1} C + e^{-rW} V_*(a'', s'') \right) \right], \\
V_*(a'', s') &= \max_{t_1 \in [0, \infty]} \mathbb{E} \left[1_{[0, t_1)}(W) \left(\int_0^W 0e^{-rt} dt + e^{-rW} V_*(a'', s'') \right) + \right. \\
&\quad \left. 1_{[t_1, \infty)}(W) \left(\int_0^{t_1} 0e^{-rt} dt + \int_{t_1}^W 1e^{-rt} dt - e^{-rY_k+t_1} C + e^{-rW} V_*(a', s') \right) \right], \\
V_*(a'', s'') &= \frac{2}{r}(1 - \kappa) + \kappa V_*(a'', s'),
\end{aligned}$$

which reduces to a system of two equations in two unknowns.

The value function equations involve two optimization problems, the one at (a', s'') and the one at (a'', s') . Monotone comparative statics show that the optimal $t_1^*(a', s'')$ at (a', s'') is smaller than the solution $t_1^*(a'', s')$ (because the flow payoffs of the switch are 2 rather than 1). Let us suppose that the solution at (a', s'') is strictly positive and less than ∞ and examine the determinants of the corresponding $t_1^*(a', s'')$. The tradeoff is between the gain in flow utility and the loss if C is incurred and the state changes back to s' in a short time, and the first order conditions should tell us that the marginal gain of switching at t_1^* is equal to the expected marginal opportunity cost.

From Theorem 4 (below), the first order conditions for $0 < t_1^*(a', s'') < \infty$ are

$$(7) \quad [u(a'', s'') - u(a', s'')] - rC = h_W(t_1^*) \mathbb{E} [C + (V_*(a', s') - V_*(a'', s''))].$$

This condition must capture indifference between switching and not switching at t_1^* . The LHS times dt is the next instant's net flow benefit from switching: the term $[u(a'', s'') - u(a', s'')] dt$ gives the change in flow benefit; and rC is the perpetual annuity flow value of C . To analyze the RHS times dt : $h_W(t_1^*) dt$ gives the probability that the state switches from s'' back to s' in the next instant; if this happens, then the decision maker has saved C plus the value difference $V_*(a', s') - V_*(a'', s')$. In this problem, it is necessary that the LHS be positive in order to ever justify switching to a'' at (a', s'') .

3. THE MODEL

We start with a brief description of the basic relations between incomplete waiting times and their hazard rates. We then turn to the class of stochastic processes describing the utility relevant parts of the changing environment in which the decision maker is immersed. We then describe the class of decision problems under consideration, reactive semi-Markovian decision problems. Following this we give an alternative interpretation of the decision maker knowing the hazard rate. The basic existence and characterization results for an optimal policy are in the subsequent section.

3.1. Hazard Rates of Incomplete Waiting Times. A random variable, $W \geq 0$, is **incomplete** if it has a mass point at ∞ . For a possibly incomplete W with density on $[0, \infty)$, the following summarizes the relation between the **density**, $f_W(t)$, the **cumulative distribution function (cdf)**, $F_W(t)$, the **reverse cdf**, $G_W(t)$, the **hazard rate**, $h_W(t)$, the **cumulative hazard**, $H_W(t)$, and the **mass at infinity**, q_W , for $t \geq 0$:

$$(8) \quad F_W(t) = \int_0^t f_W(x) dx; \quad G_W(t) = 1 - F_W(t); \quad h_W(t) = \frac{f_W(t)}{G_W(t);}$$

$$H_W(t) = \int_0^t h_W(x) dx; \quad G_W(t) = e^{-H_W(t)}; \quad \text{and } q_W = e^{-H_W(\infty)}.$$

If $H_W(t) = \int_0^t h(x) dx \uparrow \infty$ as $t \uparrow \infty$, then $q_W = 0$ so that $W < \infty$ with probability 1, so that $F_W(t) \uparrow 1$ and $G_W(t) \downarrow 0$ as $t \uparrow \infty$.

From $G_W(t) = e^{-H_W(t)}$ one sees that any non-negative h can serve as the hazard rate for some waiting time, W , and W is incomplete iff h is integrable. The following are well-known examples.

- (1) An incomplete negative exponential has cdf $(1 - q_W)(1 - e^{-\lambda t})$ and everywhere decreasing hazard rate $\lambda \left[(q_W / (1 - q_W)) e^{\lambda t} + 1 \right]^{-1}$. If $q_W = 0$, then the hazard rate is constant and the waiting time is memoryless.
- (2) A Weibull distribution is of the form $W = X^\gamma$, $\gamma > 0$, where X is a negative exponential(λ). The cdf of an incomplete Weibull is $(1 - q_W)(1 - e^{-\lambda t^{1/\gamma}})$, and the hazard rate is $h_W(t) = \frac{\lambda}{\gamma} t^{\frac{1-\gamma}{\gamma}} \left[(q_W / (1 - q_W)) e^{\lambda t^{1/\gamma}} + 1 \right]^{-1}$. If $\gamma > 1$, the $h_W(0+) = \infty$ and the hazard rate strictly decreases to 0. If $\gamma < 1$, then $h_W(0) = 0$ and the hazard rate is first increasing then decreasing if $q_W > 0$, otherwise it is strictly increasing.

3.2. Semi-Markov Processes. All random variables are defined on a probability space, (Ω, \mathcal{F}, P) . A **semi-Markov process** is a piecewise constant, right-continuous stochastic process, $X : \Omega \times [0, \infty) \rightarrow S$ where S is a state space.

The crucial part of specifying X is the Markov process $(S_k)_{k=0}^\infty$, and the associated distribution of waiting times until the next transition, $(W_k)_{k=0}^\infty$. We assume that there is a continuous stochastic kernel $s \mapsto p(s) \in \Delta(S)$, where $\Delta(S)$ is the set of Borel probabilities on S .⁸ For the reactive problems that we consider below, the Markov process $(S_k)_{k=0}^\infty$ takes values in S with transition probabilities $P(S_{k+1} \in E | S_k = s) = P(S_{k+1} \in E | S_k = s, S_{j < k}) = p(s)(E)$.

For each $s \in S$, there is a distribution Q_s on $\mathbb{R}_{++} \cup \{\infty\}$ determining the distribution of waiting times for changes, $(W_k)_{k=1}^\infty$.⁹ The building blocks for X are a countable collection, $\Xi = (S_k, W_k)_{k=0}^\infty$ where: the $W_0 \equiv 0$; $(W_k)_{k \geq 1}$ is a sequence of strictly positive random times with distributions depending only on S_k , i.e. $P(W_k \in B | S_k) = Q_{S_k}(B) = P(W_k \in B | \Xi)$.

A crucial construct for both defining semi-Markov processes and analyzing the associated optimization problems is the notion of a **stochastic interval**.

Definition 1. Define $Y_0 = W_0$ and $Y_{k+1} = Y_k + W_{k+1}$. The **stochastic interval** between Y_k and Y_{k+1} is a subset of $\Omega \times [0, \infty)$ defined as $\llbracket Y_k, Y_{k+1} \llbracket = \{(\omega, t) : Y_k(\omega) \leq t < Y_{k+1}(\omega)\}$.

⁸At the most general, the state space and the action space are assumed to be Polish, i.e. Borel subsets of complete separable metric spaces. $\Delta(S)$ has the weak* topology from S .

⁹In future research, we intend have both p and Q determined both by s and by a .

We assume $P(W_k > 0) = 1$ for all $k \in \mathbb{N}$ and that $\sup_k \mathbb{E} e^{-rW_k} < 1$. This implies that $P(Y_k \rightarrow \infty) = 1$ so that $\Omega \times [0, \infty) = \cup_{k=0}^{\infty} \llbracket Y_k, Y_{k+1} \rrbracket$.

Definition 2. The **semi-Markov process (smp)** based on Ξ is the function $X(\omega, t) = \sum_k S_k(\omega) 1_{\llbracket Y_k, Y_{k+1} \rrbracket}(\omega, t)$.

X_t denotes the random variable $\omega \mapsto X(\omega, t)$. By construction, $X_{Y_k} = S_k$.

Example 1. If $\Xi = (S_k, W_k)_{k=0}^{\infty}$ the $S_k \equiv k$ and the W_k are i.i.d. negative exponentials with parameter λ , then the associated smp X is a standard Poisson process with arrival rate λ . More generally, if $(S_k)_{k=0}^{\infty}$ is any Markov chain taking values in S , then $(S_k)_{k=0}^{\infty}$ is **embedded** in X and the transition times follow a Poisson process. For more general distributions of the W_k , the standard queueing models give rise to special cases of smp's.

Example 2. In the optimal search time problem in §2.1, take $S = \{s_0, s_1\}$ where s_0 is the initial state and s_1 is the state “source of higher profits has been found.” The transition kernel for the states is given by $P = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$, i.e. one transitions from s_0 to s_1 but never the reverse. Let $A = \{a_0, \underline{a}, \bar{a}\}$ where a_0 is the action “continue searching” and \underline{a} is the action “use the alternative (older) technology.” and \bar{a} is the action “use the newly found technology.” Q_{s_0} is the distribution of W , and $Q_{s_1}(\infty) = 1$, which corresponds to s_1 being an absorbing state. We arrange for \bar{a} not to be chosen in s_0 by making the flow utility sufficiently negative.

When actions a_k do not affect the distributions of the S_{k+1} nor the distribution of the W_k , the problem of picking the optimal a_k is inherently **reactive**.

3.3. Uncertainty About the Hazard Rate. We derive our formal results in a mathematical setting where the hazard rate functions are known to the decision maker. For expected utility maximizers, this is no loss of generality as any arbitrary prior distribution over hazard rates can be mimicked by a single hazard rate function.

Let \mathcal{P} denote the set of distributions on $[0, \infty]$. The set of extreme points is the set of point masses $\{\delta_x : x \in [0, \infty]\}$. For any $p \in \mathcal{P}$ which is not an extreme point, there are uncountably many probability distributions, μ , on \mathcal{P} such that $p = \int_{\mathcal{P}} q d\mu(q)$. This implies that any hazard rate can arise from updating a distribution over distributions. When we use subsets of \mathcal{P} , there are approximate

versions of this. We give one such result, there are uncountably many similar ones.

Let $g_W(\cdot|\lambda, \gamma)$ denote the density function of a Weibull(λ, γ).

Theorem 1. *For every strictly positive C^1 density, f , on $(0, \infty)$, for every $\epsilon > 0$, and for every compact $K \subset (0, \infty)$, there exists a density g of the form $g_\mu(x) = \int g_W(x|\lambda, \gamma) d\mu(\lambda, \gamma)$, μ a probability distribution on \mathbb{R}_{++}^2 , such that $\max_{x \in K} |f(x) - g_\mu(x)| < \epsilon$ and $\max_{x \in K} |h_f(x) - h_{g_\mu}(x)| < \epsilon$ where h_f and h_{g_μ} are the hazard rates associated with the densities f and g_μ .*

Proof. If the class of Weibulls included the point masses, we could simply draw them according to the distribution f . The first part of the proof shows that the class of Weibulls, which is C^∞ , contains distributions arbitrarily close to the point masses.

Letting $r = 1/\gamma$, $g_W(x|\lambda, 1/r) = \frac{1}{\lambda^r} r x^{r-1} e^{-(x/\lambda)^r}$. For $r > 0$, $g_W(\cdot|\lambda, 1/r)$ is single-peaked, and achieves its maximum at $x^* = \lambda \left(\frac{r-1}{r}\right)^{1/r}$, at which point it takes the value $g_W(x^*|\lambda, 1/r) = r \frac{1}{\lambda} \left(\frac{r-1}{r}\right)^{(r-1)/r} e^{(r-1)/r}$. Further, as $r \uparrow \infty$, $x^* \uparrow \lambda$, $g_W(x^*|\lambda, 1/r) \uparrow \infty$, and for any positive $c \neq 1$, $g_W(cx^*|\lambda, 1/r) \downarrow 0$.

For $r > 1$, let μ_r be the distribution concentrated on r and having λ distributed according to the target density f . As $r \uparrow \infty$, the C^∞ density of g_μ converges uniformly to the density of f . Because f is bounded away from 0 on K , this directly implies the hazard rates are also uniformly close for r sufficiently large. \square

4. REACTIVE SEMI-MARKOVIAN DECISION PROBLEMS

We study the problem of the optimal timing and choice of actions for reactive smp's when flow utility, $u(a, s)$, depends on the action a in an action space A , and the state, s in a state space S . A and S are metric spaces, A is compact, $u(\cdot, \cdot)$ is jointly continuous and bounded.¹⁰

¹⁰From an abstract point of view, the continuity assumption is without loss — if each $u(\cdot, s)$ is continuous and each $u(a, \cdot)$ is measurable, then u is jointly measurable and one can always find a metric on S making u jointly continuous.

Total reward is the expected discounted integral of the flows, and changes of action are costly. An action plan, $\mathbf{a}(\omega, t)$, will depend on t and on ω , but will only depend on ω through the realization of the process X at or before t .

4.1. Filtrations and Action Plans. Define a filtration $\{\mathcal{F}_t : t \in [0, \infty)\}$ by defining $\mathcal{G}_s \subset \mathcal{F}$ to be the smallest σ -field making $\omega \mapsto X(\omega, s)$ measurable and $\mathcal{F}_t = \sigma\{\mathcal{G}_s : s \leq t\}$. Let \mathcal{B}_t denote the Borel σ -field on the interval $[0, t]$.

Definition 3. A **policy** is a mapping $a : \Omega \times [0, \infty) \rightarrow A$ such that

- a. for all t , $a : \Omega \times [0, t] \rightarrow A$ is $\mathcal{F}_t \otimes \mathcal{B}_t$ -measurable, and
- b. for all ω , $t \mapsto a(\omega, t)$ is cadlag, piecewise constant, and has at most finitely many discontinuities on any bounded interval.

Associated with any policy $a(\cdot, \cdot)$ are the **change times**, $T(\omega)$, defined by $T = \{t_{k+1}(\omega) = \min\{t \geq t_k : a(t, \omega) \neq a(t_k, \omega)\}$ where $t_0 \equiv 0$.

4.2. Stochastic Interval Problems. The decision problem is to pick the policy that maximizes $E\left(\int_0^\infty u(\mathbf{a}, X)e^{-rs} ds - C \sum_{t \in T} e^{-rt}\right)$ given that one starts in state s_0 with action a_0 at $t_0 := 0$. We will show that the bounded continuous value function is the unique fixed point of a contraction operator. As a step in that direction and to begin to characterize the associated value function, we treat the decision problem within each stochastic interval, $\llbracket Y_k, Y_{k+1} \rrbracket$.

We suppose that the interval $\llbracket Y_k, Y_{k+1} \rrbracket$ starts in a state s_k , i.e. $X_{Y_k} = s_k$, with the action being a_k . Let $\tau = (t_n)_{n \in \mathbb{N}} \in [0, \infty]^\mathbb{N}$ be a non-decreasing sequence of times, and let $\alpha \in A^\mathbb{N}$ be a sequence of actions. The interpretation is that τ is the set of points in the interval at which the decision maker switches and α is the associated set of choices. Specifically;

- (1) if $Y_k + t_1 < Y_{k+1}$, then the decision maker incurs the cost $Ce^{-r(Y_k+t_1)}$ of switching from a_0 to a_1 , and during the interval $[Y_k, Y_k + t_1)$ the decision maker enjoys the flow utility $u(a_0, s_0)$;
- (2) if $Y_{k+1} < Y_k + t_1$, then during the interval $[Y_k, Y_{k+1})$, the decision maker enjoys the flow utility $u(a_0, s_0)$ and at Y_{k+1} , they receive the random utility $V_*(a_0, S_{k+1})$; more generally,
- (3) if $Y_k + t_{n+1} < Y_{k+1}$, then the decision maker incurs the cost $Ce^{-r(Y_k+t_{n+1})}$ of switching from a_n to a_{n+1} , and during the interval $[Y_k + t_n, Y_k + t_{n+1})$, the decision maker enjoys the flow utility $u(a_n, s_0)$; and

- (4) if $Y_k + t_n < Y_{k+1} < Y_k + t_{n+1}$, then during the interval $[Y_k + t_n, Y_{k+1})$, the decision maker enjoys the flow utility $u(a_n, s_0)$, and at Y_{k+1} , they receive the random utility $V_*(a_n, S_{k+1})$.

To discuss the problem of finding the optimal times to move and the optimal actions to choose in the stochastic interval $\llbracket Y_k, Y_{k+1} \llbracket$ with (something like) minimal notation, we condition on $Y_k = t$, subtract t , renormalize the k to 0, and start the discounting from 0. To this end, define

$$(9) \quad M_0(\boldsymbol{\alpha}, \boldsymbol{\tau}; V_\circ) = 1_{[0, t_1)}(Y_1) \left[\int_0^{Y_1} u(a_0, s_0) e^{-rs} ds + e^{-rY_1} V_\circ(a_0, S_1) \right],$$

$$(10) \quad M_n(\boldsymbol{\alpha}, \boldsymbol{\tau}; V_\circ) = 1_{[t_n, t_{n+1})}(Y_1) \left[\sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} u(a_k, s_0) e^{-rs} ds + \right.$$

$$(11) \quad \left. \int_{t_n}^{Y_1} u(a_n, s_0) e^{-rs} ds + e^{-rY_1} V_\circ(a_n, S_{k+1}) \right],$$

$$(12) \quad C_n(\boldsymbol{\alpha}, \boldsymbol{\tau}; V_\circ) = - \sum_{i=1}^n C e^{-rt_i} \cdot 1_{[t_n, t_{n+1})}(Y_1).$$

Definition 4. For a bounded continuous $(a, s) \mapsto V_\circ(a, s)$ starting action a_0 and initial state s_0 , the associated **within-interval problem (winp)** is

$$(13) \quad \max_{\boldsymbol{\alpha}, \boldsymbol{\tau}} E \left(M_0(\boldsymbol{\alpha}, \boldsymbol{\tau}; V_\circ) + \sum_{n=1}^{\infty} [M_n(\boldsymbol{\alpha}, \boldsymbol{\tau}; V_\circ) + C_n(\boldsymbol{\alpha}, \boldsymbol{\tau}; V_\circ)] \Big|_{s_0, a_0} \right).$$

4.3. The Blackwell-Pontryagin Equation. A series of Lemmas leads to the basic existence and characterization results.

Lemma 1. *Solving each winp sequentially yields an action plan.*

Proof. Within each $\llbracket Y_k, Y_{k+1} \llbracket$, the optimal $(\boldsymbol{\alpha}^*, \boldsymbol{\tau}^*)$ only depends on S_k . □

Lemma 2. *Each winp has a solution.*

Proof. Because Q_s has no atoms except perhaps at ∞ , the sum in (13) is continuous on the compact subset of $A^{\mathbb{N}} \times [0, \infty]^{\mathbb{N}}$ for which $t_n \leq t_{n+1}$ for all $n \in \mathbb{N}$. □

Lemma 3. *If $(\boldsymbol{\tau}^*, \boldsymbol{\alpha}^*)$ solves a winp, then for all $n \in \mathbb{N}$, $[t_n^* < \infty] \Rightarrow [t_n^* < t_{n+1}^*]$, and $t_n^* \uparrow \infty$.*

Proof. For the first part, note that moving twice at any $t < \infty$ incurs the fixed cost Ce^{-rt} twice with no advantage in flow payoffs. For the second part, if $\sup_n t_n^* < \infty$, then the cost incurred is infinite, and never moving is strictly better than this policy. \square

Theorem 2. *If $\beta = \inf\{E e^{-r\tau} : \tau \sim \mathcal{L}(s), s \in S\} < 1$, then the mapping from V_o to $T(V_o)$ defined by*

$$(14) \quad T(V_o)(a_0, s_0) = \max_{\alpha, \tau} E \left(M_0(\alpha, \tau; V_o) + \sum_{n=1}^{\infty} [M_n(\alpha, \tau; V_o) + C_n(\alpha, \tau; V_o)] \middle| s_0, a_0 \right)$$

has contraction factor β , and maps bounded continuous functions to bounded continuous functions. Further, the value to following the solutions to each winp is the unique fixed point of T .

Proof. From Blackwell's Lemma for contraction mappings (e.g. Corbae et. al., Lemma 6.2.33, p. 282), it is sufficient to show that T is monotonic and that for any constant c , $T(V_o + c) \leq T(V_o) + \beta c$. Monotonicity is immediate. For the contraction, note that for all policies (α, τ) with $t_m \uparrow \infty$, the difference between

$$(15) \quad E \left(M_0(\alpha, \tau; V_o + c) + \sum_{n=1}^{\infty} [M_n(\alpha, \tau; V_o + c) + C_n(\alpha, \tau; V_o + c)] \right)$$

and

$$(16) \quad E \left(M_0(\alpha, \tau; V_o) + \sum_{n=1}^{\infty} [M_n(\alpha, \tau; V_o) + C_n(\alpha, \tau; V_o)] \right)$$

is

$$(17) \quad E \sum_{m=0}^{\infty} 1_{[t_m, t_{m+1})}(Y_1) e^{-rY_1} c,$$

and because $t_m^* \uparrow \infty$, this difference is identically $e^{-rY_1} c$. Since the difference is independent of the choice of optimal policy, the solution (set) to the problem of maximizing the equation in (15) is exactly the same as the solution (set) for maximizing (16). Finally, since $e^{-rY_1} \leq \beta$, $T(V_o + c) \leq T(V_o) + \beta c$.

Finally, following the solution to the first winp associated with V_* yields value V_* . Inductively, following the solution to the first n winps also has value V_* . Since $Y_k \uparrow \infty$, the value to following the solution to each winp is V_* . \square

5. EULER EQUATIONS

We give the Euler equations in reactive smp problems for the optimal policy within the class of policies where one moves at most once within a stochastic interval¹¹. Recall that for a reactive problem, we consider semi Markov processes whose stochastic kernels do not depend on the choice of action. Also, note that we now have two choice variables, namely the choice of action a and the optimal time to execute that action t , hence we now have two Euler equations, slightly different from the usual dynamic programming problem.

5.1. Memoryless Processes.

Theorem 3. *If $(W_k)_{k=1}^\infty$ are i.i.d $\text{Exp}(\lambda)$, then*

- *the optimal waiting time, if unique, is either 0 or ∞*
- *if the waiting time is non-unique, then it must be the whole interval $[0, \infty]$*
- *the optimal action only depends on the current state and is independent of the timing of action.*

Proof. Let us re-initialize time within each interval by setting $Y_k = 0$. With the condition that at most one move is feasible given C , we have the value function

$$(18) \quad V(a_k, s_k) = \max_{t \in [0, \infty]} \max_{a \in A} \mathbb{E} \left\{ \left[\int_0^{y_{k+1}} u(a_k, s_k) e^{-ru} du + V(a_k, s_{k+1}) e^{-ry_{k+1}} \right] \mathbf{1}_{\{y_{k+1} < t\}} \right. \\ \left. + \left[\int_0^t u(a_k, s_k) e^{-ru} du + \int_t^{y_{k+1}} u(a, s_k) e^{-ru} du + V(a, s_{k+1}) e^{-ry_{k+1}} - C \right] \mathbf{1}_{\{y_{k+1} > t\}} \middle| s_k \right\}$$

First we establish that the optimal action does not depend on the timing of action, which is readily seen from the F.O.C. of the maximand in (18) with respect to a :

$$\frac{\partial}{\partial a} u(a, s_k) = -\frac{\lambda}{r} \frac{\partial}{\partial a} V(a, s_{k=1})$$

¹¹It pertains to situations where the cost is high enough so that the number of feasible moves within a stochastic interval is at most one, or the structure of the decision problem allows only a single action each time. A broad class of institutional decision making problems would fit this description.

If $t = 0$ is chosen, then the Value Function evaluated at $t = 0$ is

$$V_0 = \max_{a \in A} \left[\frac{1}{\lambda + r} u(a, s_k) + \frac{\lambda}{\lambda + r} \mathbb{E}V(a, s_{k+1}) - C \right]$$

If $t = \infty$ is chosen, then the Value Function evaluated at $t = \infty$ is

$$V_\infty = \frac{1}{\lambda + r} u(a_k, s_k) + \frac{\lambda}{\lambda + r} \mathbb{E}V(a_k, s_{k+1})$$

Finally, the derivative of the maximand in (18) with respect to t is

$$\int_0^t (V_\infty - V_0) e^{-(\lambda+r)u} du$$

where

$$\begin{aligned} V_\infty - V_0 &= \frac{1}{\lambda + r} [u(a_k, s_k) - u(a^*, s_k)] + \frac{\lambda}{\lambda + r} \mathbb{E}[V(a_k, s_{k+1}) - V(a^*, s_{k+1})] - C \\ &= \frac{1}{\lambda + r} \{ (u(a_k, s_k) - u(a^*, s_k) - rc) + \lambda \mathbb{E}[V(a_k, s_{k+1}) - V(a^*, s_{k+1}) - C] \}. \end{aligned}$$

The sign of the derivative depends only on the sign of $V_\infty - V_0$, which is constant on $(0, \infty)$. If $V_\infty - V_0 > 0$, the value is increasing in time and $t^* = \infty$. If $V_\infty - V_0 < 0$, then the value is decreasing in time and $t^* = 0$. If $V_\infty - V_0 = 0$, then any time in $[0, \infty]$ is optimal. \square

Hence, with memoryless distributions of arrival times, the optimal action depends only on the current state, and the optimal timing is either to change immediately after observing the state change, or wait until the next change.

If $V_\infty - V_0 = 0$, then

$$(19) \quad u(a_k, s_k) - u(a^*, s_k) + rC = \lambda \mathbb{E}[V(a^*, s_{k+1}) - V(a_k, s_{k+1})] - \lambda C$$

which is a special case of more general Euler equations derived in Theorem 4 that we turn to next.

5.2. Waiting Time Distributions with Memory.

Theorem 4. *If $(W_k)_{k=1}^\infty$ have distributions with memory, i.e. if $P(W > t \mid W \geq t)$ depends non-trivially on t , then there may exist interior optimal time of action. In particular, for the class of policies with at most one action within an interval,*

- any optimal interior time for action (t^*), and the optimal action a^* are characterized by the following system of Euler equations:

$$\begin{aligned}
(20) \quad & u(a^*, s_k) - u(a_k, s_k) - rC = -h_W(t^*)\mathbb{E}[V(a^*, s_{k+1}) - V(a_k, s_{k+1}) - C] \\
& \int_{t^*}^{\infty} e^{-ry} \frac{d}{da} u(a, s_k)|_{a^*} (1 - F_W(s, y)) dy \\
& + \int_{t^*}^{\infty} e^{-ry} f_W(s, y) \frac{d}{da} \mathbb{E}V(a, S_k)|_{a^*} \\
& = 0
\end{aligned}$$

- the partial second order condition for interior optimal time t^* for both increasing and decreasing hazard rates are as follows,

$$\begin{aligned}
& \frac{u(a^*, s_k) - u(a_k, s_k) - rC}{\mathbb{E}[V(a^*, s_{k+1}) - V(a_k, s_{k+1}) - C]} < \frac{f'_W(t^*)}{f_W(t^*)}, \\
& \text{and in addition for decreasing hazard rates} \\
& \left| \frac{u(a^*, s_k) - u(a_k, s_k) - rC}{\mathbb{E}[V(a^*, s_{k+1}) - V(a_k, s_{k+1}) - C]} \right| > \left| \frac{f'_W(t^*)}{f_W(t^*)} \right|
\end{aligned}$$

Proof. For the derivation of the Euler equations, see Appendix. The first Euler equation makes it clear that optimal time of action depends on the hazard function.

The rest follows from the requirements of the second order conditions. Note first that (4) implies $u(a^*, s_k) - u(a_k, s_k) - rC$ and $\mathbb{E}[V(a^*, s_{k+1}) - V(a_k, s_{k+1}) - C]$ must have opposing signs at the optima, since for any t^* , $h_W(t^*) \geq 0$. Now, for an interior optimum for t^* , the second order condition (18) w.r.t. t at the optima is¹²:

$$(21) \quad f_W(t^*)[u(a^*, s_k) - u(a_k, s_k) - rC] - f'_W(t^*)\mathbb{E}[V(a^*, s_{k+1}) - V(a_k, s_{k+1}) - C] < 0$$

which implies

$$(22) \quad \frac{u(a^*, s_k) - u(a_k, s_k) - rC}{\mathbb{E}[V(a^*, s_{k+1}) - V(a_k, s_{k+1}) - C]} < \frac{f'_W(t^*)}{f_W(t^*)}$$

It is easy to show that $h'_W(t) > 0 \Rightarrow f'_W(t) > 0$. Hence, given that the ratio on the L.H.S. must be negative if the (4) is satisfied, the S.O.C. is always satisfied

¹²easy calculation of the second derivative and then using the F.O.C

for an increasing hazard rate whenever the F.O.C is satisfied. For a decreasing hazard rate, the S.O.C. will be satisfied for small enough change in hazard rate, i.e. when

$$(23) \quad \left| \frac{u(a^*, s_k) - u(a_k, s_k) - rC}{\mathbb{E}[V(a^*, s_{k+1}) - V(a_k, s_{k+1}) - C]} \right| > \left| \frac{f'_W(t^*)}{f_W(t^*)} \right|.$$

□

5.3. Interpretation of the Euler Equations. The Euler equations have simple interpretations in terms of marginal benefits and marginal costs of change. First note that an equivalent representation of (4) is

$$(24) \quad \begin{aligned} & (1 - F_W(t^*)) [u(a^*, s_k) - u(a_k, s_k) - rC] \\ & = f_W(t^*) \mathbb{E}[V(a_k, s_{k+1}) - V(a^*, s_{k+1}) + C] \end{aligned}$$

The L.H.S represents the net expected instantaneous flow cost/benefit of moving to action a^* , given that the state change has not happened yet. We would lose/gain $u(a^*, s_k) - u(a_k, s_k)$ in terms of flow utility and rC in terms of flow cost of action. The R.H.S shows expected instantaneous benefit of waiting if the state changes in the next instant. It is the expected net benefit of starting the next stochastic interval with the action a_k instead of a^* , net of cost saved, multiplied by the conditional probability that the next change of state would occur in the next instant, given that it has not occurred till now.

The second Euler equation has straightforward interpretation as an expression of the principle that the expected marginal increase in value with a change in action must equal zero.

The important thing to note is that for (4) to be satisfied at an interior point, $u(a^*, s_k) - u(a_k, s_k) - rC$ and $\mathbb{E}[V(a^*, s_{k+1}) - V(a_k, s_{k+1}) - C]$ must have opposite signs, given optimal action a^* . This is intuitive, because if the optimal action is such that both these terms are positive, then it means we gain utility by moving to a^* regardless of whether the state remains the same or changes, so one should move to a^* immediately. On the other hand, if both these terms are negative, then even the optimal change is not worth the cost, given the current state and the expected future state. Hence one must take no action, i.e. wait till the next change.

The next thing to note are the conditions there might exist an interior optimal time under both increasing and decreasing hazard rates, but under opposing

circumstances. Observe that for a fixed t^* , we have a fixed a^* , and $u(a^*, s_k) - u(a_k, s_k) - rC$ is constant over time.

The following corollary characterizes the conditions under which interior optima exist for decreasing hazard rates.

Corollary 4.1. *For any interior optimal time t^* to exist in case of decreasing hazard rates, the following must be true:*

$$u(a^*, s_k) - u(a_k, s_k) - rC > 0$$

$$\mathbb{E}[V(a^*, s_{k+1}) - V(a_k, s_{k+1}) - C] < 0$$

and

$$u(a^*, s_k) - u(a_k, s_k) - rC < -h_W(t)\mathbb{E}[V(a^*, s_{k+1}) - V(a_k, s_{k+1}) - C] \text{ for } t < t^*$$

For any interior optimal time t^* to exist in case of increasing hazard rates, the following must be true:

$$(25) \quad u(a^*, s_k) - u(a_k, s_k) - rC < 0$$

$$(26) \quad \mathbb{E}[V(a^*, s_{k+1}) - V(a_k, s_{k+1}) - C] > 0$$

$$(27) \quad u(a^*, s_k) - u(a_k, s_k) - rC > -h_W(t)\mathbb{E}[V(a^*, s_{k+1}) - V(a_k, s_{k+1}) - C] \text{ for } t < t^*$$

Proof. Consider the various possible cases.

Case 1

Let,

$$(28) \quad u(a^*, s_k) - u(a_k, s_k) - rC > 0$$

and

$$(29) \quad \mathbb{E}[V(a^*, s_{k+1}) - V(a_k, s_{k+1}) - C] < 0$$

This is the scenario where a new optimal action provides net benefits in the current state, but it would be better to start the next stochastic interval with the status quo action. Then

$$(30) \quad -h_W(t^*)\mathbb{E}[V(a^*, s_{k+1}) - V(a_k, s_{k+1}) - C] > 0$$

Clearly, there could not be an interior solution t^* with $h_W(t)$ decreasing in t if

$$(31) \quad u(a^*, s_k) - u(a_k, s_k) - rC > -h_W(t)\mathbb{E}[V(a^*, s_{k+1}) - V(a_k, s_{k+1}) - C] \text{ for } t < t^*$$

because then one would just change immediately. So, we would only have interior optima with decreasing hazard rate in *Case 1* if $-h_W(t)\mathbb{E}[V(a^*, s_{k+1}) - V(a_k, s_{k+1}) - C]$ cuts $u(a^*, s_k) - u(a_k, s_k) - rC$ from above, as a function of t .

On the other hand, we cannot have an interior optimal t^* in *Case 1* with $h_W(t)$ increasing in t if

$$(32) \quad u(a^*, s_k) - u(a_k, s_k) - rC < -h_W(t)\mathbb{E}[V(a^*, s_{k+1}) - V(a_k, s_{k+1}) - C] \text{ for } t < t^*$$

because then the benefit of waiting is higher than the cost at $t = 0$ and keeps increasing over time, and one would keep waiting forever. We also could not have an interior optimal time in this case with $h_W(t)$ increasing in t even if

$$(33) \quad u(a^*, s_k) - u(a_k, s_k) - rC > -h_W(t)\mathbb{E}[V(a^*, s_{k+1}) - V(a_k, s_{k+1}) - C] \text{ for } t < t^*$$

because although the benefit of waiting is less than the cost at $t = 0$, the more we wait, the marginal benefit of waiting one instant longer keeps going up. Hence optimal t^* is either 0 or ∞ .

So there is no interior optimal time with increasing hazard rate in *Case 1*

Case 2

Let,

$$(34) \quad u(a^*, s_k) - u(a_k, s_k) - rC < 0$$

and

$$(35) \quad \mathbb{E}[V(a^*, s_{k+1}) - V(a_k, s_{k+1}) - C] > 0$$

This is the scenario when a new optimal action is worse than the status quo in the current state, but it would be better to start the next stochastic interval with the new optimal action. Then

$$(36) \quad -h_W(t^*)\mathbb{E}[V(a^*, s_{k+1}) - V(a_k, s_{k+1}) - C] < 0$$

We cannot have an interior solution t^* with $h_W(t)$ decreasing in t if

$$(37) \quad u(a^*, s_k) - u(a_k, s_k) - rC > -h_W(t)\mathbb{E}[V(a^*, s_{k+1}) - V(a_k, s_{k+1}) - C] \text{ for } t < t^*$$

because the R.H.S, which now denotes the cost of waiting, keeps falling over time. And of course we cannot have an interior solution with

$$(38) \quad u(a^*, s_k) - u(a_k, s_k) - rC < -h_W(t)\mathbb{E}[V(a^*, s_{k+1}) - V(a_k, s_{k+1}) - C] \text{ for } t < t^*$$

because although the benefit of waiting is less than the cost at $t = 0$, the more we wait, the marginal cost of waiting one instant longer keeps going down. Hence optimal t^* is either 0 or ∞ .

With increasing hazard rate, we have no interior optimal time t^* if

$$(39) \quad u(a^*, s_k) - u(a_k, s_k) - rC < -h_W(t)\mathbb{E}[V(a^*, s_{k+1}) - V(a_k, s_{k+1}) - C] \text{ for } t < t^*$$

one would change action immediately, since the cost of waiting is forever increasing. But if,

$$(40) \quad u(a^*, s_k) - u(a_k, s_k) - rC > -h_W(t)\mathbb{E}[V(a^*, s_{k+1}) - V(a_k, s_{k+1}) - C] \text{ for } t < t^*$$

then, although waiting gives net benefit at the beginning, the cost of waiting keeps going up, and at t^* it becomes optimal to take the new optimal action. \square

6. EXTENSIONS

We have presented results for reactive problems where the stochastic kernel of the smp does not depend on the choice of action. There are real life situations which are better modeled as controlled smp's where the choice of action effects the stochastic kernel. That would be the next step in our analysis. Here is an example where controlled stochastic processes would be relevant.

6.1. Precautionary Principle and Environmental Change. Economic models of climate change have traditionally treated the process as one of gradual change to new, stable state. Recent research in climate science has found evidence of both very rapid changes over as short period of time (around a decade) and also periods of significant fluctuations or 'environmental flickering' over periods as short as a year (Hall and Behl [2006], Stern and Treasury [2007]). These phases of rapid change and/or flickering seem to be triggered once a threshold point is reached in the ecological system. Quick changes in climate are more expensive to adapt to, and if the state changes to one where the arrival times of subsequent changes have high arrival rates, the expense is further increased, pushing policy recommendations in the direction of those that arise from the precautionary principle.¹³

¹³While the 'Precautionary Principle,' that if an action or policy has a suspected risk of causing harm to the public or to the environment, in the absence of scientific consensus that the action or policy is harmful, the burden of proof that it is not harmful falls on those taking

While ‘gradual change’ models have usually prescribed ‘adaptation to climate change’ as opposed to ‘intervention to avert it’ [Nordhaus and Boyer, 2003], the decision problem takes a new shape when we incorporate the uncertainty over the expected arrival time of a possible catastrophic change and over the issue of whether or not we are moving towards such a critical threshold. In this context, the cost of controlling the intensity of the arrival rate process seems to be the major issue, and we intend to incorporate Stone’s (Stone [1973]) work on optimizing over control functions where intensity of arrival rate of changes to different states are the object of study.

7. CONCLUSION

We began with the question as to what the optimal time is for changing a status quo policy in response to an environmental change when policy change is costly and one anticipates another change in the environment in an unknown, random time in the future. There is an immediate tradeoff between optimizing with respect to the current state and optimizing with respect to the expected future state, given that actions are costly and different actions become optimal in different states. But the main interest in such a problem stems from the fact that passage of time since the last observed change might contain information about how soon the next change is likely to occur. This would be the case when the distribution of inter-arrival times for environmental changes have hazard rates that are not constant over time. In case of increasing or decreasing hazard rates, the likelihood that the next change would happen in the next instant, given that it has not happened until now, goes up or down respectively. Hence we have an additional tradeoff in terms of timing of the action once we have seen an environmental change, namely, we lose utility every instant that the current action is not optimized to the current state, but every instant of passing time gives us more information about how far in the future the next change is likely to occur. Intuitively, it would suggest that there might be a place for ‘informative waiting’, i.e. delaying one’s action in order to have more information about the time of the next environmental change. Our results show that for non-constant hazard rates, delaying your actions could be optimal under certain circumstances.

the action, has been in use (it forms, for e.g., the basis of the Kyoto Protocol), it has also been criticized for not having proper analytical basis [Sunstein, 2002].

REFERENCES

- A. R. Amar. *America's constitution: a biography*. Random House Trade, 2006. ISBN 0812972724.
- S. I. Boyarchenko and S. Levendorski. *Irreversible decisions under uncertainty: optimal stopping made easy*. Springer Verlag, 2007.
- E. G. Carmines and J. A. Stimson. *Issue evolution: Race and the transformation of American politics*. Princeton Univ Pr, 1990. ISBN 069102331X.
- P. K. Chintagunta. Inertia and variety seeking in a model of brand-purchase timing. *Marketing Science*, 17(3):253–270, 1998. ISSN 0732-2399.
- Z. Elkins, T. Ginsburg, J. Melton, and MyiLibrary. *The endurance of national constitutions*. Cambridge University Press, 2009. ISBN 0511640064.
- E. Flexner and E.F. Fitzpatrick. *Century of struggle: the woman's rights movement in the United States*. Belknap Press, 1996. ISBN 0674106539.
- F. F. Gonul and F. Ter Hofstede. How to compute optimal catalog mailing decisions. *Marketing Science*, 25(1):65–74, 2006.
- Darwin C. Hall and Richard J. Behl. Integrating economic analysis and the science of climate instability. *Ecological Economics*, 57(3):442 – 465, 2006. ISSN 0921-8009. doi: 10.1016/j.ecolecon.2005.05.001.
- R. Hardin. *Liberalism, constitutionalism, and democracy*. Oxford University Press, USA, 2003. ISBN 0199261687.
- J. Madison. Federalist no. 49. *The federalist*, pages 338–347, 1961.
- W. D. Nordhaus and J. Boyer. *Warming the world: economic models of global warming*. the MIT Press, 2003. ISBN 0262640546.
- E. Ostrom. *Governing the commons: The evolution of institutions for collective action*. Cambridge Univ Pr, 1990. ISBN 0521405998.
- J. M. Smith. The republic of letters: The correspondence between thomas jefferson and james madison 1776-1826, 3 vols. *New York: WW Norton*, 3:1364, 1995.
- N. H. Stern and G. B. Treasury. *The economics of climate change: the Stern review*. Cambridge Univ Pr, 2007. ISBN 0521700809.

- N. L. Stokey. *The economics of inaction: Stochastic control models with fixed costs*. Princeton Univ Pr, 2009. ISBN 0691135053.
- L.D. Stone. Necessary and sufficient conditions for optimal control of semi-Markov jump processes. *SIAM Journal on Control*, 11:187, 1973.
- C. R. Sunstein. The paralyzing principle. *Regulation*, 25(4):32–37, 2002.

8. APPENDIX

Derivation of the Euler Equations in *Theorem 4*

With the condition that at most one move is feasible given C , we have the value function

$$\begin{aligned} V(a_k, s_k) &= \max_{t \in [0, \infty]} \max_{a \in A} \mathbb{E} \left\{ \left[\int_0^{y_{k+1}} u(a_k, s_k) e^{-ru} du + V(a_k, s_{k+1}) e^{-ry_{k+1}} \right] \mathbf{1}_{\{y_{k+1} < t\}} \right. \\ &\quad \left. + \left[\int_0^t u(a_k, s_k) e^{-ru} du + \int_t^{y_{k+1}} u(a, s_k) e^{-ru} du + V(a, s_{k+1}) e^{-ry_{k+1}} - C \right] \mathbf{1}_{\{y_{k+1} > t\}} \middle| s_k \right\} \end{aligned}$$

Now we calculate each of the discounting terms;

$$\begin{aligned} \mathbb{E} \int_0^y e^{-ru} du \mathbf{1}_{\{y_{k+1} < t\}} &= \frac{1}{r} (1 - e^{-rt}) F_W(t) - \int_0^t F_W(y) e^{-ry} dy \\ \mathbb{E} e^{-ry} \mathbf{1}_{\{y_{k+1} < t\}} &= \int_0^t e^{-ry} dF_W(y) = e^{-rt} F_W(t) + r \int_0^t e^{-ry} F_W(y) dy \\ \mathbb{E} \int_0^t e^{-ru} du \mathbf{1}_{\{y_{k+1} > t\}} &= \frac{1}{r} (1 - e^{-rt}) (1 - F_W(t)) \\ \mathbb{E} \int_t^y e^{-ru} du \mathbf{1}_{\{y_{k+1} > t\}} &= \frac{1}{r} \left[e^{-rt} \int_t^\infty dF_W(y) - \int_t^\infty e^{-ry} dF_W(y) \right] \\ \mathbb{E} \int_t^\infty e^{-ry} dF_W(y) &= -e^{-rt} F_W(t) + r \int_t^\infty e^{-ry} F_W(y) dy \end{aligned}$$

Gathering terms, we have,

$$\begin{aligned} \frac{1}{r} (1 - e^{-rt}) F_W(t) - \int_0^t F_W(y) e^{-ry} dy + \frac{1}{r} (1 - e^{-rt}) (1 - F_W(t)) &= \frac{1}{r} (1 - e^{-rt}) - \int_0^t F_W(y) e^{-ry} dy \\ &= \int_0^t e^{-ry} dy - \int_0^t F_W(y) e^{-ry} dy \\ &= \int_0^t e^{-ry} (1 - F_W(y)) dy \end{aligned}$$

Also,

$$\begin{aligned} \frac{1}{r} e^{-rt} (1 - F_W(t)) - \frac{1}{r} \int_t^\infty e^{-ry} dF_W(y) &= \frac{1}{r} e^{-rt} (1 - F_W(t)) + \frac{1}{r} e^{-rt} F_W(t) - \int_t^\infty e^{-ry} F_W(y) dy \\ &= \frac{1}{r} e^{-rt} - \int_t^\infty e^{-ry} F_W(y) dy \end{aligned}$$

Moreover,

$$\begin{aligned}\mathbb{E}[-Ce^{-rt}\mathbf{1}_{\{y_{k+1}>t\}}] &= -Ce^{-rt}\int_t^\infty dF_W(y) \\ &= -Ce^{-rt}(1-F_W(t))\end{aligned}$$

Replacing the discount terms in the value function and gathering terms, we have

$$\begin{aligned}V(a_k, s_k) &= \left[\int_0^t e^{-ry}(1-F_W(t))dy \right] u(a_k, s_k) \\ &+ \left[e^{-rt}F_W(t) + r \int_0^t e^{-ry}F_W(y)dy \right] \mathbb{E}V(a_k, s_{k+1}) \\ &+ \left[\frac{1}{r}e^{-rt} - \int_t^\infty e^{-ry}F_W(y)dy \right] u(a, s_k) \\ &+ \left[-e^{-rt}F_W(t) + r \int_t^\infty e^{-ry}F_W(y)dy \right] \mathbb{E}V(a, s_{k+1}) \\ &- Ce^{-rt}(1-F_W(t))\end{aligned}$$

Taking the derivative of the above expression with respect to t and a and some rearrangement give us the Euler equations.

DEPARTMENT OF ECONOMICS, UNIVERSITY OF TEXAS, AUSTIN, TX 78712-1173 USA