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 to later changes, requiring further costly change. Waiting is informative if the hazard rate for the arrival of the next change is non-constant. We identify two broad classes of situations: in the first, delayed reaction is optimal only when the hazard rate of further changes is decreasing; in the second, it is optimal only when the hazard rate of further changes is increasing. The first class of situations correspond to having waited long enough to know that future changes in circumstances are comfortably in the future, and the associated non-optimality of action in the face of an increasing hazard rate corresponds to the counsel of patience in unsettled circumstances. The second class of situations corresponds to the delay of costly precautionary steps until the danger is clear enough. These results in non-stationary dynamic optimization provide a new set of motivations for building delay into legislative and other decision systems, and arise from extensions of semi-Markovian decision theory.

Key Words and Phrases. Hazard rates, non-stationary semi-Markovian decision theory, optimal timing in non-stationary contexts, optimally delayed reactions

---

Date: September, 2012

‡Many thanks to Martin Dumav and Zahid Arefin Chowdhury for their timely and much appreciated help, to seminar participants at UC Riverside and Delhi School of Economics and to conference participants at the IHS Summer Research Conference (2010), Midwest Political Science Association Conference (2010), Midwest Theory Conference (2011), GAMES (2012), and SSCW (2012) for many helpful and insightful 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 and responding to these changes entails significant costs. Given that 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, and that actions are costly, the question is whether to take an action in response to a change in the environment or to delay any change.

Variants of this problem have been extensively analyzed in economics (for example Boyarchenko and Levendorskiĭ [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

[1] In 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?" and the dynamically consistent answer depends both on the costliness of the action and the costliness of its reversal.

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

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

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]

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.

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

[4]Flexner 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.

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

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

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 "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. **Labor Search.** One of the issues relating to long-term unemployment is depletion of human capital, which might make a candidate less and less attractive to the potential employers as the duration of unemployment gets longer and longer. This factor would have important implications for standard labor search models, as the value of the future discounted wage and thus the reservation wage for an a get would be effected. Ortego-Marti [2011] uses this insight to explain observed wage dispersion in the labor market. Also, the possibility that a long period searching without finding a job is taken, by potential employers, as an indication that there is something wrong with the person, would mean a decreasing rate of arrival of a job of any given quality. This is independent of decay of human capital, and pushes the acceptance thresh-hold downwards. One of the key ingredient of the standard DMP model is a stochastic description of labor turnover, along with a model of labor-market tightness, and a bargaining model of wage determination (Diamond [2011]; Mortensen [2011]; Pissarides [2011]; Hall [2012] and the references therein). Our model results would be relevant for an attempt to capture non-stationarities in the first component. Recent empirical work based on the DMP model, (for eg. Shimer [2005]), highlights various aspects of the wage determination phenomenon that are not well explained by the stationary DMP models so far. It would be interesting to introduce non-stationarities in the job-arrival process and work out the implications. Our framework provides a minimal way of attempting that.

1.5. **Currency Unions, International Treaties.** The benefits of joining currency unions or international treaties are variable over time. Moreover, once formed, such treaties are hard and expensive to break. These two factors combine to provide the context for countries trying to devise strategies for entry, exit and crisis management. The recent crisis in the European Monetary Union has rekindled the discussion around the impacts of countries exiting currency unions. Opinions seem to be sharply divided, although the majority opinion seems to be that breaking up of the Euro will be disastrous (Eichengreen [2010]). On the other hand, historically, there have been regular instances of countries leaving currency unions (Rose [2007]). From the design perspective, one of the interesting features of the EMU is that there are no exit clauses. Omission of an well-defined exit clause, at the least, significantly raise the procedural cost of exiting the union, and thus can be seen as a mechanism to avoid hasty reaction by individual countries that could potentially threaten all the other members as a group. Overall, European currency crisis and the responses to it by various parties highlight the importance of the main facets of our model: the trade-off between reacting quickly and flexibly on one hand, and the need to avoid precipitating a fast-moving crisis by unnecessarily hasty actions on the other.

1.6. **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, i_0 and a_0; random times Y_{k+1} at which the state changes from S_k to S_{k+1} according to a partly controlled, imbedded Markov process; and the option to engage in costly actions changes during the inter-arrival stochastic intervals, between Y_k and Y_{k+1}.

The within interval maximization problems will be central to the analysis. 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 entire problem interact with the maximization problems within intervals.

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 $\overline{\pi}$ result. If one abandons the search, the alternative yields expected net profits of $\underline{\pi}$, $\overline{\pi} > \underline{\pi} > 0$. 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 3 (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 $(\overline{\pi} - \underline{\pi})h_W(t_1)dt$, while the expected costs are $(c + r\underline{\pi})dt$ 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 $(\overline{\pi} - \underline{\pi})h_W(t_1^*) = (c + r\underline{\pi})$.

- 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, $h'_W(t_1^*) < 0$.

- Therefore the optimal t_1^* is: higher for higher $\overline{\pi}$, it is worth searching longer when the reward is larger; lower for higher c, one searches less if searching is more costly; lower for higher r, one searches less if one is more impatient; lower for higher $\underline{\pi}$, one searches less when the fallback option is better; and higher for outward shifts in $h_W(\cdot)$, one searches more if search is more productive.

To arrive at the same first order conditions more formally, 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 $\overline{\pi}e^{-rW}$; 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

$$(1)\quad \max_{t_1\in[0,\infty]} \psi(t_1) := \mathbb{E}\Bigg[1_{[0,t_1)}(W)\bigg(\int_0^W (-c)e^{-rt}\,dt + \overline{\pi}e^{-rW}\bigg) + 1_{[t_1,\infty)}(W)\bigg(\int_0^{t_1} (-c)e^{-rt}\,dt + \underline{\pi}e^{-rt_1}\bigg)\Bigg].$$

Evaluating the terms in rounded brackets and rewriting yields

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

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

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

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

$$(4)\quad \psi''(t_1^\circ) = \Big(e^{-rt_1^\circ}G_W(t_1^\circ)\Big)'[0] + (\overline{\pi}-\underline{\pi})h_W'(t_1^\circ),$$

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 $\overline{\pi}$, 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 $\psi'(t_1) \lesseqqgtr 0$ as $\lambda \lesseqqgtr \frac{c+r\underline{\pi}}{\overline{\pi}-\underline{\pi}}$, so that $t_1^* = 0, \infty$ or $[0,\infty]$ depending on the sign of ψ', which does not vary with t_1. Whichever decision is optimal at $t = 0$ is also optimal if one has already waiting to some time T.

Of particular interest are the cases of monotonically increasing and decreasing hazard rates. For example, 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 $t_1^*(T) = \infty$.

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, $\{X_k : k = 0, 1, \ldots\}$ taking values in a two state set, $S = \{i_0, i_1\}$, with the transition matrix $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. The second set of random variables are arrival times, $Y_0, Y_1, \ldots$, 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 $[\![Y_k, Y_{k+1}[\![= \{(\omega, t) : Y_k(\omega) \leq t < Y_{k+1}(\omega)\}$. The Markov chain and the waiting times combine to form the continuous time stochastic process $(\omega, t) \mapsto X(\omega, t)$ defined by

$$X(\omega, t) = \sum_k S_k(\omega) 1_{[\![Y_k(\omega), Y_{k+1}(\omega)[\![}(\omega, t). \tag{5}$$

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

Thus, if $X_0 = i_1$, then the state remains i_1 until Y_1, at which point it switches to i_0, where it stays until Y_2, when it switches back, and so on.

There are two possible actions, $A = \{a_0, a_1\}$. The flow payoffs to being in state i and taking action a are given by

	i_0	i_1
a_0	1	0
a_1	0	2

. This means that one always wants the action to match the state, matching action to state when $X_t = i_0$ earns a flow of 1, matching action to state when $X_t = i_1$ earns a flow of 2, and mismatching earns a flow of 0. For us, what may make instantaneous adjustments suboptimal is the cost of switching from a_0 to a_1 in state i_1, $c(a_0, a_1; i_1)$ is strictly positive, as is $c(a_1, a_0; i_0) > 0$. If another change in the state is expected soon, it may not be worth incuring the cost to enjoy the extra flow. The question is which changes in state to react to? and after what amount of delay?

The **value function**, $V_*(i, a)$, gives the maximal expected discounted utility to starting at $t = 0$ in state $i \in S$ with the present action being a. From Theorem 1 (below), $V_*(i, a)$ is well-defined, can be found as the fixed point to a contraction mapping, and given the value function, the optimal policy can be found by sequentially solving the optimization problems within each stochastic interval $[\![Y_k, Y_{k+1}[\![$ on the presumption that leaving an interval in the state-action pair (S_{k+1}, a_k) at time Y_{k+1} yields a payoff of $V_*(S_{k+1}, a_k)e^{-rY_{k+1}}$.

For the present, we first make the simplifying assumption that $c(a_0, a_1; i) = c(a_1, a_0; i) = C$ for all $i \in S$. The decision problem within an interval $[\![Y_k, Y_{k+1}[\![$ 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 $(a, s) = (a_0, i_0)$ or $(a, s) = (a_1, i_1)$, 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 $\beta = \mathbb{E}\, e^{-rW} < 1$, the value function satisfies

$$V_*(i_0, a_0) = \tfrac{1}{r}(1-\beta) + \beta V_*(i_1, a_0), \tag{6}$$

$$V_*(i_1, a_0) = \max_{t_1 \in [0,\infty]} \mathbb{E}\left[1_{[0,t_1)}(W) \left(\int_0^W 0 e^{-rt}\, dt + e^{-rW} V_*(i_0, a_0) \right) + \right. \tag{7}$$
$$\left. 1_{[t_1,\infty)}(W) \left(\int_0^{t_1} 0 e^{-rt}\, dt + \int_{t_1}^W 2 e^{-rt}\, dt - e^{-rY_k + t_1} C + e^{-rW} V_*(i_0, a_1) \right) \right],$$

$$V_*(i_0, a_1) = \max_{t_1 \in [0,\infty]} \mathbb{E}\left[1_{[0,t_1)}(W) \left(\int_0^W 0 e^{-rt}\, dt + e^{-rW} V_*(i_1, a_1) \right) + \right. \tag{8}$$
$$\left. 1_{[t_1,\infty)}(W) \left(\int_0^{t_1} 0 e^{-rt}\, dt + \int_{t_1}^W 1 e^{-rt}\, dt - e^{-rY_k + t_1} C + e^{-rW} V_*(i_1, a_0) \right) \right],$$

$$V_*(i_1, a_1) = \tfrac{2}{r}(1-\beta) + \beta V_*(i_0, a_1), \tag{9}$$

which, because of the first and the last equations, reduces to a system of two equations in two unknowns.

The value function equations involve two optimization problems, the one at (a_0, i_1) and the one at (a_1, i_0). Monotone comparative statics show that the optimal $t_1^*(a_0, i_1)$ at (a_0, i_1) is smaller than the solution $t_1^*(a_1, i_0)$ (because the flow payoffs of the switch are 2 rather than 1). Let us suppose that the solution at (a_0, i_1) is strictly positive and less than ∞ and examine the determinants of the corresponding $t_1^*(a_0, i_1)$. The tradeoff is between the gain in flow utility and the loss if C is incurred and the state changes back to i_0 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 3 (below), the first order conditions for $0 < t_1^*(a_0, i_1) < \infty$ are

$$[u(a_1, i_1) - u(a_0, i_1)] - rC = h_W(t_1^*) \mathbb{E}\left[C + (V_*(i_0, a_0) - V_*(i_0, a_1)) \right]. \tag{10}$$

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 $[u(a_1, i_1) - u(a_0, i_1)]$ 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 i_1 back to i_0 in the next instant; if this happens, then the decision maker has saved C plus the value difference $V_*(i_0, a_0) - V_*(i_0, a_1)$. In this problem, it is necessary that the LHS be positive in order to ever justify switching to a_1 at (a_0, i_1).

For simplicity, we analyzed the case in which all the switching costs were equal, and this is another potential source of asymmetry. A higher $c(a_0, a_1; i_1)$ would delay optimal switching time from the mismatched state-action pair (a_0, i_1) by shrinking the set of circumstances in which it is worth switching. There would also be an indirect effect — as switches into i_1 from i_0 become more costly, if presently in the mis-matched (a_1, i_0), the opportunity cost of matching action to state is raised, and this implies that there are also delays in the optimal switching time from (a_1, i_0) to (a_0, i_0). Parallel arguments apply to higher values of $c(a_1, a_0; i_0)$, and all of these arguments are reversed for lower switching costs.

3. The Model

We begin 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. There is some delicacy involved in correctly specifying strategies and how they lead to distributions over outcomes, but the essential idea is quite simple: the distribution of the change *time* for the state of the system depends only on the present state, while the distribution of the new state depends on the action the decision maker is taking when the transition occurs. This is a *semi*-Markovian structure because the arrival rate of the transition times can be non-stationary, which implies that the optimal choices of actions may also be non-stationary. The crucial definition that allows us to find and use a recursive structure in this class of problem is the notion of a *stochastic interval*. The basic existence results, and necessary conditions 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 $t \geq 0$:

$$(11)\quad F_W(t) = \int_0^t f_W(x)\,dx;\; G_W(t) = 1 - F_W(t);\; h_W(t) = \frac{f_W(t)}{G_W(t)};$$
$$H_W(t) = \int_0^t h_W(x)dx;\; G_W(t) = e^{-H_W(t)};\text{ and } q_W = e^{-H_W(\infty)}.$$

If $H_W(t) = \int_0^t h(x)\,dx \uparrow \infty$ as $t \uparrow \infty$, then $q_W = 0$. This means 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) If W is an incomplete negative exponential, then its cdf is $F_W(t) = (1-q_W)(1-e^{-\lambda t})$, and its everywhere decreasing hazard rate is $h_W(t) = \lambda\left[(q_W/(1-q_W))e^{\lambda t}+1\right]^{-1}$. If $q_W = 0$, then the hazard rate is constant and the waiting time is memoryless.

(2) An incomplete Weibull distribution is of the form $W = X^\gamma$, $\gamma > 0$, where X is an incomplete negative exponential. Its cdf is $F_W(t) = (1-q_W)(1-e^{-\lambda t^{1/\gamma}})$, and the hazard rate is $h_W(t) = \frac{\lambda}{\gamma} t^{\frac{1-\gamma}{\gamma}}\left[(q_W/(1-q_W))e^{\lambda t^{1/\gamma}}+1\right]^{-1}$.

   (a) If $\gamma > 1$, the Weibull is a convex transformation of the negative exponential, $h_W(0+) = \infty$, and the hazard rate strictly decreases to 0 whether or not $q_W > 0$.

   (b) If $\gamma < 1$, the Weibull is a concave transformation of the negative exponential, $h_W(0) = 0$, if $q_W > 0$, the hazard rate is first increasing then decreasing to 0, if $q_W = 0$, it is strictly increasing.

(3) An Erlang distribution with shape parameter M is the sum of M i.i.d. negative exponentials. It has cdf $F_W(t) = 1 - \sum_{m=0}^{M-1} \frac{1}{m!} e^{-\lambda t}(\lambda t)^m$, $h_W(0) = 0$, and the hazard rate is increasing and concave with $\lim_{t\uparrow\infty} h_W(t) = \lambda$.

There are also a wide variety of mixture model interpretations of hazard rates. For example, if W is a negative exponential with parameter λ where λ is itself random with non-degenerate mixture distribution μ, then the cdf is $F_W(t) = 1-\int e^{-\lambda t}\,d\mu(\lambda)$, the density is $f_W(t) = \int \lambda e^{-\lambda t}\,d\mu(\lambda)$, $h_W(0) = \int \lambda\,d\mu(\lambda)$, and $h_W(t)$

is a strictly decreasing function with tail behavior determined by the behavior of μ at the upper end of its support. More generally, from Edgar's non-compact version of Choquet's theorem Edgar [1975], arbitrary smooth non-monotonic hazard rates can arise from mixtures of different classes of smooth densities.[8]

3.2. **Controlled Semi-Markov Processes.** There is a compact *state space*, S, with generic elements denoted i_0, i, j, and a compact *action space*, A, with generic elements denoted $a, a_0, a_1, \ldots$. For many applications, both S and A are finite.

3.2.1. *Feasible Time Paths.* The time horizon for the optimization problems is $\mathbb{T} := \{0-\} \cup [0,\infty)$ where $0-$ is an "initialization point" assumed to lie strictly to the left of 0, i.e. $0- < 0$. In particular, the interval $[0-, 0)$ contains only the point $0-$. Optimization in our model gives rise to a distribution over *feasible* time paths, $h : \mathbb{T} \to S \times A$, also denoted $t \mapsto (h_S(t), h_A(t))$.

Definition 1. A time path $(h_S(\cdot), h_A(\cdot))$ is **feasible** if:

(a) $h_S(0-) = h_S(0)$;

(b) $h(\cdot)$ is right-continuous, for all $t \in [0,\infty)$, $h(t) = \lim_{\epsilon \downarrow 0} h(t+\epsilon)$; and

(c) for all $(i,a) \in S \times A$, $h^{-1}(i,a)$ is either empty or is a disjoint union of intervals $[r_k, s_k)$ with $r_k \uparrow \infty$ if the union is countably infinite. H denotes the set of feasible time paths.

Combined, parts (a) and (b) of the definition require that the intial state, $h_S(0)$, lasts for a strictly positive amount of time. By contrast, $h_A(0-) \neq h_A(0)$ is possible, and corresponds to the decision maker changing the initial action choice, $a_0 = h_A(0-)$, as soon as $h_S(0)$ is realized. Part (c) requires that the time paths be piece-wise constant with only finitely many jumps in any finite interval of time.

[8] Edgar's theorem implies that any density, f, in a reflexive Sobolev space can be represented as the result of some mixture, μ, over the extreme points of a closed convex set containing f. To find interesting classes of extreme points, note that every distribution on $[0,\infty]$ has a (unique) representation as a distribution on the extreme points of the set of distributions, that the extreme points are point masses, and that point masses are arbitrarily well approximated by smooth distributions, e.g. a Weibull distribution with parameters λ and γ converges to point mass on λ as $\gamma \uparrow \infty$.

3.2.2. *Stochastic Intervals.* For any $h \in H$, $Y_k(h)$ will denote the time at which the k'th change in $h_S(t)$ occurs. Formally, define $Y_0(h) = 0$, define $Y_{k+1}(h)$ as $\min\{t > Y_k(h) : h(t) \neq \lim_{s \uparrow t} h(s)\}$ with the convention that $\min \emptyset = \infty$. The distribution on H that arises depends on the decision maker's choice of policy f. Given f, we define a distribution on H in two parts: by specifying the distribution of the waiting times between state changes, $W_0 := 0$ and $W_k := Y_k - Y_{k-1}$ for $k \geq 1$; and by specifying the transition kernel of the state changes at the Y_k's. Stochastic intervals are a useful tool for this.

Definition 2. The **stochastic interval** between Y_k and Y_{k+1} is a subset of $H \times [0, \infty)$ defined as $[\![Y_k, Y_{k+1} [\![= \{(h, t) : Y_k(h) \leq t < Y_{k+1}(h)\}$.

One standard use of stochastic intervals is to define distributions over sets of rcll paths.

Example 1 (Queues and embedded Markov processes)**.** Suppose that $(S_k)_{k=0}^{\infty}$ is a discrete time Markov process in S, $W_0 \equiv 0$, and $(W_k)_{k=1}^{\infty}$ are and independent collection of i.i.d. negative exponentials, and $Y_k := \sum_{j \leq k} W_j$ where the S_k and W_k are defined on a probability space $(\Omega, \mathcal{F}, P)$. In $\Omega \times [0, \infty)$, the stochastic intervals are defined by $[\![Y_k, Y_{k+1} [\![= \{(\omega, t) : Y_k(\omega) \leq t < Y_{k+1}(\omega)\}$. With these stochastic intervals, the random rcll path, $X(\omega, t) := \sum_k S_k(\omega) 1_{[\![Y_k(\omega), Y_{k+1}(\omega) [\![}(\omega, t)$, has the distribution of a Poisson process with an embedded Markov chain. Simple queueing models of the number of customers in line waiting to be served start with a collection (A_k, B_k) of independent negative exponentials, with the A_k corresponding to interarrival times of new customers and the B_k to the interarrival times of service completions, defines $W_k = \min\{A_k, B_k\}$, sets

$$S_{k+1} = \begin{cases} S_k + 1 & \text{if } A_k < B_k \\ \max\{S_k - 1, 0\} & \text{if } A_k > B_k. \end{cases}$$

If the B_k are instead a sequence independent Erlang distributions with shape parameter M, the queueing model corresponds to each customer needing a total of M services, with each service have an independent negative exponential time till completion. More general arrival and service distributions, and more general transition rules for the state space give rise to the standard queueing models.

Because we are interested in decision makers whose choice of actions determine the distribution over the feasible rcll paths, we replace the probability space, Ω, with the space of feasible paths, H, in the definition of stochastic intervals. We will define probability distributions on H by specifying the distributions of the stochastic intervals in $H\times[0,\infty)$, the distributions of the states at the beginnings of the intervals, and the distribution of the actions taken during the intervals.

3.2.3. *Policies.* For each $t \in \mathbb{T}$, let $\mathcal{H}_t$ denote the smallest σ-field of subsets of H making the evaluation mappings, $h \mapsto h(s)$, $s \leq t$ measurable. For each Y_k, let $\mathcal{H}_{Y_k}$ be the smallest σ-field of subsets of H making the restriction mapping $h \mapsto h_{|[0,Y_k(h)]}$ measurable.

We will represent policies by specifying what they do *within* a stochastic interval as a function of what has happened before the beginning of the interval. To this end, let A_∞ denote the compact $A \cup \{a_\infty\}$, where the "book-keeping" point a_∞ is, by assumption, at a distance equal to the diameter of A from each $a \in A$. Action plans are then elements, $(\boldsymbol{\tau}, \boldsymbol{\alpha})$, of $\mathfrak{I} \subset [0,\infty]^{\mathbb{N}} \times A_\infty^{\mathbb{N}}$

$$\begin{aligned} \mathfrak{I} = \{(\boldsymbol{\tau},\boldsymbol{\alpha}) \in [0,\infty]^{\mathbb{N}} \times A_\infty^{\mathbb{N}} : \boldsymbol{\alpha}_n = a_\infty \text{ iff } \boldsymbol{\tau}_n = \infty, \\ \boldsymbol{\tau}_n \uparrow \infty,\ \boldsymbol{\tau}_n \leq \boldsymbol{\tau}_{n+1} \text{ with equality iff } \boldsymbol{\tau}_n = \infty, \text{ and} \\ \boldsymbol{\alpha}_n \neq \boldsymbol{\alpha}_{n+1} \text{ for all } \boldsymbol{\alpha}_n \neq a_\infty.\} \end{aligned} \tag{12}$$

The vector $\boldsymbol{\tau}$ gives the planned change times, and the vector $\boldsymbol{\alpha}$ gives the new actions chosen at the change times.

For $k = 0$, the initial state-action pair at Y_k is defined as $(h_S(0), h_A(0-))$, for $k \geq 1$, it is defined as $(h_S(Y_k), h_A(Y_k))$.

Definition 3. A policy is a sequence $(f_k)_{k=0}^\infty$, of $\mathcal{H}_{Y_k}$-measurable functions from H to $\mathfrak{I}$ with the property that the first action called for in any interval is different from the action in the initial state-action pair. A policy is **semi-Markovian** if the f_k depend only on the initial state-action pair at Y_k.

A semi-Markovian policy can be represented by a single measurable function $f : S \times A \to \mathfrak{I}$.

3.2.4. *Policy Induced Distributions Over Outcomes.* In our model, policies induce distributions over feasible histories, utility depends on the feasible history that is realized, and the decision maker will chose the policy yielding the highest expected utility. To make this coherent, we must specify how policies induced distributions. For simplicity, we do this only for semi-Markovian strategies, changing the indexing appropriately yields the general case.

We assume

(1) that there is a continuous mapping, $i \mapsto Q_i(\cdot)$, from S to $\Delta^d((0,\infty])$, the set of distributions on $(0,\infty]$ having $d$ensities with respect to Lebesgue measure on $(0,\infty)$;

(2) the expected waiting times under Q_i, $i \in S$ are uniformly bounded away from 0; and

(3) that there is a jointly continuous mapping, $(i,a) \mapsto p(\cdot|i,a)$ from $S \times A$ to $\Delta(S)$, the set of distributions on S, and that it satisfies $p(\{i\}|i,a) = 0$ for all $i \in S$ and $a \in A$.

Given a semi-Markovian policy, $f = ((\boldsymbol{\tau}_n)_{n=1}^{\infty}, (\boldsymbol{\alpha}_n)_{n=1}^{\infty})$, we define the induced distribution by induction through the stochastic intervals.

$k = 0$: By definition, the interval $[\![Y_0, Y_1[\![$ starts in the state-action pair $h(0-) = (i_0, a_0)$. Let $f(i_0, a_0) = ((\boldsymbol{\tau}_n)_{n=1}^{\infty}, (\boldsymbol{\alpha}_n)_{n=1}^{\infty})$. The distribution of Y_1 is given by Q_{i_0}. The distribution of S_1 is given by $p(\cdot|i_0, \boldsymbol{\alpha}_{N_0})$ where $N_0 := \min\{n : Y_0 + \boldsymbol{\tau}_n < Y_1\}$.

$k \geq 1$: The interval $[\![Y_k, Y_{k+1}[\![$ starts in the state action pair $h(Y_k) = (i_k, a_k)$. Let $f(i_k, a_k) = ((\boldsymbol{\tau}_n)_{n=1}^{\infty}, (\boldsymbol{\alpha}_n)_{n=1}^{\infty})$. The distribution of $Y_{k+1} - Y_k$ is independent of the previous waiting times and given by Q_{i_k}. The distribution of S_{k+1} is given by $p(\cdot|i_0, \boldsymbol{\alpha}_{N_k})$ where $N_k := \min\{n : Y_k + \boldsymbol{\tau}_n < Y_{k+1}\}$.

The following "book-keeping" result implies that every choice of policy has a well-defined expected utility.

Lemma 1. *For every semi-Markovian f and starting point (i_0, a_0), there exists a unique distribution, $\mathcal{L}_f$, on the rcll paths $t \mapsto (i(t), a(t))$ from $\mathbb{T}$ to $S \times A$ having the properties described above:*

(a) *the distributions of the waiting times between state-jumps, $W_{k+1} = Y_{k+1} - Y_k$ are independent and distributed according to Q_{S_k};*

(b) *within any stochastic interval, $[\![Y_k, Y_{k+1}[\![$, action changes take place at each $\boldsymbol{\tau}_n$ with $Y_k + \boldsymbol{\tau}_n < Y_{k+1}$, and the action changes to $\boldsymbol{\alpha}_n$; and*

(c) *if $Y_{k+1} \in [Y_k + \boldsymbol{\tau}_n, Y_k + \boldsymbol{\tau}_{n+1})$, then $P(S_{k+1} \in E) = p(E|S_k, \boldsymbol{\alpha}_n)$.*

Proof. From [Billingsley, 1999, §12], it is sufficient to show that conditions (a)-(c) determine the finite dimensional distributions of the stochastic process, which is slightly tedious, but entirely straightforward. □

3.2.5. *Payoffs.* Payoffs have two components, an integrated discounted flow depending on the state and the present action, and a discounted sum of change costs. A stochastic interval, $[\![Y_k, Y_{k+1}[\![$, starts at a state-action pair (i_0, a_0). Until a change of action, the decision maker receives a flow payoff, $u(i_0, a_0)$, continuously discounted at a rate r. If the decision maker changes to an action a_1 at $Y_k + t_1 < Y_{k+1}$, the decion maker incurs a cost $c(a_0, a_1; i_0) > 0$ which has discounted present value $c(a_0, a_1; i_0)e^{-r(Y_k+t_1)}$, they change the flow payoff to $u(i_0, a_1)$, and they change the distribution of S_{k+1} to $p(\cdot|i_0, a_1)$. Further changes within the interval $[\![Y_k, Y_{k+1}[\![$ are treated the same.

(1) We assume that $u : S \times A \to \mathbb{R}$ is bounded and jointly continuous.

(2) We also assume that $c : [(A \times A) \setminus D] \times S$ is jointly continuous, uniformly bounded away from 0, and satisfies $c(a_0, a_1; i) + c(a_1, a_2; i) \geq c(a_0, a_2; i)$ for all $i \in S$ and all $a_0, a_1, a_2 \in A$ where $D \subset (A \times A)$ denotes the diagonal.

(3) Combining, the utility of a feasible path h given by $t \mapsto (i(t), a(t))$ from $\{0-\} \cup [0, \infty)$ to $S \times A$ is

$$U(h) := \int_0^\infty u(h(t))e^{-rt}\,dt - \sum_{t_k \in T_a} c(a_k, a_{k+1}; i(t_k))e^{-rt_k} \tag{13}$$

where T_a is the set of time points where the actions change along the path $t \mapsto a(t)$, and T_a includes 0 if $h_A(0-) \neq h_A(0)$.

Two comments on the modeling choices seem appropriate. First, unlike the subsequent intervals, at $t = 0$ there is no previous interval of time at which the pre-change action is being chosen. This is what necessitates the introduction of the "time" 0−.[9] Second, it is quite possible that $c(a_0, a_1; i) \neq c(a_1, a_0; i)$. This allows the model to capture partial irreversibilities of various strengths.

Example 2. In the optimal search time problem of §2.1, take $S = \{i_0, i_1\}$ where i_0 is the initial state and i_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

[9]See Simon and Stinchcombe [1989] and Stinchcombe [1992] for related solutions to this representation problem.

from i_0 to i_1 but never the reverse. Let $A = \{a_0, \underline{a}, \overline{a}\}$ where a_0 is the action "continue searching" and $\underline{a}$ is the action "use the alternative (older) technology." and $\overline{a}$ is the action "use the newly found technology" where we include the costs of changing action in the profits. Q_{i_0} is the distribution of W, and $Q_{i_1}(\infty) = 1$, which correpsonds to i_1 being an absorbing state. This means that the problem has only one relevant stochastic interval, $[\![0, Y_1[\![$. We arrange for $\overline{a}$ not to be chosen in i_0 either by making the transition cost $c(\cdot, \overline{a}; i_0)$ prohibitive or by making the utility flow $u(i_0, \overline{a})$ sufficiently low.

3.3. **The Bellman Equation.** The value function $V_* : S \times A \to \mathbb{R}$ is defined as

$$V_*(i_0, a_0) = \sup_f E^f U(h) := \sup_f \int U(h)\, d\mathcal{L}_f(h), \tag{14}$$

where f is a policy, $\mathcal{L}_f$ is the distribution on H induced by the policy f, and the supremum is taken over all policies.

The essential recursive structure is found in the observation that $U(h)$ can be broken into a sum of payoffs within stochastic intervals. When one leaves a stochastic interval $[\![Y_k, Y_{k+1}[\![$ to the state S_{k+1} while taking an action a_k, one starts the new stochastic interval in the state-action pair (S_{k+1}, a_k). If one behaves optimally thereafter, one receives $V_*(S_{k+1}, a_k)e^{-rY_{k+1}}$. With this structure, it becomes clear that if we choose optimally within each interval $[\![Y_k, Y_{k+1}[\![$ taking into account that we will receive the value $V_*(S_{k+1}, a_k)e^{-rY_{k+1}}$ when we exit to (S_{k+1}, a_k) at Y_{k+1}, then we will have solved the problem in (14).

If one starts with $h(Y_k) = (i_k, a_k)$, then the contribution of the interval $[\![Y_k, Y_{k+1}[\![$ to $E^f U(h)$ is

$$E^f \left(U(h) \cdot 1_{[\![Y_k, Y_{k+1}[\![} \middle| h(Y_k) = (i_k, a_k) \right) = \tag{15}$$

$$E^f \left(\int_0^\infty u(h(t))e^{-rt} 1_{[\![Y_k, Y_{k+1}[\![}(h, t)\, dt - \right.$$

$$\left. \sum_{t_k \in T_a \cap [\![Y_k, Y_{k+1}[\![} c(a_k, a_{k+1}; i(t_k))e^{-rt_k} \middle| h(Y_k) = (i_k, a_k) \right). \tag{16}$$

If one receives $V_0(i_{k+1}, a_{k+1})$ for starting the interval $[\![Y_{k+1}, Y_{k+2}[\![$ at the state-action pair (i_{k+1}, a_{k+1}), then the the problem at Y_k at the state-action pair $h(Y_k) =$

(i_k, a_k) is given by

$$(17) \qquad V_1(i_k, a_k) = \max_f E^f \left(U(h) \cdot 1_{[\![Y_k, Y_{k+1}[\![} + V_0(h(Y_{k+1}))e^{-rY_{k+1}} \middle| h(Y_k) = (i_k, a_k) \right).$$

To minimize on the notational burden, we condition on $Y_k = t$, subtract t, renormalize the k to 0, and start the discounting from 0, which changes (17) to

$$(18) \qquad V_1(i_0, a_0) = \max_f E^f \left(U(h) \cdot 1_{[\![Y_0, Y_1[\![} + V_0(h(Y_1))e^{-rY_1} \middle| h(Y_0) = (i_0, a_0) \right).$$

The set of bounded continuous functions on $S \times A$ is denoted $C_b(S \times A)$.

Theorem 1. *If $\beta = \sup\{E\, e^{-rW} : W \sim Q_i,\ i \in S\} < 1$, then the mapping from $V_\circ$ to $T(V_\circ)$ defined by*

$$(19) \qquad T(V_\circ)(i_0, a_0) = \max_{f=(\boldsymbol{\alpha},\boldsymbol{\tau})} E^f \left(U(h) \cdot 1_{[\![Y_0, Y_1[\![} + V_\circ(h(Y_1))e^{-rY_1} \middle| h(Y_0) = (i_0, a_0) \right)$$

is a well-defined contraction mapping on $C_b(S \times A)$, its contraction factor is at most β, its unique fixed point is the true value function, V_*, and following the policy

$$(20) \qquad f^* = \operatorname*{argmax}_{f=(\boldsymbol{\alpha},\boldsymbol{\tau})} E^f \left(U(h) \cdot 1_{[\![Y_0, Y_1[\![} + V_*(h(Y_1))e^{-rY_1} \middle| h(Y_0) = (i_0, a_0) \right)$$

in every stochastic interval achieves the value V_*.

The condition that $\sup\{E\, e^{-rW} : W \sim Q_i,\ i \in S\} < 1$ is equivalent to the assumption that the expectation of the necessarily strictly positive waiting times is uniformly bounded away from 0.

The detailed proof is in the Appendix. It consists of several steps: showing that one can restrict attention to a compact subset of $\mathbb{I}$; that the expected utility function is continuous on this set; that the mapping is a contraction; and that the fixed point is indeed the true value function.

4. Euler Equations

We give the Euler equations, i.e. the necessary conditions, for the optimal policy within stochastic intervals for which one move is optimal. Formally, this corresponds to the set of stochastic intervals for which the optimal $\boldsymbol{\tau}_n^*$ is equal to ∞ and $\boldsymbol{\alpha}_n^* = a_\infty$ for all $n \geq 2$.[10] We briefly discuss the Euler equations for multiple moves after the proof of the following. Note that there are two Euler equations, one for $\boldsymbol{\tau}_1$, the optimal time to change actions, and one for $\boldsymbol{\alpha}_1$, the optimal choice of action to change to.

4.1. **Memoryless Processes.** To focus attention on the effect of the non-stationarity of the waiting time distributions, we first show that: the optimal policy is always immediate response or infinite patience if the waiting times are memoryless; and that, if forced to make a change, the optimal choice does not depend on the time at which the choice is made.

Theorem 2. *If $(W_k)_{k=1}^\infty$ are independent negative exponentials with parameter $\lambda(S_k)$, then in each stochastic interval $[\![Y_k, Y_{k+1}[\![$,*

(a) *the optimal waiting time, $\boldsymbol{\tau}_1^*$, if unique, is either 0 or ∞,*

(b) *if $\boldsymbol{\tau}_1^*$ is not unique, then any time in $[0,\infty]$ is indifferent, and*

(c) *the optimal action if the decision maker is **forced** to move at some time $Y_k + t < Y_{k+1}$ depends only on the current state, $h_S(Y_k)$, and is independent of t.*

Proof. By the memorylessness of the negative exponentials, at every time $Y_k + t < Y_{k+1}$, the choice between different actions leads to the same distribution over S_{k+1} and the distribution of $Y_{k+1} - t$ is the same as it was at Y_k. Thus, if it is optimal to change actions at $Y_k + 0$, then it is optimal to change to the same action if one arrives unchanged at any $Y_k + t < Y_{k+1}$, and if it is optimal to leave the action unchanged at Y_k, the same is true at any $Y_k + t < Y_{k+1}$. □

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 else wait at least until the next change. Also, within this class of problems, there is always at most one change of action.

4.2. **Processes with Memory.** We say that a waiting time process, $(W_k)_{k=1}^\infty$, has memory if the hazard rates, $h_{W_k}(\cdot)$, are non-constant.

[10]The class of problems for which this is applicable models situations where the cost of changes, $c(\cdot,\cdot;\cdot)$, is high relative to the differences in the flow utility, $u(\cdot,\cdot)$. A broad class of institutional decision making problems fit this description.

Theorem 3. *Suppose that* $(W_k)_{k=1}^{\infty}$ *is a waiting time process with memory and that* $[\![Y_k, Y_{k+1}[\![$ *is a stochastic interval in which the hazard rate is strictly positive and exactly one change of action is optimal, and that it happens at an interior* $\boldsymbol{\tau}_1^* = t_1^*$ *and* $\boldsymbol{\alpha}_1^* = a_1^*$*. Then* (t_1^*, a_1^*) *must satisfy the following conditions,*

$$(21)\quad h_W(t_1^*)\mathbb{E}[C + (V_*(a_k, S_{k+1}) - V_*(a_1^*, S_{k+1}))] = [u(a_1^*, i_k) - u(a_k, i_k)] - rC,$$

$$(22)\quad \int_{t_1^*}^{\infty} e^{-ry}\frac{\partial}{\partial a}u(a, i_k)|_{a_1^*}(1 - F_W(y))dy + \int_{t_1^*}^{\infty} e^{-ry} f_W(s, y)\frac{\partial}{\partial a}\mathbb{E}V_*(a, S_{k+1})|_{a_1^*} = 0$$

where $C = c(a_k, a_1^*; i_k)$*. Further, if* $\mathbb{E}[C + (V_*(a_k, s_{k+1}) - V_*(a_1^*, s_{k+1}))] > 0$ *(resp.* < 0*), then the second order conditions are strictly satisfied iff the hazard rate is increasing (resp. decreasing) at* t_1^**.*

Proof. We set the expected benefit to waiting for an instant dt at t_1^* equal to the expected benefit of acting at that instant.[11] The benefit to waiting is dt times $h_W(t_1^*)\mathbb{E}[C+(V_*(a_k, S_{k+1})-V_*(a_1^*, S_{k+1}))]$ because: $h_W(t_1^*)dt$ is the probability that the state changes; one saves the cost $C = c(a_k, a_1^*; i_k)$; and one has a change in expected value of $\mathbb{E}(V_*(a_k, S_{k+1}) - V_*(a_1^*, S_{k+1}))$. The benefit to acting is dt times $[u(a_1^*, i_k) - u(a_k, i_k)] - rC$ because: the perpetual annuity flow value of the cost of moving from a_k to a_1^* is $rC \cdot dt$; and the change in utility flow is $[u(a_1^*, i_k) - u(a_k, i_k)] \cdot dt$.

For the first order condition for a_1^*, the first term is directly seen to be the derivative of $E^f\left(U(h) \cdot 1_{[\![Y_0, Y_1[\![}\middle|h(Y_0) = (i_0, a_0)\right)$ with respect to a_1, and the second is the derivative of $E^f\left(V_*(h(Y_1))e^{-rY_1}\middle|h(Y_0) = (i_0, a_0)\right)$ with respect to a_1.

For the last part, note that because $h_W(t_1^*) > 0$, to satisfy (21), the terms on both sides must have the same sign. If the left-hand sign is positive and $h_W'(t_1^*) > 0$, then the benefits of waiting are increasing while the benefits of acting are constant. Thus, for strict satisfaction of the second order conditions, one must have $h_W'(t_1^*) < 0$. The reasoning for the left-hand side being negative leading to an increasing hazard rate at the solution is analogous. □

4.3. **Interpretation of the Euler Equations.** The increasing and decreasing hazard rates cases in Theorem 3 have very different intuitions, one consonant

[11] An alternative to the present derivation writes out all of the integrals, evaluates and simplifies, takes derivatives, gathers terms and simplifies. The example in §2.1 gives an example of what is involved.

with the search theory model in §2.1, the other consonant with delaying expensive preventive measures until danger presses.

4.3.1. *Increasing Hazard Rates at the Optimum.* Suppose that one is in a state-action pair (i_k, a_k) that is mis-matched in the sense that $u(a_k, i_k) < \max_{a \in A} u(a, i_k) - rc(a_k, a; i_k)$. Suppose also that, given the action a_k, one expects the arrival of S_{k+1} at Y_{k+1} to be good news, that is, a_k will come very close to solving $\max_{a \in A} u(a, S_{k+1}) - rc(a_k, a; S_{k+1})$. If the likelihood of the arrival of Y_{k+1} is climbing and the cost of waiting is constant, it cannot be optimal to incur large action change costs right now because one might be able to stay put and have the world come to you. On the other hand, if the likelihood of the arrival of Y_{k+1} decreases to a low enough level, it becomes optimal to more closely adapt the present action to the present state, that is, to solve $\max_{a \in A}[u(a, i_k) - rc(a_k, a; i_k)]$.

4.3.2. *Decreasing Hazard Rates at the Optimum.* In many situations, there are costly safety measures that one would rather avoid if the danger they are meant to protect against, S_{k+1}, is far enough in the future, that is, if Y_{k+1} looks comfortable remote. Immunizations for possible diseases are costly, at the very least in terms of time, and one would rather delay them even if their effectiveness does not decay over time. Imposing some kind of control on powerful institutions whose behavior has the potential to inflict huge negative externalities is costly, at the very least in terms of the political capital that must be expended. However, as the arrival of S_{k+1} comes closer, i.e. as the hazard rate $h_{W_{k+1}}$ increases, at some point taking the expensive precautionary measures becomes the optimal choice.

To see how this works in the left- and right-hand sides of (21), consider first the term

$$[c(a_k, a_1^*; i_k) + \mathbb{E}\left(V_*(a_k, S_{k+1}) - V_*(a_1^*, S_{k+1})\right)]. \tag{23}$$

Suppose that S_{k+1} is likely to be the sort of disaster that drastically lowers utility flows unless some action, a_1^*, is taken. Suppose further that $c(a_k, a_1^*; i_k) \ll c(a_k, a_1^*; S_{k+1})$, that is, suppose than an ounce of prevention is worth a pound of cure. This means that we can expect $V_*(a_1^*, S_{k+1}) \gg V_*(a_k, S_{k+1})$, enough so that the whole term in (23) is negative. This negative term, corresponding to the 'benefit' to waiting, is multiplied by a decreasing hazard rate term, that is, the 'benefit' is negative and becoming more negative as the hazard rate increases.

Now, a maximizing choice between two negative terms involves picking the one closest to 0, and on the other side of the equation are the utility flow terms,

$$[u(a_1^*, i_k) - u(a_k, i_k)] - rc(a_k, a_1^*; i_k). \tag{24}$$

The preventive action being costly corresponds to this term being negative, either because $u(a_k, i_k) > u(a_1^*, i_k)$ or because the cost, $-rc(a_k, a_1^*; i_k)$, dominates. While this is negative, it is not decreasing over time. Eventually, at t_1^*, this negative term becomes better than the first one.

5. 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.

Using our framework, we can glean some insights about examples presented at the beginning. Women's suffrage was long delayed. Part of the delay may be explained by near-irreversibility of any enfranchisement. It is politically much more costly to disenfranchise a group than to enfranchise them, this added to 'optimal' delay in that case. On the other hand, in case of Prohibition, although banning a popular item of consumption by itself was an expensive proposition, it was easily and cheaply reversible, which probably goes some way to explain the relativeness swiftness of both the enactment and repeal. Cuban missile crisis, Vietnam war, Watergate scandal — none of these led to any particularly extensive checks on, or built-in delays to, executive power. It could partly be for the reason that the typical class of problems the executive branch is called on to solve often involve the need for quick decisions.

We have presented results for reactive problems where the stochastic kernel of the SMP depends on the choice of action only through the Markov state transition process. There are real life situations which are better modeled as fully controlled SMP's where the choice of action effects also the distribution of arrival times. That would be the next step in our analysis. Here are some examples where such stochastic processes would be relevant.

Economic models of climate change have traditionally treated the process as one of gradual change to new, stable state. Recent research in 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 [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.[12]

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

[12] While 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 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].

context, the cost of controlling the intensity of the arrival rate process seems to be the major issue.

Immunizations, breaking up banks "too big to fail," — the costs of these preventive actions before the crisis may not be negligible, but the costs of achieving the same outcomes after the crisis has arrived are immensely larger (these are also cases where the optimal delay can happen with an increasing hazard rate). Choice problems involving preventive actions would warrant models incorporating hazard rate controls.

Future research would tackle these extended set of models incorporating hazard rate controls. Cases where we only need controls executed at discrete times would be straightforward extensions of our current work. Substantive work would be needed to deal with models involving continuous control of hazard rates.

References

Akhil R. Amar. *America's constitution: a biography*. Random House Trade, 2006.

Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999.

Svetlana Boyarchenko and Sergei Levendorskiĭ. *Irreversible decisions under uncertainty: Optimal stopping made easy*, volume 27 of *Studies in Economic Theory*. Springer, Berlin, 2007.

Edward G. Carmines and James A. Stimson. *Issue evolution: Race and the transformation of American politics*. Princeton University Press, 1990.

P. K. Chintagunta. Inertia and variety seeking in a model of brand-purchase timing. *Marketing Science*, 17(3):253–270, 1998.

Dean Corbae, Maxwell B. Stinchcombe, and Juraj Zeman. *An introduction to mathematical analysis for economic theory and econometrics*. Princeton University Press, 2009.

Peter Diamond. Unemployment, vacancies, wages. *American Economic Review*, 101(4):1045, 2011.

Gerald A. Edgar. A noncompact Choquet theorem. *Proc. Amer. Math. Soc.*, 49: 354–358, 1975.

Barry Eichengreen. The breakup of the euro area. *NBER Working Paper*, 2010.

Zachary Elkins, Tom Ginsburg, and James Melton. *The endurance of national constitutions*. Cambridge University Press, 2009.

Eleanor Flexner and Ellen F. Fitzpatrick. *Century of struggle: The woman's rights movement in the United States*. Belknap Press, 1996.

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.

Robert E. Hall. The 2010 nobel prize in economics: How the dmp model explains current high unemployment. 2012.

Russell Hardin. *Liberalism, constitutionalism, and democracy*. Oxford University Press, USA, 2003.

James Madison. The federalist no. 46. *The Federalist Papers*, page 299, 1961.

Dale T. Mortensen. Markets with search friction and the dmp model. *American Economic Review*, 101(4):1073, 2011.

William D. Nordhaus and Joseph Boyer. *Warming the world: economic models of global warming*. the MIT Press, 2003.

Victor Ortego-Marti. Unemployment history and frictional wage dispersion. *Working Paper, Department of Economics, Stanford*, 2011.

Elinor Ostrom. *Governing the commons: The evolution of institutions for collective action*. Cambridge University Press, 1990.

Christopher A. Pissarides. Equilibrium in the labor market with search frictions. *American economic review*, 101(4):1092, 2011.

Andrew K. Rose. Checking out: exits from currency unions. *Journal of Financial Transformation*, 19:121–128, 2007.

Robert Shimer. The cyclical behavior of equilibrium unemployment and vacancies. *American economic review*, pages 25–49, 2005.

Leo K. Simon and Maxwell B. Stinchcombe. Extensive form games in continuous time: pure strategies. *Econometrica*, 57(5):1171–1214, 1989.

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.

Nicholas H. Stern. *The economics of climate change: the Stern review*. Cambridge University Press, 2007.

Maxwell B. Stinchcombe. Maximal strategy sets for continuous-time game theory. *Journal of Economic Theory*, 56(2):235–265, 1992.

Nancy L. Stokey. *The economics of inaction: Stochastic control models with fixed costs*. Princeton University Press, 2009.

Cass R. Sunstein. The paralyzing principle. *Regulation*, 25(4):32–37, 2002.

6. Appendix

Proof of Theorem 1: To show that the mapping is well-defined, we must show that the problems in (19) have a solution and that the maximized value of the solution is continuous in (i_0, a_0). To this end, we will show that there is no loss in ruling out all but a compact subset of within interval policies $(\boldsymbol{\tau}, \boldsymbol{\alpha})$. Because costs are bounded away from 0 by some strictly positive $\underline{c}$, there is no loss in assuming that we are optimizing over the subset of $\mathbb{I}$ with $\underline{c}\sum_n e^{-r\boldsymbol{\tau}_n} \leq \int_0^\infty Be^{-rt}\,dt$ where $B > \sup_{i,a} u(i,a) - \inf_{i,a} u(i,a)$. Since costs $c(\cdot,\cdot;i)$ are defined on the open complement of the diagonal, we must establish that there is no loss in restricting attention to the subset of $\mathbb{I}$ having $d(\boldsymbol{\alpha}_n, \boldsymbol{\alpha}_{n+1}) \geq \epsilon$ for some $\epsilon > 0$ and all n with $\boldsymbol{\tau}_n < \infty$. Because $S \times A$ is closed, hence compact, and the continuous mapping $(i,a) \mapsto (Q_i, p(\cdot|i,a))$ is therefore uniformly continuous. Hence, for some $\epsilon > 0$, $d(a,a') < \epsilon$ implies that for all $i \in S$, the potential gain of moving from (i,a) to (i,a') is less than $\underline{c}$. The subset of $\mathbb{I}$ with $\underline{c}\sum_n e^{-r\boldsymbol{\tau}_n} \leq \int_0^\infty Be^{-rt}\,dt$ and $d(\boldsymbol{\alpha}_n, \boldsymbol{\alpha}_{n+1}) \geq \epsilon$ for some $\epsilon > 0$ and all n with $\boldsymbol{\tau}_n < \infty$ is compact, and on this set the utility function is continuous. By the Theorem of the Maximum, T is a well-defined mapping from bounded continuous functions to bounded continuous functions.

From Blackwell's Lemma for contraction mappings (e.g. [Corbae et al., 2009, Lemma 6.2.33]), to show that T is a contraction mapping with contraction factor β, it is sufficient to show that (i) T is monotonic, and (ii) that for any constant κ and $V_\circ \in C_b(S \times A)$, $T(V_\circ + \kappa) \leq T(V_\circ) + \beta\kappa$. Monotonicity is immediate. For (ii), note that the optimized value of $E^f(V_\circ(h(Y_1)) + \kappa)e^{-rY_1}$ is at least as large as the optimized value of $E^f V_\circ(h(Y_1))e^{-rY_1}$, so that the difference between the two optimized values must be less than or equal to $\kappa E\, e^{-rW}$ where $W \sim Q_{i_0}$, and $E\, e^{-rW} \leq \beta$.

The arguments for the last two parts of the theorem parallel the standard discrete-time dynamic programming arguments: if $V_\dagger$ is the unique fixed point for T, then following f^* for the first stochastic interval and receiving $V_\dagger(h(Y_1))$ when the interval ends must have value $V_\dagger$. Therefore following f^* for the first two stochastic intervals has value $V_\dagger$. Since $Y_k \uparrow \infty$ because the expectation of the strictly positive W_k is bounded away from 0, following f^* in each interval yields $V_\dagger$. By definition, $V_* \geq V_\dagger$, and if $V_*(i_0, a_0) > V_\dagger(i_0, a_0)$ for some (i_0, a_0), then $V_\dagger$ is not the fixed point of T. $\square$