Winter 2017

Conservatism and Switcher's Curse

Aaron Edlin

Conservatism and Switcher's Curse

Aaron Edlin, *University of California, Berkeley*

Send correspondence to: Aaron Edlin, Department of Economics and School of Law, and National Bureau of Economic Research; Email:edlin@berkeley.edu

This article formally models the virtues of Edmund Burke's conservatism, characterizes the optimal level of conservatism, and applies the model to management, law, and policy. I begin by introducing "switcher's curse," a trap in which a decision maker systematically switches too often. Decision makers suffer from switcher's curse if they forget the reason that they maintained incumbent policies in the past and if they naively compare rival and incumbent policies with no bias for incumbent policies. Conservatism emerges as a heuristic to avoid switcher's curse. The longer a process or policy has been in place, the more conservative one should be. On the other hand, the more conservative were past decision makers, the more progressive one should be today.

1. Introduction

Consider a CEO who rarely if ever changes a long-standing policy, even when she cannot identify a good reason to stick with the status quo. For her,

I thank George Akerlof, Jonathan Carmel, Stefano Dellavigna, Daniel Hemel, Benjamin Hermalin, Louis Kaplow, Brent McDonnell, David Kennedy, Barry Nalebuff, Mitch Polinsky, Catherine Roux, Alan Schwartz, Carl Shapiro, Joseph Stiglitz, Philipp Strack, Eric Talley, Abraham Wickelgren, Justin Wolfers, Jeffrey Zwiebel, and seminar participants at Columbia, the University of Chicago, the University of California Berkeley, NHH in Bergen, Norway, University of St. Gallen, Switzerland, the American Law and Economics Association Meetings, and the National Bureau of Economic Research Summer Institute for helpful conversations. I owe particular thanks to George Akerlof and Benjamin Hermalin for encouraging me to write up my ideas about conservatism and the switcher's curse.

American Law and Economics Review
doi:10.1093/aler/ahw022

the fact that things have always been done a certain way is reason itself not to change.

In management, this CEO would be seen as a dinosaur and a clear victim of status-quo bias. But in political philosophy, this CEO's position has a long tradition. Edmund Burke wrote in 1790:

> You see, Sir, that in this enlightened age I am bold enough to confess that we [the English] . . . instead of casting away all our old prejudices, we cherish them ... and, to take more shame on ourselves, we cherish them because they are prejudices; and the longer they have lasted, and the more generally they have prevailed, the more we cherish them.

This article formally models the virtues of Burkean conservatism, characterizes the optimal level of conservatism, and applies the model in management, law, and government policy, to individuals and to organizations. The fact that things have always been done a certain way can be a perfectly good reason not to change, *particularly* if one does not know why they have always been done that way.

Not everyone would agree. Oliver Wendell Holmes (1897), for example, wrote that "[i]t is revolting to have no better reason for a rule of law than that so it was laid down in the time of Henry IV." Many judges no doubt wish they could ignore precedent and simply cite to Holmes.

Likewise, it is surely tempting to a *new* CEO to channel Holmes and announce: "We will look at everything with fresh eyes." But, the problem with such an approach is that if past rationales have been forgotten, or are simply unknown to a new decision-maker, as they often will be to a new CEO, then this progressive approach will lead to too much switching. The new CEO will be a victim of what I call "switcher's curse."

Looking at alternative policies with fresh eyes ignores the fact that a long-standing policy has likely competed against other policies in the past and only survived because it won those competitions. Perhaps it even competed against the very rival it faces today. A policy's longevity testifies to its virtue, and those with imperfect recall should only switch to policies that appear substantially better. The new CEO could run into switcher's curse if she applies her decision-making philosophy and if she (and her organization)

cannot remember the reasons that policies were put in place and stayed in place.[1]

Switcher's curse is a trap for the unwary, inexperienced, or boundedly rational decision maker. The trap is that a new policy looks better than it actually is, or an existing policy looks worse than it is.

Two colleagues have provided examples which I hope ring bells for the reader. One pointed out that he has frequently wasted days, weeks, and even months pursuing ideas for simpler proofs of a theorem in a years-old working paper only to finally realize that he had tried these approaches before, sometimes many times before. Perhaps he should have put weight on the idea that his proof was the way it was for good reason, even if he did not remember the reason, and been more skeptical of new approaches.

Another colleague recently changed an important grading policy in a course he had taught for many years. To alleviate student worries about exams and to improve his teacher ratings, this colleague decided to put more weight on problem sets and ceased grading them on a "check/no-check" basis (whereby nearly everyone got a check). This switch turned out poorly when it emerged that some students had access to the problem sets and solutions from earlier years through fraternities and the like. This was a big problem now that there was substantial weight on problem sets. The irony was that this colleague had early in his teaching career adopted the check/no-check, low-weight grading policy for problem sets exactly so that he could reuse his perfected problem sets year after year without worrying about some students having solutions. He had kept the policy for many years for exactly that reason. Eventually, though he forgot the reasons for his own grading policy and changed it to improve ratings only to have his ratings plummet.

1. Organizations have difficulty remembering the past because employees come and go, because even when they stay, they suffer from human frailties of poor memory, and because information can become siloed. Hirshleifer and Welch (2002) provide an extensive review of psychological literature to the effect that actions will be easier to remember than the reasons for actions. In this article's vernacular, actions are policies and I assume that the CEO can tell the difference between a new policy and an incumbent one even though she cannot remember why the incumbent policy was put in place. Anand et al (1988) provide an instructive example in which managers at a major aerospace company after "another wave of changes in management" decided that a technology was critical but "no one remembered that there was an expert already on staff."

The experiences of Company X, a technology company, serve as another illustration of switcher's curse.[2] During the dot com boom, a few individuals who created and consumed intellectual property (content), concluded that both creators and consumers of content were poorly served by existing systems of production and distribution. They invented a radical new model. Established organizations, though, were far too conservative to try their approach, so they founded Company X.

As entrepreneurs and innovators, they evaluated everything with fresh eyes and never privileged the status quo. They were constantly changing the systems that they put into place at Company X. At the outset their flexibility and progressivism paid off and they rarely regretted switching to a new policy.

After about five years, however, "improvements" to their system began to go awry. The new systems they put in place had flaws that they did not anticipate and old systems had virtues they had forgotten. Frequently, they decided to switch back to their old system. Sadly, even this full circle was not necessarily the end. Several years later, this cycle might be repeated as the whole lesson was forgotten and someone again had a new clever idea or even the same clever idea. Once switching began to go awry at Company X the problem seemed to get worse with time.

This story raises a question. Why did Company X profit from early switches but eventually suffer switcher's curse? It is not simply that X's managers ran out of good ideas. That would simply mean a slower arrival rate of good ideas and less switching. Instead, the problem was that over time Company X began to switch when it should not, suffering from switcher's curse.

This article presents a very simple model that provides a possible answer. Policies compete sequentially after signals of their quality are drawn. The decision maker has limited information in that she does not know the past signals of a policy's quality or the past performance of a policy. What she does know is how long a policy has been in place.

I begin by comparing a "naive" decision maker who does not take into account the past at all (including the longevity of a policy) with a fully rational Bayesian updater who, even though she does not remember past

2. The real name of Company X is withheld.

signals, makes the most of the fact that a status-quo policy must have won past contests.

In a simple two-period example with no switching cost, the naive decision maker switches if contemporaneous information about a new policy is better than that for a status-quo policy, and is indifferent to switching if information is the same and there are no switching costs. Such a decision maker will switch too often, suffering switcher's curse.

In contrast, I find that a fully-rational decision maker who compares a status-quo policy that appears otherwise equal (or close) to an alternative policy should not switch even if there are no switching costs. The reason is straightforward. The incumbent policy competed against another policy last year (or might have done so) and its incumbency is itself a signal that it is a good policy, even if the specific reasons that the policy was kept are forgotten.

Given that an incumbent policy was judged good in the past, it is likely that the reason that it looks relatively bad today is either because today's evaluation study drew a relatively pessimistic signal of the policy's quality or a relatively optimistic signal of the alternative. I find that the longer an incumbent policy has survived, the more pronounced these effects and the heavier the thumb should be on the scale in favor of the incumbent policy.

Conservatism is a heuristic that addresses switcher's curse. A conservative decision maker will hesitate to make changes, particularly to a long-standing policy, even when she cannot articulate or pinpoint reasons not to switch. Only if some new policy is truly compelling will she switch, and the longer a policy has been in place, the more compelling any alternative must be. A naive decision maker, in contrast, estimates with "fresh eyes" the virtues of a new policy and the status-quo policy and switches if the estimated virtues of switching justify the transition cost. She will suffer switcher's curse. Although Holmes finds conservatism "revolting," conservatism has a big virtue: it avoids switcher's curse.

In practice, of course, conservatism is a matter of degree and different people and organizations display different degrees of conservatism. In this article's model, as in life, it is possible to be too conservative. How conservative one should be depends on many factors. One factor turns out to be how conservative past decision makers were. If past decision makers were extremely conservative, little is learned from their decision to keep a policy

and it will therefore pay to be relatively progressive today. In contrast, if a policy survived a long time despite being subject to the vicissitudes of progressive management, which is exactly what happened at Company X, then the policy is very likely good and a great degree of conservatism is warranted today. Thus, the innovators at Company X eventually needed to change their progressive approach and become extremely conservative, or they needed to be replaced with conservative management. In turn, extremely conservative decision makers might enjoy success for a while but should eventually moderate or be replaced with progressives.

To summarize, our main propositions state:

1. Conservatism and Switcher's Curse. A naive decision maker switches too often because she puts existing and alternative policies on the same footing. It is frequently optimal to be conservative and stick to the status quo even when contemporaneous information suggests switching.
2. The older the policy, the more conservative one should be today. This is true regardless of a decision-maker's level of conservatism in the past.
3. The more progressive were past decision makers, the more conservative it is optimal to be today, and conversely, the more conservative were past decision makers, the more progressive it is optimal to be today.

Propositions 1 and 2 formalize Burke's intuition. Proposition 3 is new, as far as I know.

Section 2, below, discusses related literature. Section 3 presents a simple two-period example that captures the intuition of switcher's curse and its cure, conservatism. Subsequently, Section 4 explores a multi-period model with variable switching costs where conservatism is a matter of degree. After developing the main model, the article applies its basic lesson to several contexts: the system of *stare decisis* in which courts are loath to change legal rules; administrative law; venture capital; separation of powers; the evolution of a firm; the cases of a new CEO and a new employee; cycles of meta-policy and policies; and finally, the decision to stay married or to divorce. Appendix B presents a case study of Company X, providing further motivation for the model.

2. Related Literature

Conservatism is closely related to the well-known psychological phenomenon of status-quo bias (Porter and McIntyre, 1984; Samuelson and Zeckhauser, 1988; Kahneman et al., 1991; Burmeister and Schade, 2007). Status-quo bias is generally understood to be a flaw, but this article's analysis sugggests that far from a flaw, status-quo bias could be a healthy psychological adaptation that helps avoid "switcher's curse."

In turn, switcher's curse in this article's model, can be understood to arise from either of two other well-documented psychological phenomena, the "availability bias" (Tversky and Kahneman, 1973) or "prior neglect" (Kahneman and Tversky, 1973, 1982; Bar-Hillel, 1980). The availability bias means that agents may overemphasize recent or contemporaneous information and deemphasize past information which is less available. Even without memory loss, this would explain our naive agent's mistakes. Likewise, the naive agent in our model can be seen as a victim of prior neglect because she neglects the difference in priors between an incumbent policy and a proposed alternative.[3] In sum, then, this article shows that status-quo bias may actually compensate for two other biases, prior neglect and the availability bias, which would otherwise cause switcher's curse.

Several economics papers have explored various aspects of conservatism. Prendergast and Stole (1996) and Zwiebel (1995) explore how conservatism can emerge when manager's talents are unobservable and they seek to increase their pay by protecting or improving their reputations. Conservatism protects the managers in these models at the expense of the companies they serve. In this article, in contrast, managerial conservatism protects companies and might be undertaken by wise principals not just by misbehaving agents.

Other related papers are Rasmusen (1992), Hirshleifer and Welch (2002) and Baliga and Ely (2011). Rasmusen argues based on reversion to the mean that when the status quo has known payoff and an innovation receives a particularly good signal, reversion to the mean suggests that the innovation

3. In our context such prior neglect is even more likely to arise than in the experiments where it is discovered, because sophisticated inference is required in our context even to figure out that a status-quo policy is likely good. In contrast in many prior-neglect experiments, the difference in base rates is either stated or supposed to be common knowledge to subjects.

is not as good as it appears. Rasmusen's reversion to the mean argues for caution when estimating any uncertain return, not for prioritizing the status quo; in fact, in contrast to this article's model, Rasmusen points out that "[i]f the true profitability of the status quo were unknown, the conservative bias would disappear."

Hirshleifer and Welch (2002), like this article, consider a setting of limited memory. They are interested, however, in fully-rational decision making. This article, in contrast, studies the impact of potentially irrational rule-of-thumb decision-making styles, characterized on a spectrum of progressivism to conservatism, and asks what metapolicy or heuristic makes most sense given the level of progressivism or conservatism of prior decision makers. Switcher's curse only arises when decision-makers are irrationally progressive (in the sense of this article) and cannot arise for the rational agents of Hirshleifer and Welch (2002).

Baliga and Ely (2011) also explore the consequences of limited memory. They argue that "sunk cost bias is a optimal heuristic that compensates for the constraints of limited memory." Their point is closely related to this article in that sticking with a project (sunk cost bias) can be seen as a form of conservatism. Their paper differs from this one in a couple of ways. First, I study a multi-period model in a stationary environment, whereas they study a single project that will be at a different state at date 2; this means that while their conclusion can be seen as a form of this article's Proposition 1, this article's Proposition 2, which states that the longer a policy has been in place the more conservative the decision maker should be does not have a counterpart in their paper. Second, they are studying rational behavior, so nothing analogous to switcher's curse arises, whereas this article compares rational and irrational behavior in a setting where past decision makers deviated from rationality. This distinction leads to the discovery that past progressivism makes conservatism today more sensible (Proposition 3), which also has no counterpart in their model.

3. A Simple Two-Period Example of Switcher's Curse and Conservatism

Here, I present a simple two-period example that illustrates switcher's curse and its cure, conservatism. Policies are either good (G) or bad (B),

with good policies having a higher expected payoff than bad policies. A good policy is always good, and a bad policy always bad. For a court, a policy could be that the court and its lower courts only enforce liquidated damages for breach of contract if the damages are a reasonable estimate of expectation damages. For Congress, the Affordable Care Act is a policy. For a company, a policy might be requiring managers to confer with their own bosses before firing employees or it could be using Microsoft Windows instead of Apple computers. For an individual, a policy might be being single or being married.

A decision maker does not directly observe whether a policy is good or bad, but instead observes a noisy contemporaneous signal of its quality. Past signals are forgotten, but a decision maker knows that a status-quo policy was used in the past. The fact that a policy was used in the past means that it won some contest with another policy.

The basic question in this article is how naive decision-making compares with optimal decision making under limited information. By "naive," I mean that the decision maker assumes that the base rate probability that the incumbent policy is good is the same as that for the general population from which a competing policy is drawn. This captures the idea of comparing policies with "fresh eyes," as the founders of Company X did. It also captures the well-known phenomenon in which people ignore differences in base rates (the base-rate fallacy or prior neglect).

Every policy is independently drawn with probability $p \in (0, 1)$ of being good (G) and $1 - p$ of being bad (B).

3.1. Time 0

At time 0, a decision maker must choose whether to use policy A or policy D. She does not observe whether either policy is good or bad. Instead she observes independent signals $d_0 \in \{g, b\}$ of D and $a_0 \in \{g, b\}$ of A. A signal "g" indicates that the associated policy *appears* good and a signal "b" indicates that the policy *appears* bad.

Every signal is statistically independent of every other signal. The signals are accurate with probability $s > 1/2$, which represents signal strength. Thus, the chance that any good policy appears to be good or that any bad policy appears to be bad is s; the chance that any good policy appears to be

bad or any bad policy appears to be good is $1 - s$. Formally,

$$Pr(a_0 = g \mid A = G) = s = Pr(a_0 = b \mid A = B) \tag{1}$$

$$Pr(d_0 = g \mid D = G) = s = Pr(d_0 = b \mid D = B). \tag{2}$$

The time 0 decision rule is:

1. If $a_0 = d_0$, then flip a coin to choose between A and D.
2. Otherwise, choose the policy that appears to be good.

This decision rule is fully rational because the prior probability that A is good equals the probability that D is good. It is also what a naive decision maker would do who simply assumed that the base rates were identical for the two policies.

3.2. Time 1

Without loss of generality, assume that policy A was chosen at time 0 and so A is the status-quo policy at time 1. The question is whether to switch to some other policy $E \in \{G, B\}$.

The time 1 decision maker does not directly observe whether E or A are good or bad but instead observes a signal $e \in \{g, b\}$ of E and signal $a \in \{g, b\}$ of A, where

$$Pr(a = g \mid A = G) = s = Pr(a = b \mid A = B) \tag{3}$$

$$Pr(e = g \mid E = G) = s = Pr(e = b \mid E = B). \tag{4}$$

The time 1 decision maker either has limited recall or is a new decision maker without records of the time 0 contest. Either way, she does not know a_0 or d_0.

After comparing a with e, the decision maker chooses whether to keep A or switch to E. I compare "naive" and "conservative" decision-making approaches as specified below:

Naive	$e = b$	$e = g$
$a = b$	Flip coin	Switch to E
$a = g$	Keep A	Flip coin

Conservative	$e = b$	$e = g$
$a = b$	Keep A	Switch to E
$a = g$	Keep A	Keep A

The two decision makers differ on the diagonal when signals are equal. The naive decision maker is indifferent between policy A and E in that case and flips a coin to choose between them. In contrast, the conservative decision maker favors the status quo and so keeps A when signals are equal. She puts a thumb on the scales in favor of the status quo and would not switch in that case even if you paid her a small amount to do so.

Which approach is better? The conservative approach turns out to be better because positive information is embedded in incumbency as shown by the following lemma.

LEMMA 1 $Pr(A = G \mid A \textit{ won at time } 0) > Pr(A = G) \equiv p$

Proof. See Appendix A. □

The intuition for Lemma 1 is straightforward. Winning at time 0 is a good signal and Bayesian updating implies that once A wins at time 0, the probability that it is good increases above p.

In the example, incumbency entails a certainty that a prior contest occurred. In reality, it only entails a positive probability of a prior contest. Lemma 1 would hold in either case.

Lemma 1 is the basis for concluding that the conservative approach dominates the naive approach.

OBSERVATION 1 The conservative approach is strictly better than a naive approach in the two-period example.

Proof. See Appendix A. □

The key to the observation is that the conservative and naive decision maker only disagree on the diagonals when both policies appear bad or both appear good. The naive decision maker is indifferent to the policies in these cases, but should not be.

According to Lemma 1, the incumbent policy is likely better (prior to time 1 signals), so the naive decision maker errs by flipping a coin when time 1 signals are equal. This is an example of what I call switcher's curse. Just as *winner's curse*, in an auction, involves a failure to take into account the negative information entailed in another party's lower bids, and can lead a bidder to bid too high and win an auction too often, here *switcher's curse* involves a failure to take into account the positive information entailed in winning past contests, and can lead a decision maker to switch too often.

The naive decision maker fails to take account of the fact that A being a status-quo policy is itself a good signal about A. The conservative approach incorporates this information by not switching in cases where contemporaneous signals indicate that the two policies are equally good.

In practice, of course, conservatism comes in degrees and entails more than just refusing to switch when contemporaneous information is in equipoise. A true conservative will often refuse to switch to policies that appear better, and sometimes even significantly better. That turns out to be sensible. If we generalize the example to allow the rival policy to appear slightly better than the incumbent policy, the best decision rule would be to keep the incumbent policy provided the rival policy did not appear to be too much better.

Below, I generalize the example in a slightly different direction. I explore contests where switching costs may be positive and are a continuous random variable. This allows me to study a continuum of conservatism. At the same time I extend the model to a multiperiod model. These two changes allow me more fully to explore the phenomenon observed at Company X. Switcher's curse was not a big factor early on there, but it eventually bit severely for two reasons. First, switcher's curse is larger for older policies, and second, switcher's curse is larger when past decision makers were highly progressive as were the innovators at Company X. Each of these factors suggests that an optimal decision maker should have become highly conservative at Company X over time.

4. Multiperiod Model: Conservatism on a Continuum

I consider here the decision problem of choosing between a status-quo policy A and an alternative policy E^t at time t (today) in a multi-period

model. First, I will ask how the optimal choice depends upon the length of time that an incumbent policy has been in place. Second, what kinds of mistakes would be made by a naive decision maker who ignored the importance of incumbency? And third, how are the answers to the above two questions affected by the extent of prior or future decision makers' conservatism or progressivism? This third question asks how a decision maker today should optimally account for past and future nonoptimal decision making (either by others or by she herself).

In order to assess degrees of conservatism, I introduce a cost of switching to E^t given by x^t. Each period this cost is independently drawn from the uniform distribution $U[0, x_{max}]$ and is observable.[4] A conservative person switches from a policy that appears bad to one that appears good only if x^t is low. A more progressive (or impulsive) decision maker will have a higher threshold for switching costs. The decision maker's level of progressivism or conservatism defines a rule of thumb for decision making.

At time 0 a policy is drawn that has probability p of being good and $1-p$ of being bad. At each subsequent time t, there is a contest: an alternative policy E^t is drawn (also with probability p of being good) and a decision maker draws a signal e^t of E^t, a new signal a^t of the status-quo policy A, and a new switching cost x^t; each signal is statistically independent of all others. The signals each have strength s, as described in the two-period example above.

Below, I will suppress the time superscripts whenever confusion is unlikely to result. As in the two-period example, while the decision maker prefers a good policy, she does not know whether a policy is of type G or B, and so must form probability estimates. Because the model now has switching costs, even a naive decision maker would not switch if signals are equal.

Switching is tempting only if A appears bad and E appears good (i.e. if $a = b$ and $e = g$) and if the switching cost x is sufficiently low. The decision maker's problem is to figure out how low constitutes "sufficiently low."

4. The difficulty of knowing the cost of switching prior to any switch is another important source of switcher's curse that I will explore in a different paper. If Kahneman and Tversky's (1979) "planning fallacy" applies, then, people will have a tendency to underestimate switching costs.

Define a decision-maker's threshold for switching as x^*, which is to say that the decision maker switches only if $a = b,\ e = g$, and $x < x^*$.

This framework admits a continuum of conservatism or progressivism as follows.

Definition 1 The greater is the switching threshold x^* the more **progressive** is the decision maker.

Definition 2 The lower is the switching threshold x^* the more **conservative** is the decision maker.

Definition 3 A decision maker is **completely conservative** if $x^* = 0$.

The idea behind these definitions is that a conservative decision maker is relatively skeptical that a new policy is better and therefore will only switch under a relatively low switching cost. A progressive decision maker is open to the idea that a new policy is better than the status quo and willing to switch even at relatively high switching cost if an alternative policy appears better. An extremely conservative decision maker switches only if switching costs are very low (and alternative policies appear better). The higher is x^* the more prone is a decision maker to switch and the more progressive and less conservative we say she is.

Let us begin by considering an optimal decision rule today at time $t > 1$ taking as given the level of conservatism of past and future decision-makers. I study the optimal decision rule today given the following information set:

> **Information Set:** *The decision maker knows (1) current signals, (2) the longevity of the status-quo policy, and (3) the conservatism or progressivism of past and future decision rules.*

The motivation for studying such an optimum, rather than optimal behavior today assuming past and future optimal behavior, is that in practice, it is likely that past decision makers behaved nonoptimally (they were human), and the current decision-maker cannot depend upon future decision-makers to behave optimally.[5]

5. The decision maker may only have beliefs about the conservatism or progressivism of past and future decision rules in which case the optimal decision rule should be understood to be optimal conditional on these beliefs.

The current decision maker does not know past signals either because she was not the past decision maker or because she has imperfect recall. The assumption is extreme in many settings where at least a signal of the signal is available — in law, for example, past opinions constitute an imperfect signal of the signal. The model highlights the value of preserving institutional memory, such as through written judicial opinions or corporate management information systems. Because a conservative philosophy emerges as a substitute for good memory, the value of institutional memory is higher for progressive decision makers.

The current decision maker also does not know whether the alternative policy "E" has ever been tried in the past or is entirely new. Future research might explore what happens if a decision maker knows that an alternative policy "E" was used in the past and for how long.

In this model, a rational decision will depend upon the probability that A is good, the probability that E is good, the switching cost x, and the difference in present discounted value between having a good and bad policy, a difference that I study in the next subsection.

4.1. The Present Discounted Value of Good and Bad Policies

The present discounted values of a good policy, denoted $V(G)$, and of a bad policy, denoted $V(B)$, depend upon flow payoffs and the frequency of switching, which in turn depends upon the degree of conservatism or progressivism of future decision makers. For tractability, I shall assume that the level of conservatism of future decision makers does not vary with time.

Let the expected per-period payoff of a good policy be π_G and of a bad policy be π_B. The decision maker does not observe the expected payoff, or equivalently whether the policy is good or bad, but only observes the realized payoff which takes values of 0 or 1 with probability $1-s$ and s for a good policy G and with probability s and $1-s$ for a bad policy B. The realized payoff is simply the signal of the policy's quality and has strength $s > 1/2$.

Policies follow a Markov process with the state G, B being unknown to the decision maker and transition probabilities $p_{G\to B}$ and $p_{B\to G}$ depending upon the degree of conservatism of future decision makers. A future decision

maker will inadvertently switch away from a good policy to a bad one if and only if (1) the incumbent good policy A appears to be bad; (2) the alternative policy E is bad but appears to be good; and (3) a switching cost below the decision-maker's threshold is drawn. Likewise, a future decision maker who unknowingly has a bad policy will switch and luck into a good policy if and only if (1) the incumbent bad policy appears to be bad, (2) the alternative policy E is good and appears to be good, and (3) the switching cost is below the decision-maker's threshold.

Let x^*_{future} be the threshold for switching of future decision makers, so that they switch when $x < x^*_{future}$. Then, these transition probabilities are as follows:

$$p_{G\to B} = (1-s)\Big[(1-p)(1-s)\Big]\left[\frac{x^*_{future}}{x_{max}}\right], and \tag{5}$$

$$p_{B\to G} = s\Big[ps\Big]\left[\frac{x^*_{future}}{x_{max}}\right]. \tag{6}$$

Then the expected present discounted value of the random future stream of payoffs from having a good policy G can be calculated as follows. A flow payoff of π_G is earned. Then with probability $p_{G\to B}$, the policy will be switched to a bad policy next period so that a payoff of π_B would be earned in the following period. Of course, in the case of switching, a switching cost will be incurred that will on average be $x^*/2$. With probability $1-p_{G\to B}$ the good policy will be maintained and π_G will again be earned in the following period. Assuming a discount factor $\delta < 1$ for future payoffs, $V(G)$ and $V(B)$ are found by solving the following pair of simultaneous equations:

$$\begin{aligned} V(G) &= \pi_G + \delta\left[p_{G\to B}\big(V(B) - x^*/2\big) + (1-p_{G\to B})V(G)\right] \\ V(B) &= \pi_B + \delta\left[p_{B\to G}\big(V(G) - x^*/2\big) + (1-p_{B\to G})V(B)\right]. \end{aligned} \tag{7}$$

4.2. Optimal Decision Making at Time t

Below I characterize the optimal switching rule for a decision maker today, where "optimal" means the best decision given the limited information available (i.e. that past signals are forgotten) and given that past and future decision makers have rule-of-thumb switching thresholds that may not be optimal. I will then compare the decisions of an optimizing decision

maker at time $t > 1$ to one who is naive and ignores the information embedded in incumbency. I will also explore how optimal decisions depend upon the progressivism or conservatism of past and future decision makers.

Given that switching costs are positive, the decision maker will consider switching only when the incumbent policy appears bad and an alternative appears good: i.e. when $a = b$ and $e = g$ (where I have suppressed the time t superscripts on a & e). The question is how low switching costs x must be to make switching worthwhile.

Consider any time $t > 1$. Let I denote the event that $a = b$ and $e = g$. The expected value of policy E given I is

$$V(E \mid I) = Pr(E = G \mid I)V(G) + [1 - Pr(E = G \mid I)]V(B) \quad (8)$$

and the expected value of policy A is

$$V(A \mid I) = Pr(A = G \mid I)V(G) + [1 - Pr(A = G \mid I)]V(B). \quad (9)$$

Hence, the expected amount by which E is better than the status quo A equals

$$V(E \mid I) - V(A \mid I) = [Pr(E = G \mid I) - Pr(A = G \mid I)]\,[V(G) - V(B)]. \quad (10)$$

This makes sense: the extra expected value of policy E over A equals (1) the increased probability that the policy E is good over the baseline probability that policy A is good, times (2) the expected extra value derived from a good policy.

For a decision maker who seeks to maximize the expected value of the policy, it is optimal to switch if and only if the switching cost is less than the above quantity: that is, if and only if $x < x^*_{optimal}$, where

$$x^*_{optimal} \equiv [\Pr(E = G \mid I) - \Pr(A = G \mid I)]\,[V(G) - V(B)]. \quad (11)$$

The above expression defines the optimal switching threshold, $x^*_{optimal}$, at any time $t > 1$. If prior to receiving signals at time t, the decision maker thought it much more likely that A is good than that E is good, then it is possible that the right-hand side of the above expression is negative. That would mean that it is optimal to ignore time t signals so that even if E appeared good, A appeared bad, and $x = 0$, it would still be optimal not

to switch. For fully-rational decision makers that would imply being in an informational cascade (see Bikhchandani et al., 1992).

Let p_A denote the prior probability that the decision maker believes that A is good and p_E denote the prior that E is good, where each prior is formed immediately before seeing the period t signals about the quality of the policies. I assume that $p_A < 1$ and $p_E < 1$.

The decision maker can form posteriors according to Bayes's rule. Thus, for example, if E appears to be good and A appears to be bad, the posteriors are as follows:

$$Pr(E = G|e = g, p_E) = \frac{sp_E}{p_E s + (1 - p_E)(1 - s)} \tag{12}$$

$$Pr(A = G|a = b, p_A) = \frac{(1 - s)p_A}{p_A(1 - s) + (1 - p_A)s}. \tag{13}$$

These posteriors are the relevant probabilities to use to find $x^*_{optimal}$ using Equation (11).

It is intuitive that the posterior probability that A is good despite receiving a bad signal will be increasing in the prior probability that A is good. This intuition is confirmed by the following lemma.

Lemma 2 $Pr(A = G|a = b, p_A)$ is increasing in p_A, the probability that $A = G$ prior to the signal a.

Proof. From Equation (13) we have that

$$Pr(A = G|a = b, p_A) = \frac{(1 - s)}{(1 - s) + s(1 - p_A)/p_A}, \tag{14}$$

which is increasing in p_A because $\frac{1-p_A}{p_A}$ is decreasing in p_A. □

Corollary 1 follows directly from the lemma above, the formula for $x^*_{optimal}$, the fact that A and E are statistically independent, and the fact that the event under consideration is the event where A appears bad and E appears good.

Corollary 1 The optimal switching-cost threshold $x^*_{optimal}$ is decreasing in p_A.

4.3. Switcher's Curse from Naive Decision Making and Conservatism as a Cure

I will now show that a naive decision maker who ignores the information embedded in incumbency will switch too often, suffering what I dub "Switcher's Curse." At the same time, I will show that conservatism is a heuristic that can eliminate switcher's curse.

Formally, I define a naive decision maker as follows:

DEFINITION 4 A **naive** decision maker does not consider the importance of past decisions for the probability that A is good; she instead assumes that $p_A = p_E = p$ and then computes the optimal switching threshold given these beliefs.

A naive decision maker's switching threshold is:

$$\begin{aligned} x^*_{naive} &= [Pr(E = G|e = g, p_E = p) - \Pr(A = G|a = b, p_A = p)] \\ &\quad \times [V(G) - V(B)] \\ &= \left[\frac{sp}{ps + (1-p)(1-s)} - \frac{(1-s)p}{p + s - 2ps}\right][V(G) - V(B)] \\ &> 0. \end{aligned} \tag{15}$$

As in the two-period model, a naive decision maker does not consider the possibility that incumbent policies have a higher base-rate of being good than new policies. Perhaps, the naive decision maker commissions a study of the incumbent policy and of the alternative without telling those doing the study which is the incumbent policy (perhaps to avoid status-quo bias) and switches whenever the new policy appears better than the old taking into account transition costs.

I will now argue that a naive decision maker will switch too often. Too often, the naive decision maker will switch to a bad E when A is good, and this will not be compensated for by the increased chance of abandoning a bad status-quo A and getting a good alternative E. This is what I mean by switcher's curse and why I say that naive decision makers will suffer switcher's curse.

I demonstrate switcher's curse in several steps. The first is to observe that if A wins a contest against E under an arbitrary switching threshold

$x^* > 0$, then the probability that A is good increases as a result of the win. I formally state this in the following lemma.

Lemma 3 If at any time t, (1) A wins a contest with E, (2) the decision maker is not completely conservative and (3) $p_A \in (0,1)$, then the posterior probability that A is good will exceed the prior: that is, $Pr(A = G \mid \text{A wins at t}, x^* > 0, p_A) > p_A$ for any $p_A \in (0,1)$. When the decision maker is completely conservative, then winning does not increase the likelihood that A is good: that is, $\Pr(A = G \mid A \textit{ wins at } t, x^* = 0, p_A) = p_A$ for any p_A.

Proof. See Appendix A. □

The intuition behind Lemma 3 is that when $x^* > 0$, winning is a good signal of A's quality. Of course, A might have won even though $a = b$, but it is more likely to win if $a = g$, and a good signal g is more likely to occur if $A = G$. This observation leads a fully-rational decision maker to a positive updating of p_A after a win. Of course, in the extreme case that a decision maker is completely conservative, winning provides no information and so the posterior equals the prior.

Because a naive decision maker is boundedly rational and neglects the fact that the prior p_A should reflect the updating from past victories of the status-quo policy, a naive decision maker does not appreciate how good the status quo is before observing signals. A naive decision maker treats A and E symmetrically and starts with the unupdated probability p. If prior decision makers were completely conservative, refusing to switch under all circumstances, then the naive decision maker's belief that $p_A = p$ would be unbiased.

The optimality of being more conservative than a naive decision maker follows from Lemma 3.

Proposition 1 The optimal decision making rule is more conservative than the naive one — that is, $x^*_{optimal} < x^*_{naive}$ — provided that at least one earlier decision maker was not completely conservative.

Proof. When A was first drawn, it had probability p of being good. Winning with a completely conservative decision maker does not update the

probability that A is good. However, Lemma 3 shows that each time A wins a contest in which $x^* > 0$, the probability that $A = G$ increases. Thus, $p_A > p$.

The naive switching cost threshold would be optimal by construction if $p_A = p$. Now recall that Corollary 1 states that $x^*_{optimal}$ is decreasing in p_A. It follows that $x^*_{optimal} < x^*_{naive}$. □

The above proposition states that naive decision makers switch too often. This phenomenon, I call switcher's curse because the decision maker too often switches away from a good policy because she failed to account for the good signal embedded in incumbency. Even if she cannot remember the reasons that past decision makers stuck with A (or that she herself did in the past), the fact that they did is informative: A is better on average than the contemporaneous signal suggests.

Switcher's curse can be caused by the "base-rate fallacy" also known as "prior neglect," a well-known phenomenon that has been documented in the psychology literature by Kahneman and Tversky (1972, 1973, 1985) and Bar-Hillel (1980). In experiments, people often ignore the differences in base rates for two choices A and E, and focus on specific information, even when told explicitly that A and E have different base rates.[6] Our decision-maker's problem is harder than those in the Kahneman–Tversky laboratory because no one tells our decision maker that status-quo policies are very likely to be good. Somehow he needs to figure out that status-quo policies tend to be good. Because this problem is subtle, even someone not subject to the base-rate fallacy in those experiments might still be subject to switcher's curse.

Switcher's curse also can result from the "availability bias," another well-known heuristic discovered by Tversky and Kahneman (1973). The availability bias is a tendency to overemphasize current information and underemphasize past information. Our naive agent suffers from the availability bias.

6. In a famous example, a decision maker who thinks she sees a blue cab typically ignores the base rate that almost all cabs are green and concludes that in fact she probably saw a blue cab because that is how it appeared. Remarkably, the base-rate fallacy occurs in the lab even when people are told that cabs are mostly yellow (that is the Kahneman and Tversky experiment). Such mistakes can cause switcher's curse.

The fact that it is optimal to resist change, even when change seems good, shows that conservatism is a heuristic that can avoid switcher's curse. Status-quo bias is not so much a bias as it is a useful adaptation to counterbalance other biases, namely prior neglect and the availability bias. Of course, more conservatism is not always better, for there is an optimal degree of conservatism given by Equation (11).

4.4. Switcher's Curse and the Need for Conservatism Grow for Older Policies

Below, I show an important fact about switcher's curse. Switcher's curse becomes a bigger problem, the longer the status-quo policy has been in place.

Proposition 2 Suppose that prior decision makers were not completely conservative and had identical levels of conservatism $x^* > 0$. Then, the optimal decision rule becomes more conservative the longer the status-quo policy A has been in place — that is, $x^*_{optimal}$ is decreasing in the longevity of the incumbent policy A.

Proof. As observed in the proof of the prior proposition, when A was first drawn, it had probability p of being good. And, Lemma 3 shows that each time A wins a contest in which $x^* > 0$, the probability that $A = G$, which is the p_A for next period, increases. Corollary 1 then shows that every time A wins a contest in which $x^* > 0$, the optimal switching cost threshold $x^*_{optimal}$ decreases in the next period. Hence, $x^*_{optimal}$ decreases over time with the longevity of A; it pays to be more conservative the longer A has been in place. □

Proposition 2 implies that the optimal conservative policy is not just to be resistant to change, but to be particularly resistant to switching away from long-standing policies. All change is not equal. Switcher's curse is larger for older policies, because these policies will have survived more past contests. A wise decision maker avoids switcher's curse by being conservative, and according to Proposition 2, the longer A has been in place the more conservative she should be.

4.5. Why it is Optimal to be More Conservative if Past Decision Makers were Progressive

Here I explore the effects of past decision makers' degree of conservatism or progressivism.

The posterior that a winning policy is good will increase with the progressivism of a decision maker. Put precisely,

LEMMA 4 After the status-quo policy A wins at any time t, where $x^* > 0$, the posterior that A is a good policy increases with the progressiveness of a decision maker and decreases with her conservatism: that is, $\frac{\partial \Pr(A=G|\text{A wins})}{\partial x^*} > 0$.

Proof. See Appendix A. □

The intuition behind Lemma 4 is straightforward. If a decision maker is extremely conservative then very little is learned from the fact that a policy won. Even if signals were bad, it would be likely to win, so updating would not be strong. On the other hand, if the decision maker is progressive (a large x^*), then winning becomes meaningful and leads to substantial updating of the prior that $A = G$.

I will now explore how the conservatism or progressivism of past decision makers affects the level of conservatism that is appropriate today. As a preliminary, I state a straightforward lemma.

LEMMA 5 For a policy A, where the prior $p_A \in (0, 1)$, the posterior probability that A is good conditional on winning at t is increasing in the prior probability that A is good: that is, $\frac{\partial \Pr(A=G|A \text{ wins}, p_A)}{\partial p_A} > 0$ for any $p_A \in (0, 1)$.

Proof. See Appendix A. □

PROPOSITION 3 Suppose that past decision makers all had the same $x^* > 0$. Then, the more conservative they were, the more progressive the current period-t decision maker should be, and the more progressive they were, the more conservative the current decision maker should be: that is, $x^*_{optimal}$ is decreasing in the switching-cost threshold x^* that past decision makers applied.

Proof. See Appendix A. □

The basic logic of the proposition's proof is that increases in x^* at any time t directly increase the posterior at that time t that A is good, holding time t's priors fixed, according to Lemma 4. There is an indirect effect as well that moves in the same direction: Increases in x^* in still earlier periods will increase the priors at time t that A is good, and higher priors increase posteriors as shown in Lemma 5. These two effects combine to ensure that higher x^* prior to time t increases the time t priors that A is good. Higher priors at time t, according to Corollary 1, increase the optimal level of conservatism thereby proving the theorem.

The intuition for the proposition is that if the past decision makers who chose to keep A were very progressive then A very likely survived past contests by getting a lot of good signals. This suggests that A is probably good and that even if A appears bad according to contemporaneous information, abandoning A would be a mistake unless switching costs are very low.

For ease of exposition in the proof, avoiding extra subscripts, the proposition is stated assuming that past decision makers had the same level of conservatism. The reader can check that levels can differ and that increasing, or decreasing, the level of conservatism of a single past decision maker rather than all at once leads to the same results. The situation would be more complex in a model where signal strength varied, though it would continue to be the case that progressive decision makers keeping a policy for a long time is a very strong signal that a policy is good. [7]

This proposition explains why switcher's curse became so severe at Company X over time. Because Company X was run by entrepreneurs, they were extremely progressive, looking at all policies with fresh eyes and always

7. In a model where signal strength varied, immediately after a policy A were chosen, the probability that it was good would depend upon the strength of the signal saying it was good. A more conservative decision maker would tend to switch only after seeing very strong signals, so immediately after a switch, a new policy is more apt to be good the more conservative was the person deciding to switch. From that point forward, progressive decision makers deciding to keep the policy will most quickly and strongly update the priors that A is good. Thus, *conservative* decision makers switching to a policy tends to suggest it is good, while *progressive* ones maintaining the policy tends to suggest it was good. It is a reasonable hypothesis that over a sufficiently long period of time, the latter effect would dominate in a model of varying signal strength, but that hypothesis is left for future research.

coming up with new ideas. If a policy survived a long time in that environment it was very likely to be good. Switching a long-standing policy was almost always a mistake, so Company X faced switcher's curse again and again.

5. Applications and Interpretation

Here I interpret the lessons from the model and suggest a variety of applications. Each application is brief and intended to be stimulating and provocative, not a thorough treatment of the issues. Hopefully, the applications suggest avenues for further research, whether empirical, experimental, or theoretical. First, though, a few comments on general interpretation.

5.1. General interpretation

Imperfect information, limited memory, and processing limitations (bounded rationality) all play roles in this article's model. Information is imperfect in that decision makers do not observe the actual quality (i.e. expected payoff) of a given policy; instead, they observe a signal of quality, possibly a realization of payoff. Second, decision makers do not remember prior signals, though they may know how long a policy has been in place. Third, I consider decision makers that follow rule-of-thumb behavioral rules characterizing a degree of conservatism or progressivism. Naive decision makers ignore all information embedded in incumbency which leads them to suffer switcher's curse and to be too progressive as compared with an optimal fully-rational decision maker (who is still handicapped by imperfect information and limited memory).

For simplicity, I have assumed that there is a contest every period, so that a decision maker can infer that a status-quo policy has definitely won prior contests. A richer model would have a probability of a contest at each stage. This complexity would not affect the existing results (switcher's curse would remain and grow with time, but just be smaller).

I have emphasized that naive decision makers will be subject to switcher's curse and that conservatism emerges as a remedy for switcher's curse, or viewed differently a substitute for full memory. Of course, it is entirely possible to be too conservative and the multiperiod model encompasses

that possibility. Conservatism, like most things, has an optimal level and should be practiced in moderation.

The model highlights the value of memory. If decision makers knew why they had kept a policy in the past, they would not be subject to switcher's curse, at least not so long as they were able to process and properly use that information. If processing were overly costly or complex a conservative rule of thumb might still be a useful heuristic.

Environmental changes were also not considered. The model has a stationary environment. If the environment changed radically, so that what was a good policy in the past is no more likely than any other policy to be good today, then conservatism loses all its value and a naive approach is optimal. In less extreme cases, it is reasonable to suppose that the results are moderated but not eliminated. Note also that if a policy has survived many environmental changes, it is reasonable to imagine it is very robust and a good deal of conservatism is called for.

Finally, I note that experimentation played no role in the discussion. Because signals are not remembered, there is actually no role to experiment in the model. Experimentation is an interesting issue when past signals are remembered, an issue that has been explored in the multi-arm bandit literature among others. Likewise if past signals were remembered there would be a real option value of not switching.

5.2. Law and *stare decisis*

Oliver Wendell Holmes (1897, at 469) had a point when he complained of laws that are maintained merely because they have been maintained since Henry IV. Yet, under the common law, and many other systems as well, courts continue to believe in *stare decisis*, which means that prior court decisions are not lightly overturned. Moreover, older decisions, at least if continuously applied, are even less apt to be overturned. This system is perhaps most commonly defended because it provides "uniformity of treatment to litigants, and ... stability and certainty in the law."

In the words of the U.S. Supreme Court in *Kimball v. Marvel* [2015], "[I]t is not alone sufficient that we would decide a case differently now than we did then." Although Kimball argued that [the Brulotte Court] "just made the wrong call," the Court says "[t]hat claim, even if itself dead-right, fails to clear *stare decisis*'s high bar."

The Burkean conservatism that this article develops would assert that however revolting this seemed to Holmes, and even apart from the benefits of predictability and stability, putting weight on the decisions of the past, even if they seem wrong, makes a current court more likely to be right.

To be sure, the legal system differs from our model because decisions are written down and reasons are given for past decisions in court opinions. However, not all the reasons for past decisions are contained in the opinions. Sometimes the most important reasons are absent; indeed, legal realists often argue that opinions are reverse engineered with decisions made first and opinions constructed as *ex post* rationalizations. Regardless, it is clear that *stare decisis* means more than simply considering the arguments made in the past. *Stare decisis* does not require that courts read prior courts' opinions before deciding, but instead directs them to follow prior decisions. Even without *stare decisis*, advocates could bring forward sound arguments found in past cases and current judges could take them for what they are worth independent of precedential value. But *stare decisis* means abiding by past decisions even when they seem wrong, so long as they are not too wrong.

A court is not "inexorably bound by its own precedents," and can overrule them if the court is "clearly convinced that the rule was originally erroneous or is no longer sound because of changed conditions" (See Moore and Oglebay, 1943, pp. 539–40). The model indirectly provides strong support for overruling precedents when the situations that lead to them no longer exist. If the arguments to switch are tied to changed circumstances then switcher's curse is not likely to bite. This point is consistent with Holmes who found adherence to an old rule "still more revolting if the grounds upon which is was laid down have vanished long since, and the rule simply persists from blind imitation of the past." I would, however, put the matter somewhat differently than Holmes. If you are truly blind, then adherence to the past is sound. If on the other hand you know the prior grounds for the rule and know they no longer hold, then abandoning the rule is reasonable.

This article can be understood as formally modeling one reason why precedent matters, and why old precedents carry more weight than newer ones. Likewise, because the article's conservative conclusions evidently

depend upon a stationary environment, the article suggests that those seeking to overturn a precedent should demonstrate that overruling the precedent is complementary with environmental changes.

Of course, different judges give different weight to precedent. The idea that progressive judges who give less weight should follow conservative ones who give more weight and vice-versa is a new claim coming out of the model.

5.3. Administrative Law

Administrative law appears broadly to incorporate the conservative philosophy modeled in this article. For example, "a court is not to substitute its judgment for that of the agency." *Motor Vehicle Mfrs. Assn. of United States, Inc. v. State Farm Mut. Automobile Ins. Co.*, 463 U.S. 29, 43, 103. *Chevron* deference is a form of conservatism in that a thumb is on the scale in favor of agency judgments and these judgments are not revisited with fresh eyes.

A recent decision by the Supreme Court in *FCC v. Fox Television*, 129 S. Ct. 1800 (2009), however, runs counter to what this article advocates. The FCC had changed its rules for what constituted indecent broadcasting over the radio waves. The Second Circuit overruled the Commission relying in "in part on Circuit precedent" and on the Circuit's interpretation of Supreme Court precedent in *State Farm* that required "a more substantial explanation for agency action that changes prior policy," including making clear "why the original reasons for adopting the [displaced] rule or policy are no longer dispositive" *ibid.* at 1810. The Supreme Court, in an opinion by Scalia, overturned the Second Circuit's conservative rule that guarded against switcher's curse with respect to agency decisions. Scalia argued that the FCC did not need to provide "a more detailed justification than what would suffice for a new policy created on a blank slate."

Breyer dissented in *Fox Television* joined by Ginsberg, Stevens, and Souter. Breyer championed a heightened burden on agencies that change policy. Breyer argued that the FCC had not met its burden to change a rule under *State Farm* because the FCC's explanation of its policy change largely discussed factors "well-known to it the first time around" which provide "no significant justification for a *change* in policy." Breyer's view of the FCC's

change of position is similar to his view of the Supreme Court's change of position on vertical price fixing in *Leegin Creative Leather Products v. PSKS*, 551 U.S. 877 (2007). In *Leegin*, Breyer, again joined by Ginsberg, Stevens, and Souter, argued that the court was unwise to overturn a 100-year old precedent against vertical price fixing given that all the majority's arguments could as easily have been made 30 or 40 years ago and there had been no significant new learning that justified a change of Supreme Court antitrust policy.

This article's arguments favor the Breyer–Ginsberg–Stevens–Souter conservative philosophy over their progressive brethren.

5.4. Venture Capital

Venture capitalists get a lot of proposals. How should they identify those with potential? A common first question is whether there is competition already. Venture capitalists seek green fields, not crowded fields.

If there is no competition, a good follow-up is: "Why is there no competition?" A strong version of conservatism would say that if there is no competition, then it is a bad idea because there must be good reasons that no one has entered or that those who entered exited.

A more nuanced follow-up would ask if anything dramatic changed that makes this business possible now where it would have been impossible a decade or two earlier. If the business proposal would have been equally possible a decade earlier, then conservatism suggests putting it at the bottom of the pile. After, all, a great deal of negative information is contained in the fact that no firm in a related business has started this business already, or started it and survived.

Consider Netflix. Netflix proposed in 1998 to distribute rental movies by mail. There was no competition so it passed the venture capitalist's first screen. However, neither mail nor the movie rental business were new in 1998, which might suggest this was a bad idea according to the model here. The absence of competition at that point was a very bad signal about the idea. But let us consider the nuanced question.

The emerging DVD standard made movies much cheaper to mail than the old bulky VHS tapes. Moreover, the internet allowed people, even with the low bandwidth available in those days, to easily choose a movie, store

and update a ranked list, and communicate their orders to Netflix. This made Netflix a valuable investment in 1998, even though a pure mail order movie rental business would have failed in the 1980s. Prophesies of future broadband and future streaming made Netflix a golden opportunity.

The point is that environmental changes meant that Netflix could not have been successful in the past, but could be successful in 1998. In the face of environmental changes that were complementary with the idea, there was no reason to be conservative. And, progressivism paid off big.

5.5. Separation of Powers

Although America's founders were revolutionary and sought fundamental change, they had the prescience to set up a system of government with separation of powers. The consequence is that it is frequently extremely hard to get anything changed. The system is conservative.

The difficulty of achieving change in Washington D.C. is supremely frustrating to anyone who works there and can lead Presidents to dream of a parlimentary system where the prime minister and ruling legislative party are unified. Even when both parties of Congress and the President manage to agree on new legislation, the courts may declare it unconstitutional. Yet, despite its obvious frustrations, according to the model of this article, a conservative system has virtues. Perhaps it is good that little gets done. Switcher's curse might be worse.

5.6. Evolution of a Firm

A new firm needs to create a business model and many policies to effectuate that model. Likely enough, the firm will not get anything right on its first try. Conventional wisdom has it that the business plan will itself survive (at most) until it hits its first customer. And, this is much as it should be according to the model: Initially, the model suggests that the firm should be relatively flexible and progressive. If an alternative business model appears better, why not switch? If someone suggests abandoning cubicles in favor of an open floor plan, and the benefits appear to justify the costs of switching, by all means switch. The mindset of the entrepreneur, which is naturally progressive, is well suited to managing such a firm.

Eventually, though, the business model and policies will have longevity. Longevity means that they will have survived many past contests with alternative policies. Unless times have changed, their age should not stand against them, but rather in their favor. There are reasons, and likely sound reasons, that the business model or operating policies survived prior contests and are the way they are. These reasons should be respected even if they are forgotten. One way to respect them is to require as in the model that switching costs be very low before switching. Another would be to require that a new alternative policy appear extremely good (in a model with variable goodness) or be very sure to be good (in a model where a more precise signal of E might be possible).

One benefit of bringing in senior managers later on in a firm's evolution could be that they are naturally more conservative.[8]

5.7. The New CEO and the New Employee

A new CEO comes to a firm. Should he fire the senior management team? The answer is: it depends.

No doubt the new CEO will have some impulse to make changes and prove himself to the Board, and no doubt he would like to have a new team loyal to him, but firing the senior managers could subject the new CEO to switcher's curse.

If the firm has been successful, the existing management team is of long standing, and the CEO retired because she was getting older, the new CEO should proceed with caution. The existing team knows a lot about why the firm came to do things the way it does. In the model, the existing team will remember many of the rationales behind incumbent policies (older signals). That will no doubt help him avoid switcher's curse as one cause of switcher's

8. It is common in Silicon Valley to bring in older more experienced, and likely more conservative, managers to run a company after a point. "Once a company takes off, the next round of staffing often involves bringing in people with experience – 'grown ups' as they are sometimes called – who know how to manage a maturing business ... " (Bernstein, 2015). As one Example, Eric Schmidt (age 46 years) was brought in to be CEO of Google in 2001 succeeding Larry Page (age 28 years), who continued with the firm, serving as President under Schmidt. See http://googlepress.blogspot.com/2001/08/google-names-dr-eric-schmidt-chief.html.

curse is lack of memory. As long as the CEO respects their judgment and does not push too hard for change he could avoid a lot of switcher's curse.

On the other hand, suppose that the old CEO was extremely conservative and fired by the Board because she refused to adapt to an evolving world. If she surrounded herself with extremely conservative managers, perhaps it is time for them to go. But if the existing managers are prudent people who have been frustrated by the old CEO's absolute refusal to make any changes, the existing managers may be the perfect partners for the new CEO, as their memories will help her sort out the good incumbent policies from the bad ones.

Similar issues arise with any new employee. The employee will come with fresh eyes, but that is not necessarily a good thing, however. After all, a lack of institutional memory will subject him to switcher's curse if he is not conservative by nature. So when the employee comes to his boss and suggests a change in the first week on the job, a wise reply would be: "wait a year." Watch and consider over the next year how things would be with the alternative policy. If it still seems a good idea, then it is worth considering. This is the policy Company X came to after leniently letting many new employees pursue their ideas.

5.8. Cycles of Policies and Cycles of Metapolicies like Conservatism itself

Political cycles are common with conservatives followed by progressives, and in turn progressives followed by conservatives. This article says such cycles of metapolicy may be sensible at least in the following sense. Once one has a long sequence of conservatives, it is likely time for a progressive. Likewise, once one has a long sequence of progressives, it is likely time for a conservative. Similarly, if the judiciary has been dominated by conservatives and strict adherants to *stare decisis*, then it may be time for some activist judges.

A different kind of cycle is when policy A switches to E and then back to A. This is a policy cycle rather than a metapolicy cycle. Such policy cycles could easily result from switcher's curse, limited memory, and decision makers who are too progressive. Companies, universities, and other organizations may go through cycles of centralization and decentralization. When purchasing is decentralized to departments, for example, eventually

someone may observe that costs could be reduced with centralized purchasing. After a period of centralized purchasing, departments may rebel because they perceive the centralized purchasing services as too slow or unresponsive to ideosyncratic needs of departments. It is possible, of course, that such changes reflect optimal adjustment to changed technology or circumstances. Frequently, though they are like a dog chasing its tail, but more costly.

Company X recently almost suffered a cycle but was saved by the conservatism it had adopted after suffering switcher's curse. Four years ago, it had formed an outreach group, whose purpose was to evangelize for its software-as-a-service (SAAS) product. The outreach group wrote blogs, newsletters, and conducted webinars, each one-to-many marketing activities. It also did certain one-to-one customer service activities for individual customers. Recently, the CEO contemplated an internal reorganization in which he put all the one-to-one services into another group, the customer support group, that specialized in one-to-one interactions. The rationale for the switch seemed clear cut. Having one group do all one-to-one activities and another do only one-to-many ones was logical and efficient, and having each customer have a single point of contact would minimize confusion.

In the end, though, the CEO decided to stick with the status quo even though he admitted he had no reason to do so other than conservatism. Only later, long after he decided to maintain the status quo, did the CEO remember a key virtue of the existing structure, a virtue that was part of why he himself had set things up this way years before. Only through their one-on-one interactions with live customers, helping them with concrete issues, could the outreach group gain the texture and wisdom it needed to write meaningful one-to-many marketing materials. Without that experience, the marketing materials would be generic and off-point marketese. Had the CEO switched, he would eventually have switched back. Company X's CEO was saved by his conservative management philosophy from entering a wasteful policy cycle. He developed the conservative philosophy after suffering switcher's curse one too many times.

5.9. Divorce

Consider what this model has to say about the following situation.

Baseline facts. Harold has been married ten years to his wife, Wilma. Harold is miserable in his marriage, which makes him just plain miserable. He cannot remember why he married Wilma and cannot see any good reason to stay in this marriage.

5.9.1. *Variation 1* Assume also that Harold is a level-headed rational decision maker. This article suggests that Harold should likely not leave his wife. If it had always been this bad he would *rationally* have probably left one or two years into his marriage. So there must be good times Harold has simply forgotten. This article's advice to Harold is: "Hang in there Harold. Probably, you and Wilma are at each other's throats now only because each of you has unusual stress at work, and the baby has been sick for so long, keeping you sleepless. This too shall pass."

5.9.2. *Variation 2* Assume instead that Harold tends to be impulsive and confident when it comes to decision making.

Now the inference above is much stronger. If the past had all been as bad as the present, he would surely have left her long ago given how impulsive he is and and how confident he is in his opinions.

5.9.3. *Variation 3* Assume finally that Harold is extremely conservative and cautious by nature and avoids big changes.

Sadly, in this case, this article implies that it could easily be that the past of this marriage is like the present. Miserable. Why did Harold stick with it so long? It takes an avalanche of negative signals to get an extremely conservative person to make a change. If he forgets past signals, he may never leave unless he changes his decision-making philosophy. Perhaps it is time to do that.

6. Conclusions

Many expressions capture the essential intuitions of this article. We say, "If it ain't broke, don't fix it;" "The grass is always greener;" and "The devil you know is better than the devil you don't." Each maxim became a mantra at Company X after switcher's curse had bitten enough times.

These age-old expressions suggest that the wisdom of conservatism has long been recognized, no doubt even before Edmund Burke. They also suggest that many people have overly progressive tendencies and need urging to be more conservative, else there would be no need for the maxims.

This article has studied problems in which a decision maker decides between a status-quo policy and another policy. Naively looking at the two policies with fresh eyes will lead the decision maker to switch too often, suffering what I call switcher's curse. Switcher's curse arises if the status-quo policy may have competed against (and won against) other policies before to achieve its incumbent status, and if there was a tendency for the better policy to win the prior contest(s). In such circumstance, if a decision maker does not give adequate weight to the positive information embedded in incumbency she will suffer switcher's curse. Very progressive people, such as the innovators who started Company X are likely to do just that. Limited memory and either prior neglect or the availability heuristic, two widely documented psychological phenomena, can lead agents to be naive and overly progressiveness so that they suffer switcher's curse in our model.

I show that conservatism is a rule-of-thumb antidote for switcher's curse. Decision-makers should put a thumb on the scale in favor of a status-quo policy. They should be more conservative the longer a policy has been around. I also show that if past decision makers were progressive, then it pays to be especially conservative today. If past decision makers were extremely conservative, progressivism is called for today.

One lesson of the analysis is that what psychologists describe as status-quo bias (see, e.g. Burmeister and Schade, 2007; Kahneman et al., 1991) is often not a bias but a heuristic that can compensate for these other biases. The analysis also points to what may be the best arguments to overturn a long-standing policy. Instead of just arguing that some other policy is better than the incumbent policy, it is more convincing to show that the other policy is better now, but only becasue of environmental changes, and that absent those changes, the incumbent policy would have been better in the past. Such an argument strikes at the heart of Burkean conservatism and truly justifies change.

Appendix A. Proofs

A.1. Proof of Lemma 1

For brevity, we write "*A won*" in place of "*A won at time* 0." Bayes's rule implies:

$$Pr(A = G \mid A \text{ won}) = \frac{Pr(A \text{ won} | A = G)}{Pr(A \text{ won})} Pr(A = G). \tag{A1}$$

The numerator can be calculated as follows:

$$\begin{aligned} Pr(A \text{ won} | A = G) = {} & 1/2 Pr(a_0 = g | A = G) Pr(d_0 = g) \\ & + 1/2 Pr(a_0 = b | A = G) Pr(d_0 = b) \\ & + Pr(a_0 = g | A = G) Pr(d_0 = b). \end{aligned} \tag{A2}$$

To simplify the expression, let $w = Pr(a_0 = g \mid A = G)$ and $z = Pr(d_0 = b)$. We then have

$$\begin{aligned} Pr(A \text{ won} | \, A = G) &= 1/2 \; w(1 - z) + 1/2 \; (1 - w)z + wz \\ &= w/2 + z/2 \\ &= 1/2 \; Pr(a_0 = g \mid A = G) + 1/2 \; Pr(d_0 = b). \end{aligned} \tag{A3}$$

A similar derivation shows that

$$Pr(A \text{ won}) = 1/2 \; Pr(a_0 = g) + 1/2 \; Pr(d_0 = b). \tag{A4}$$

Bayes equation can then be rewritten as

$$Pr(A = G \mid A \text{ won}) = \frac{Pr(a_0 = g \mid A = G) + Pr(d_0 = b)}{Pr(a_0 = g) + Pr(d_0 = b)} Pr(A = G). \tag{A5}$$

Since $Pr(a_0 = g \mid A = G) > Pr(a_0 = g)$, this establishes that

$$Pr(A = G \mid A \text{ won}) > Pr(A = G). \tag{A6}$$

■

A.2. Proof of Observation 1

Define $p_A \equiv Pr(A = G|A \text{ won at time } 0)$, and $p_E \equiv p$. These are the probabilities that A and E respectively are good immediately prior to receiving the time 1 signals a and e of their quality.

The naive and conservative decision maker only disagree when signals are equal.

Consider first the case $a = b, e = b$.

Observe that

$$\begin{aligned}\Pr(A = G \mid a = b) &= \frac{\Pr(a = b \mid A = G)p_A}{\Pr(a = b)} \\ &= \frac{(1-s)p_A}{(1-s)p_A + s(1-p_A)} \\ &= \frac{1-s}{1-s+s\frac{1-p_A}{p_A}} \end{aligned} \tag{A7}$$

which is increasing in p_A. By symmetry and the fact that $p_A > p_E = p$, we therefore have

$$\Pr(A = G \mid a = b) > \Pr(E = G \mid e = b), \tag{A8}$$

so that being conservative and keeping A is better in this case.

Consider now the case: $a = g, e = g$**.**

The logic in this case is similar. Observe that

$$\begin{aligned}\Pr(A = G \mid a = g) &= \frac{\Pr(a = g \mid A = G)p_A}{\Pr(a = g)} \\ &= \frac{sp_A}{sp_A + (1-s)(1-p_A)} \\ &= \frac{s}{s+(1-s)\frac{1-p_A}{p_A}} \end{aligned} \tag{A9}$$

which is increasing in p_A. By symmetry and the fact that $p_A > p_E = p$, we therefore have

$$\Pr(A = G \mid a = g) > \Pr(E = G \mid e = g). \tag{A10}$$

This implies that $\Pr(A = G \mid a = g, e = g) > \Pr(E = G \mid a = g, e = g)$, so that being conservative and keeping A is better in this case as well. ■

A.3. Proof of Lemma 3

We abbreviate "*A wins at t*" with "*A wins*." According to Bayes's rule

$$\Pr(A = G \mid A \text{ wins}) = \frac{\Pr(A \text{ wins} \mid A = G) p_A}{\Pr(A \text{ wins})}. \tag{A11}$$

Let us unpack the updating ratio $u \equiv \frac{\Pr(A \text{ wins}|A=G)}{\Pr(A \text{ wins})}$. We can rewrite the numerator as follows:

$$\begin{aligned}\Pr(A \text{ wins} \mid A = G) &= 1 - \Pr(A \text{ loses}|A = G) \\ &= 1 - \Pr(a = b, e = g, x < x^* \mid A = G) \\ &= 1 - Pr(a = b \mid A = G) Pr(e = g) \frac{x*}{x_{max}}. \end{aligned} \tag{A12}$$

In words, A wins only if it did not lose and it loses only if (i) A appears bad (i.e. if $a = b$), (ii) E appears good (i.e. $e = g$), and (iii) switching costs are below the switching cost threshold (i.e. $x < x^*$).

Similarly, we can rewrite the denominator as follows:

$$\begin{aligned}\Pr(A \text{ wins}) &= 1 - \Pr(A \text{ loses}) \\ &= 1 - \Pr(a = b) \Pr(e = g) \frac{x^*}{x_{max}}. \end{aligned} \tag{A13}$$

Comparing (A12) and (A13), and observing that $\Pr(a = b \mid A = G) < \Pr(a = b)$, we find that whenever $x^* > 0$ it follows that, $u > 1$. In contrast, when $x^* = 0$, then expression (A12) equals expression (A13) and $u = 1$.■

A.4. Proof of Lemma 4

Recall that

$$\begin{aligned}\Pr(A = G \mid A \text{ wins}) &= \frac{\Pr(A \text{ wins} \mid A = G) p_A}{\Pr(A \text{ wins})} \\ &\equiv u\, p_A. \end{aligned} \tag{A14}$$

We will now show that

$$\frac{\partial u}{\partial x^*} > 0. \tag{A15}$$

Observe that (explanation follows the equations)

$$\begin{aligned} sgn\left(\frac{\partial u}{\partial x^*}\right) &= sgn\left(\frac{\partial \frac{\Pr(A \text{ wins} \mid A=G)}{\Pr(A \text{ wins})}}{\partial x^*}\right) \\ &= sgn[-\Pr(A \text{ wins})\Pr(a=b \mid A=G)\Pr(e=g)\frac{1}{x_{max}} \\ &\qquad + \Pr(A \text{ wins} \mid A=G)\Pr(a=b)\Pr(e=g)\frac{1}{x_{max}}] \\ &= sgn[-\Pr(A \text{ wins})\Pr(a=b \mid A=G) \\ &\qquad + \Pr(A \text{ wins} \mid A=G)\Pr(a=b)] \\ &= 1 \end{aligned} \tag{A16}$$

where $sgn(w)$ takes the value "1" when w is positive and "-1" when w is negative.

The second equality above follows by (i) substituting into the right-hand expression of the top line using Equations (A13) and (A12), (ii) evaluating the derivative using the quotient rule, and then (iii) eliminating the squared term in the denominator, which is of necessity positive.

The third inequality follows by eliminating the positive quantities $\Pr(e = g)$ and $\frac{1}{x_{max}}$.

The fourth inequality follows because

$$\Pr(a=b) > \Pr(a=b \mid A=G) \tag{A17}$$

and

$$\Pr(A \text{ wins} \mid A=G) > \Pr(A \text{ wins}). \tag{A18}$$

■

A.5. Proof of Lemma 5

$$\begin{aligned} &\Pr(A=G \mid A \text{ wins}; p_A) \\ &\quad = \frac{\Pr(A \text{ wins} \mid A=G)p_A}{\Pr(A \text{ wins})} \\ &\quad = \frac{\Pr(A \text{ wins} \mid A=G)p_A}{\Pr(A \text{ wins} \mid A=G)p_A + \Pr(A \text{ wins} \mid A=B)(1-p_A)}. \end{aligned} \tag{A19}$$

Hence

$$sgn(\frac{\partial \Pr(A = G \mid A\ wins)}{\partial p_A}) = sgn(\frac{\partial}{\partial p_A}\left(\frac{1}{1 + \frac{(1-p_A)\Pr(A\ wins|A=B)}{p_A \Pr(A\ wins|A=G)}}\right)$$
$$= -sgn(\frac{\partial \frac{(1-p_A)}{p_A}}{\partial p_A})$$
$$= 1. \tag{A20}$$

■

A.6. Proof of Proposition 3

Suppose that A is a status-quo policy that has been in place for T periods at time t.

For any time m, let p_A^m denote the probability that $A = G$ immediately prior to the time m contest and prior to receiving any time m signal about A.

Updating is special in the first period for policy A and so our proof will depend upon whether policy A began at time 0 or at a later time.

Case #1: A began at time 0.

We suppose that A began at time 0 and had probability p of being good and $1 - p$ of being bad, $p \in (0, 1)$. At time 1, A faced its first contest. The prior probability that A is good at time 1 equals p: that is,

$$p_A^1 = p. \tag{A21}$$

Let x^* be the decision rule applied at times $m = 1, ..., T - 1$. Observe that p_A^1 does not depend upon x^*, directly or indirectly, so that

$$\frac{\mathrm{d}p_A^1}{\mathrm{d}x^*} = 0. \tag{A22}$$

.

Let $g(p_A^m, x^*) \equiv \Pr(A = G \mid A\ wins, p_A^m, x^*)$ be a function that maps priors p_A^m at any given time m for a status-quo policy A into a posterior that $A = G$ given that A wins at time m under the switching threshold x^*.

An equation of motion of p_A^m is then given by

$$p_A^m = g(p_A^{m-1}, x^*), \text{ for } m = 2, ..., T. \tag{A23}$$

We can now show by induction that

$$\frac{\mathrm{d}p_A^m}{\mathrm{d}x^*} > 0, m = 2, ..., T. \tag{A24}$$

First, we establish inequality (A24) for m=2. Observe that

$$\frac{\mathrm{d}p_A^2}{\mathrm{d}x^*} = g_1 \frac{\mathrm{d}p_A^1}{\mathrm{d}x^*} + g_2, \tag{A25}$$

where g_1 and g_2 denote the first and second partial derivatives of g. Since $\frac{\mathrm{d}p_A^1}{\mathrm{d}x^*} = 0$, we can simplify the above to yield:

$$\frac{\mathrm{d}p_A^2}{\mathrm{d}x^*} = g_2. \tag{A26}$$

Observe that $g_2 > 0$ because of Lemma 4. This establishes inequality (A24) for the case $m = 2$. To complete our inductive proof, assume that

$$\frac{\mathrm{d}p_A^{k-1}}{\mathrm{d}x^*} > 0. \tag{A27}$$

Differentiating (A23) yields

$$\frac{\mathrm{d}p_A^k}{\mathrm{d}x^*} = g_1 \frac{\mathrm{d}p_A^{k-1}}{\mathrm{d}x^*} + g_2. \tag{A28}$$

Observe now that $\frac{\mathrm{d}p_A^k}{\mathrm{d}x^*} > 0$ because

- $g_2 > 0$ (see Lemma 4)
- $g_1 > 0$ (see Lemma 5), and
- the inductive hypothesis that $\frac{\mathrm{d}p_A^{k-1}}{\mathrm{d}x^*} > 0$.

This establishes inequality (A24), and in particular that $\frac{\mathrm{d}p_A^T}{\mathrm{d}x^*} > 0$. Now Corollary 1 proves the proposition for case # 1.

Case #2. A is first chosen at some time $m > 0$.

The only meaningful difference from case #1 is that the formula for p_A^{m+1} differs from what the formula was for p_A^1 above. We need to ensure that

$$\frac{\mathrm{d}p_A^{m+1}}{\mathrm{d}x^*} = 0. \tag{A29}$$

Let policy F be the status-quo policy at time m, let p_F be the prior that F is good, and let f be the signal of $F's$ quality.

$$\begin{aligned} p_A^{m+1} &= \Pr(A = G \mid A \textit{ wins at time } m; x^*, p_F) \\ &= \frac{\Pr(A \textit{ wins at time } m \mid A = G; p_F, x^*)p}{\Pr(A \textit{ wins at time } m;\ p_F, x^*)} \\ &= \frac{\Pr(a = g, f = b, x < x^* \mid A = G; p_F, x^*)p}{\Pr(a = g, f = b, x < x^*;\ p_F, x^*)} \\ &= \frac{\Pr(a = g \mid A = G)p}{\Pr(a = g)}. \end{aligned} \tag{A30}$$

The last equality follows because a^m, f, x^m are independent and f and x^m are unaffected by whether A is good or bad. This implies that

$$\frac{\mathrm{d}p_A^{m+1}}{\mathrm{d}x^*} = 0. \tag{A31}$$

The remainder of the proof in case #2 follows that of case #1. ■

Appendix B. Company X: A Case Study in Switcher's Curse and Conservatism

As mentioned in the introduction, Company X began with innovations and an entrepreneurial mindset. For a while all went well and changes were mainly for the better, but eventually the firm began to suffer switcher's curse. Changes often went hay wire and had unintended consequences. Frequently the virtues of an existing policy had been forgotten and only became evident after a switch.

B.1. Login Screens

A simple example is login screens. Company X developed its own software. When Company X began, the internet was young and login's were

especially frustrating to users. Many users turned off upon seeing a login screen. As usage meant success, the company had a strong interest in minimizing login hassles. Early on, the company had a smart innovation (independently invented by many firms) to put a "Remember me" box under a login so that users could click it and never have to login in again from that computer (so long as the user's browser allowed the system to place a cookie on his machine). This was a valuable early switch in policy. Unfortunately, most users did not bother clicking that box so Company X continued to worry about how to make logging in easier.

Eventually one executive had an epiphany. Why not make the "Remember me" box be default "on" so that a user needed to click it to turn it off? Defaults matter a lot (see Madrian and Shea, 2001; and Sunstein and Thaler, 2003), so this switch was expected to dramatically increase usage and to be well worth the effort of reprogramming legacy software.

Although the switch seemed a no-brainer, it went wrong. Librarians, who were Company X's paying customers and are especially concerned with privacy issues, complained vociferously, because some patrons when using public computers did not uncheck the box so that other patrons later found themselves in another person's account when sitting down at the same commonly used machine. The complaints were sufficient that Company X had to switch back quickly; again this required programming effort because the old code needed to be regression tested again to make sure it was compatible with other independent changes that were concurrently made.[9]

The reader may wonder if this was a good gamble that went awry for an unforeseeable reason. In a sense that is right, but here is the catch. This same executive had this same idea a couple of years earlier and the firm had tried it with the same "unforeseeable" result. In fact, every year or two the executive regularly had the same epiphany. Usually he got lucky and was saved by the memory of a long-standing employee. Company X's managers believe they only made the mistake of switching the login default twice before, but cannot be sure.

9. According to Wikipedia (downloaded 2014.08.28 and available from author), regression testing is needed because "Experience has shown that as software is fixed, emergence of new faults and/or re-emergence of old faults is quite common."

Eventually, the once progressive executive became more conservative and even if no one could remember the reason that "default off" was best, trusted that if it were not, the firm would have switched long ago.

Company X may seem quite a mess. The question, though, is "compared to what?" Compared to perfection, surely. But, despite its bobbles, Company X has grown and thrived in a difficult market. Other companies must suffer similar issues, or Company X could not thrive. The model in this article is intended to explain why firms can have such troubles. A few more examples will help to motivate the model.

B.2. Company Y's Switch in Sales Strategy

Company X licensed its software but faced reputational barriers as an early software-as-a-service company. It therefore entered a partnership with Company Y, a large established company, as the exclusive distributor of its software service. Company Y had a large salesforce with hundreds of salespeople around the world selling to libraries, but its reputation was more important to the partnership. After consideration Company Y decided not to use this established organization to sell Company X's software service and instead started a specialized sales force. The reasons were sound: Company Y's products were content, not software, and X's software product would also have been lost among much higher ticket items in the portfolio of the main sales force. Y therefore started a specialized sales force which did quite well arguably proving the decision sound.

After management changes at Company Y, Company Y switched tacts and decided to expand sales not by hiring more specialized sales people but by jettisoning the specialized sales model and using Y's main sales force. Sales slowed rather than rose, tensions ensued between X and Y, and the partnership ended. Company X bought the rights back to sell its software services, returned to a specialized sales force and sales grew strongly.

In short, Company Y, after management changes, forgot the reasons for its original policy and suffered switcher's curse. It lost a great deal of money from this mistake.

B.3. An Outside CEO

A few years later, Company X hired an experienced successful outside CEO from the same industry. This executive made a great many changes at

Company X. Most of these were to conform with her own experience at other companies. In a sense, one might think that these changes were conservative in that they conformed to outside norms. But they did not work. Unnoticed was that Company X was ideosyncratic and its policies fit together and worked tolerably well. Changing to industry norms was an option that had always been available to X's managers and a student of this article's model might surmise that perhaps they had not switched for good reason. One example: Company X had found it difficult to sell subscriptions to new content and therefore early on decided to sell subscriptions to its content together as a bundle. X's bundled subscription system worked well and so X kept the system. Unlike many other firms in the industry, however, as it added new content to its portfolio it added that content to the bundled subscriptions of existing customers and raised the price accordingly.

The new CEO decided to switch from this policy to those of her previous firm's, which created a new content bundle each year without upgrading old subscriptions. There were two problems. The first was one of organization. It was difficult for a small firm to build new systems to keep track of different customers subscribing to different packages—the 2007, 2008, or 2009 bundle of content, depending upon the year of original purchase and whether the customer upgraded to a subsequent year—and to build systems to provide electronic access to arbitrary collections of content. The bigger problem was that content could not be sold effectively a la carte by Company X (which is what lead Company X to bundle in the first place) and few existing customers wanted to upgrade their packages to include new content when they now had the option not to. Company X consequently suffered switcher's curse and lost substantial revenue and profit.

B.4. The Switches of Company Z, an Acquirer

Company X eventually sold its content business. Company Z, a much larger firm in the industry bought it. For six months all was well. Then the acquirer changed a number of policies, including ceasing to use Company X's software. In doing so, it failed to appreciate that Company X's policies were not chosen randomly but refined over time and kept when they worked. It failed to appreciate that X's managers could at any time have switched to industry standard practices or software if those would have made sense. It

is not that industry standard practices are wrong, nor that X's policies were, but Company X had a niche and its policies worked for that niche. The result of the acquirer's changes was that many critical content producers quit. Company Z suffered switcher's curse because it did not put sufficient weight on the idea that Company X's policies were there for a sound reason, even if that reason was not apparent to Company Z.

All these examples are of course open to other interpretations, like any case study. The interpretations above, however, serve to motivate the model in this article and to illustrate the ideas of this article. The reader is encouraged to check this article's ideas against her own experience, or even better, to develop empirical tests of these ideas.

References

Anand, Vikas, Charles C. Manz, and William H. Glick. 1998. "An Organizational Memory Approach to Information Management," 23 *Academy of Management Review* 796–809.

Baliga, Sandeep, and Jeffrey Ely. 2011. "Mnemonics: The Sunk Cost Fallacy as a Memory Kludge," 3 *American Economic Journal: Microeconomics* 35–67.

Bar-Hillel, Maya. 1980. "The Base-rate Fallacy in Probability Judgments," 44 *Acta Psychologica* 211–233.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch. 1992. "A Theory of Fads Fashion, Custom, and Cultural Change as Informational Cascades," 100 *Journal of Political Economy* 992–1026.

Bernstein, Mark F. 2015. "Over the Hill in Silicon Valley?" *Princeton Alumni Weekly*. March 18, 24–27.

Burke, Edmund. 1790. *Reflections on the Revolution in France*. London: J. Dodsley.

Burmeister, Katrin, and Christian Schade. 2007. "Are Entrepreneurs' Decisions More Biased? An Experimental Investigation of the Susceptibility to Status Quo Bias," 22 *Journal of Business Venturing* 340–362.

Dertouzos, Michael, Richard K. Lester, and Robert M. Solow. 1989. *Made in America, Regaining the Productive Edge*. Cambridge, MA: MIT Press.

Hirshleifer, David and Ivo Welch. 2002. "An Economic Approach to the Psychology of Change: Amnesia, Inertia, and Impulsiveness," 11 *Journal of Economics and Management Strategy* 379–421.

Holmes, Oliver Wendell. 1897. "The Path of the Law," 10 *Harvard Law Review* 457–478.

Kahneman, Daniel, and Amos Tversky. 1973. "On the Psychology of Prediction," 80 *Psychological Review* 237–251.

Kahneman, Daniel, and Amos Tversky. 1979. "Intuitive Prediction: Biases and Corrective Procedures" in S. Makridakis, and S.C. Wheelwright, eds., *Forecasting: TIMS Studies in the Management Sciences*, Vol. 12. Amsterdam: North-Holland.

Kahneman, Daniel, and Amos Tversky. 1982. "Evidential Impact of Base Rates" in Daniel Kahneman, Paul Slovic, and Amos Tversky, eds., *Judgment under Uncertainty: Heuristics and Biases*, 153–160. New York: Cambridge University Press.

Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler. 1991. "Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias," 5 *Journal of Economic Perspectives* 193–206.

Madrian, Brigitte C. and Dennis F. Shea. 2001. "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior," 116 *Quarterly Journal of Economics* 1149–87.

Moore, James, and Robert Oglebay. 1943. "The Supreme Court, Stare Decisis and Law of the Case," 21 *Texas Law Review* 514–53.

Porter, Maureen, and Sally McIntyre. 1984. "What is, must be Best: a Research Note on Conservative or Deferential Responses to Antenatal Care Provision," 19 *Social Science & Medicine* 1197–200.

Prendergast, Canice, and Lars Stole. 1996. "Impetuous Youngsters and Jaded Old-Timers: Acquiring a Reputation for Learning," 104 *Journal of Political Economy* 1105–34.

Rasmusen, Eric. 1992. "Managerial Conservatism and Rational Information Acquisition," 1 *Journal of Economics and Management Strategy* 175–201.

Samuelson, William, and Richard Zeckhauser. 1988. "Status quo bias in Decision Making," 1 *Journal of Risk and Uncertainty* 7–59.

Thaler, Richard, and Cass Sunstein. 2003. "Libertarian Paternalism," 93 *American Economic Review* 175–9.

Tversky, Amos, and Daniel Kahneman. 1973. "Availability: A Heuristic for Judging Frequency and Probability," 5 *Cognitive Psychology* 207–32.

Zwiebel, Jeffrey. 1995. "Corporate Conservatism and Relative Compensation," 103 *Journal of Political Economy* 1–25.