European Journal of Political Economy 5 (1989) 219–227. North-Holland

OPTIMAL FLEXIBILITY OF RULES

The Tale of the Wise Principal and the Naive Agent

Ariel RUBINSTEIN*

The Hebrew University of Jerusalem, Jerusalem, Israel

The governed has to choose an action and his decision is influenced by a rule made by the Governor, who announces it retroactively. The Governor may want to change the rule from time to time because of a change in the state of the world. Changing the rule affects the uncertainty about the correlation between current and past rules. Too much flexibility will create unnecessary uncertainty and too much rigidity will often result in unsuitable rules being in force. The paper studies the properties of the golden mean between these two extremes.

In a typical Stakelberg situation, the *governed* (the agent), g, has to choose an action. His decision is influenced by a rule made by the *Governor*, G, who chooses it so as to maximize a target function, given the pattern of behaviour he assumes for g.

As time passes, some of the conditions of G's problem may change and he may want to alter the rule. However, changing the rule in itself introduces new factors into G's considerations which may restrain him from immediately changing the rule as soon as it ceases to be optimal.

One factor in G's considerations is the administrative costs he incurs whenever the rule is changed, costs which are condusive to a rigid policy. Another factor is possible imperfections in the transmission of the rules from G to g – there are usually delays in the diffusion of a new rule. The more frequent the changes, the more confusion there is about what is the current rule.

These two factors are ignored here; the paper deals with principal-agent systems in which the ruling is done by doing. There is no advance announcement of G's behaviour. In any specific case, G reacts to g's behaviour. The governed does not know G's rule. He constructs expectations about the rule effective at a particular time, taking past rules into account. It is assumed that if a rule has not been changed g believes it will remain that way. However, if the sequence of past rules includes changes, g evaluates the

*I am indebted to James Mirrless, Barry Nalebuff, Joseph Raz, Murat Sertel, and Asher Wolinsky for their comments.

probability of a new change according to the frequency of past changes.

Governors who rule by doing are confronted with a dilemma. On the one hand, it is desirable to change the rule whenever it ceases to be the best. On the other, a change of rule creates uncertainty because individuals are not sure which rule is in force in any specific case. Such uncertainty is undesirable.

A legal system based on a doctrine of precedent is an example of a ruling-by-doing system. My own interest in the subject was aroused by the common explanation of the traditional rigidity of the law by the desire to establish certainty as to the existing law. One scholar – Holdsworth – says that legal systems aim to:

> 'hit the golden mean between too much flexibility and too much rigidity ... the rigidity a legal system must have if it is to possess a definite body of principles, and the flexibility which it must have if it is to adapt itself to the needs of a changing society' [as cited by Cross (1966, p. 118)].

Ruling by doing is not confined to the law. An economic example is a long-term seller-buyer relationship in which the seller quotes the price after the buyer has ordered. Another example is the learning process in which a father transmits signals to his children designed to instruct them on how to behave in different situations.

Before introducing the model, I must admit that it is a somewhat artificial one. In spite of its limitations, however, I believe that it captures a certain limited element of a ruling-by-doing situation.

1. The tale of the naive agent and the wise principal

Once upon a time there was a naive person, g, who lived in a small village. He divided his day between two activities; immemorial village tradition has it that only one of them is the 'right' activity. By himself, g does not know the difference between right and wrong.

By great good fortune there was also a wise person, G, living in the village. The naive g had complete faith in G. Every evening he went to G to be told what the day's right activity had been, and then obtained remuneration, spiritual or otherwise, for that part of the day which he had spent on the activity pronounced by G to have been the right one.

G's judgement derives from deep insight into the state of the world. The state may occasionally change according to a Markov process, known to him. Whenever he announces a new rule, he reveals the state of the world to g, and thereby maintains g's unbounded faith in him.

Both G and g are interested in g spending his whole day in the right activity. However, there is uncertainty about what that will be. G and g may have different risk attitudes. Thus even if G and g coincide in their belief as to whether or not the rule will change, g may still not allocate his time between the two activities in the proportions G thinks preferable. If G changes the rule, g's allocation may not be optimal for G, who might therefore prefer to be rigid and not to adjust the rule to the state of the world.

2. The model

The basic situation

The two players in the model are G (the Governor/Principal) and g (the governed/agent). The governed, g, has to choose a mixture of two activities 1 and 2. Thus he has to choose from among the set $A=\{a_1,a_2)|0\leqq a_1,\ 0\leqq a_2,\ a_1+a_2=1\}$, where a_1 denotes the share of the day devoted by g to activity 1. Time is discrete. At time t $(t=1,2,3,\ldots)$ g chooses action $a^t\in A$. In each period G rules that one of the activities is the right one; his decision at time t is denoted by r^t. Let $R=\{1,2\}$ be the set of possible rules. Denotes $\hat{1}=2$ and $\hat{2}=1$.

The behaviour of g

The governed has a von Neumann–Morgenstern utility function over $A\times R$, of the form

$$v(a,r)=v(a_r).$$

That is, g's utility is a function of the fraction of the day that he has devoted to the activity which is declared the right one. The function v is twice differentiable, strictly increasing, strictly concave, and $v'_+(0)=\infty$. In each period g seeks to maximize his expected utility. He does not know the rule r^t when he chooses a^t. All he knows at time t is what the previous rules $r^1,\ldots,r^{t-1}$ were; he builds his expectations about the next rule on the basis of this knowledge, and evaluates the probability of a change in the rule as the relative frequency of past changes in the rule $r^{t-1}=r$, that is as c_r^t, where

$$c_r^t=\frac{|\{h<t-1\,|\,r^h=r^{t-1}\ \text{ and }\ r^{h+1}\neq r^{t-1}\}|}{|\{h<t-1\,|\,r^h=r^{t-1}\}|}.$$

(If the divisor is 0, define it arbitrarily as 0).

Thus g acts as if he believed that the rule changes by a Markov process. As will appear, g's Markov hypothesis is confirmed. Note that his choice is short-run minded; one plausible interpretation being that the g in the model stands for a sequence of individuals, each of whom makes one decision, knowing the history and taking the same lessons from it.

The state of the world

The problem is that G's utility depends on a random state of the world. Let $W=\{1,2\}$ be the set of possible states. G knows that the state is determined by a Markov chain with a symmetric transition matrix

$$\begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}.$$

In order to express the requirement that today's state is a positive indicator for tomorrow's, assume $p<1/2$. To prevent degeneracy of the problem, assume $p>0$.

The governor's goal

G's is a function of g's choice and of the state of the world. Let

$$u(a,w)=u(a_w)$$

be G's utility where g chooses $a=(a_1,a_2)$ and w is the state. The function u is also assumed to be twice differentiable, strictly increasing and strictly concave, and $u'_+(0)=\infty$. Notice that u is independent of the rule, which excludes buyer–seller relationships in which the seller quotes the price after the buyer has ordered. The function u is only a component of G's maximization. His goal is to maximize the limit of the averages of the single-period utilities, that is G's long-run utility from a sequence of actions (a^t) is

$$\lim_{T\to\infty} \sum_{t=1}^{T} u(a^t,w^t)/T.$$

(We shall look for the strategies by which G induces sequences of actions such that this limit almost certainly exists.)

The governor's set of strategies

A strategy for G is a sequence of decisions. The t-th decision determines the rule in period t as a function of history up to time $t-1$ and of the state of the world at time t. It is assumed that G cannot change the rule r^{t-1} at time t unless $w^t \neq r^{t-1}$. It is well known [see Derman (1970)] that there exists a stationary Markovian optimal strategy. Therefore, the discussion is restricted to stationary mixed strategies (f_1, f_2), with the interpretation that f_r is the probability that if $r^{t-1}=r$ and $w^t \neq r$, G changes the rule and makes $r^t \neq r^{t-1}$.

Note that if G uses a stationary mixed strategy, the rule is a Markovian random variable and g's hypothesis is confirmed.

The problem

The situation at time t is:

	Information	Choice	Utility
G	complete	r^t	$u(a^t, w^t)$
g	$r^1, \ldots r^{t-1}$	a^t	$v(a^t, r^t)$

Thus the problem we are looking into is (1):

$$\max_{0 \leq f_1, f_2 \leq 1} \lim_{T \to \infty} \sum_{t=1}^{T} u(a^t, w^t)/T$$

s.t. $a^t \in A$ maximizes $v(a, r^{t-1})(1-c_q^t) + v(a, \hat{r}^{t-1})c_q^t$ (where $q = r^{t-1}$)

$$\text{prob}\,(r^t = i \,|\, r^{t-1} = j \quad \text{and} \quad w = {}^t h) = \begin{bmatrix} 0 & i \neq j = h \\ 1 & i = j = h \\ f_j & i = h \neq j. \\ 1 - f_j & i = j \neq h \end{bmatrix}$$

Note that $\{r^t\}$ is a sequence of random variables determined by (f_1, f_2). The initial rule r^0 is arbitrary.

Before continuing let me emphasize two assumptions which are embedded in the construction of the model.

(1) The utility functions u and v are functions over two different spaces, $A \times W$ and $A \times R$ respectively. We identify R and W and this enables us to compare G and g in terms of their risk aversion.

(2) The model is symmetric in the states of the world. For g, choosing (a_1, a_2) when the rule is 1, is like choosing (a_2, a_1) when the rule is 2. For G, the choice (a_1, a_2) when the state is 1, is like the choice (a_2, a_1) when the state is 2.

3. An auxiliary problem

Consider the following static principal–agent problem. Assume that the state is w and G believes that the probability of a change is p. Assume that G has the power to determine g's belief about the rule which will be in force. Denote by c the probability that the rule will be $\hat{w}$. G's problem becomes (2):

$$\max_{\substack{0\leq\alpha\leq 1\\0\leq c\leq 1}} (1-p)u(\alpha)+pu(1-\alpha)$$

$$\text{s.t.}\quad \alpha\in\arg\max_{0\leq\alpha\leq 1}(1-c)v(\alpha)+cv(1-\alpha).$$

Such a problem is discussed in Paroush and Rubinstein (1982) and the following proposition is a direct application of the analysis there. Define G as more risk averse than g if there exists a strictly increasing and concave function T such that $u=T(v)$.

Proposition 1. If G is more risk averse than g and if (c^*,α^*) is a solution of (2), $\alpha^*<1$, *and α^* maximizes $(1-p)u(\alpha)+pu(1-\alpha)$; then $c^*\geqq p$.*

Thus, if G is more risk averse than g he will persuade g to assign probability to the less likely state of the world, which is at least as large as the true probability.

4. Results

Proposition 2. Let c^* be a solution of the problem

$$\max_{0\leq c\leq p}[1-s(c)]u[\alpha(c)]+s(c)u[1-\alpha(c)]$$

(3) s.t. $\alpha(c)$ *maximizes* $(1-c)v(\alpha)+cv(1-\alpha)$

$$s(c)=\frac{1}{2}-c\left(\frac{1}{2p}-1\right).$$

Then

(a) $f_1=f_2=f^*$ *is a solution of the original problem* (1), *where f^* satisfies* $c^*=f^*/[2+f^*(1/p-2)]$.

(b) *For the strategy $f_1=f_2=f^*$, c^* is the limit of c_r^t (the frequencies of $r^{t-1}\neq r^t$ where $r^{t-1}=r$)*

(c) $s(c^*)$ *is the limit of the frequencies of $w^t\neq r^{t-1}$.*

The proof of the proposition appears in the appendix.

When g expects the rule to be changed with probability c, he spends a fraction $\alpha(c)$ of his time on the activity pronounced to be the 'right' one the day before. It follows from Proposition 2 that (1) is very close to the auxiliary problem (2). G determines c, the uncertainty about the rule facing g. He does it by manipulating the strategy (f_1, f_2). When $f_1 = f_2 = f$ varies from 0 to 1, c varies from 0 to p.

The differences between (3) and the auxiliary problem (2) are in the target function. In (3) $s(c)$ replaces the p which appears in (2). G weights the utility of the activity r^{t-1} by the probability that $r^{t-1} = w^t$, which differs from the probability that $w^{t-1} = w^t$. The second difference between (3) and (2) is in the restriction that $c \leqq p$.

The target function in (3) may be rewritten as $u_f - u_r$, where $u_f = (1-p)u[\alpha(c)] + pu[1-\alpha(c)]$ and $u_r = (s-p)\{u[\alpha(c)] - u[1-\alpha(c)]\}$. The rigidity loss, u_r, is inflicted on G by the fact that with frequency $(s-p) \geqq 0$, G refrains from changing the rule even though the previous rule is not suitable in the current state of the world. When $c = p$, $s - p = 0$. The function u_r is always non-negative. Therefore $c = p$ is the solution of (3) whenever $c = p$ is the solution of $\max_{0 \leqq c \leqq p} u_f$. We thus get:

Proposition 3. $f_1 = f_2 = 1$ *is a solution of* (1) *whenever the soluton of* $\max_{0 \leqq c \leqq p} u_f$ *is* p.

It is easy to apply Proposition 1 to conclude that, if G is more risk-averse than g, then the solution of $\max_{0 \leqq c \leqq p} u_f$ is $c = p$, and thus we get:

Proposition 4. $f_1 = f_2 = 1$ *is a solution of* (1) *if* G *is more risk-averse than* g.

Remark. The original problem (1) has been reduced to a simple steady-state problem. It is easy to check that the two steady-state equations

$$s = sfp + s(1-f)(1-p) + (1-s)p \quad \text{and} \quad c = sf$$

are satisfied by

$$c = f/[2 + f(1/p - 2)] \quad \text{and} \quad s = \frac{1}{2} - c\left(\frac{1}{2p} - 1\right).$$

Thus Proposition 2 may be considered a justification for looking at the steady-state model for this problem.

5. Concluding remarks

There are many potential reasons for why one might wish to avoid excessive rule changes. This paper focuses only on one reason which is due to differences in attitude toward risk. Let me admit that this explanation

Table A.1
The transition matrix.

	(1, 1)	(1, 2)	(2, 1)	(2, 2)
(1, 1)	$1-p$	p	0	0
(1, 2)	$(1-f_{12})p$	$(1-f_{12})(1-P)$	$f_{12}p$	$f_{12}(1-p)$
(2, 1)	$f_{21}(1-p)$	$f_{21}p$	$(1-f_{21})(1-p)$	$(1-f_{21})p$
(2, 2)	0	0	p	$1-p$

does not get to the heart of the problem. Notice that if we relax the assumption that the governor can change the rule only if it is currently incorrect we may get the result that the governor changes the rule in order to create excess uncertainty. We do observe such behaviour in real life (especially in non-democratic organizations) but, however, I doubt that it is an outcome of differences in risk-aversion. Given my own reservation the paper should be regarded as just pointing out one consideration from among many which should be included in a more complete analysis of the problem.

Appendix: Proof of Proposition 2

Let (f_1, f_2) be a strategy. Define a Markov chain on $R \times W$ where the occurrence of (r, w) at time t means that $f^{t-1}=r$ and $w^t=w$. Let Q be the transition matrix induced by f, as shown in table A.1.

When the matrix is irreducible [see Feller (1950)] the relative frequencies of the states in $R \times W$ converge a.s. (and independently of the initial state) to the unique vector $\alpha=(\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22})$ which satisfies

$$\alpha Q=\alpha \quad \text{and} \quad \sum \alpha_{rw}=1.$$

When the matrix is not irreducible, the limit of the frequencies is $(0,0,\frac{1}{2},\frac{1}{2})$ or $(\frac{1}{2},\frac{1}{2},0,0)$.

Clearly $\{c_r^t\}$ converges a.s. to

$$c_r=\frac{\alpha_{r\hat{r}}}{\alpha_{rr}+\alpha_{r\hat{r}}} f_r, \quad \text{where} \quad \hat{1}=2 \quad \text{and} \quad \hat{2}=1.$$

It is simple to calculate that

$$c_r=\frac{f_r}{2+f_r(1/p-2)}.$$

Our assumptions ensure that $\alpha(c)$ is continuous [see (3)]. Then, from the continuity of u, the target function (1) exists and is equal to

$$\alpha_{11}u[\alpha(c_1)]+\alpha_{12}u[1-\alpha(c_1)]+\alpha_{21}u[1-\alpha(c_2)]+\alpha_{22}u[\alpha(c_2)]$$

$$=(\alpha_{11}+\alpha_{12})\left\{\frac{\alpha_{11}}{\alpha_{11}+\alpha_{12}}u[\alpha(c_1)]+\frac{\alpha_{12}}{\alpha_{11}+\alpha_{12}}u[1-\alpha(c_1)]\right\}$$

$$+(\alpha_{21}+\alpha_{22})\left\{\frac{\alpha_{21}}{\alpha_{21}+\alpha_{22}}u[\alpha(c_2)]+\frac{\alpha_{22}}{\alpha_{21}+\alpha_{22}}u[1-\alpha(c_2)]\right\}.$$

Note that

$$\frac{\alpha_{12}}{\alpha_{11}+\alpha_{12}}=\frac{1}{2+f_1(1/p-2)}$$

depends only on f_1 and similarly $\alpha_{21}/(\alpha_{21}+\alpha_{22})$ depends only on f_2. Moreover, $\alpha_{12}/(\alpha_{11}+\alpha_{12})$ and c_1 depend on f_1 precisely as $\alpha_{21}/(\alpha_{21}+\alpha_{22})$ and c_2 depend on f_2. Using also the symmetry of u and v, it follows that the target function of (1) has the form

$\varepsilon(f_1,f_2)\delta(f_1)+[1-\varepsilon(f_1,f_2)]\delta(f_2)$. Such a function is maximized by $f_1=f_2=f^*$, where f^* maximizes $\delta(f)$.

References

Cross, R., 1961, Precedent in English law (Clarendon Press, Oxford).
Derman, C., 1970, Finite state Markovian decision processes (Academic Press, New York).
Feller, W., 1950, An introduction to probability theory and its applications (John Wiley, New York).
Paroush, J. and A. Rubinstein, 1982, Pessimism and risk aversion, Department of Economics research report no. 145 (The Hebrew University, Jerusalem).