A Formal Model of Learning and Policy Diffusion: A Comment on Multiple Unknown Policies

Craig Volden
Department of Political Science
The Ohio State University

Michael M. Ting
Department of Political Science and SIPA
Columbia University

Daniel P. Carpenter
Department of Government
Harvard University

July 1, 2008

Abstract

Volden, Ting, and Carpenter (2008) present a theoretical model of learning and policy choice across governments. This model is restricted to the simple case of two policy choices, only one of which has initially unknown payoffs. In this note we extend the model to include a second unknown policy. We find that while pure strategy equilibria are no longer assured in this variant of the model, many of the properties of the original model remain intact.

The model of Volden, Ting, and Carpenter (2008) (hereafter, VTC) lays the groundwork for studying numerous empirical implications of decentralized learning-based policy diffusion. One of their crucial simplifying assumptions is its limitation to two policies; one with known effects and one with initially unknown payoffs. In the policy world there are often competing policy ideas, each with various unknown effects. As such, it is valuable to assess whether the main results of VTC continue to hold in a more complicated policy arena. In this comment, we extend the model to also include a "rightist" experimental policy.

We assume that the reader is familiar with the VTC model, and extend the model and notation as follows. The new alternative is labeled policy 3, and thus policies are chosen from the set $\{1, 2, 3\}$. Where appropriate, the variables μ, ρ, r, θ, $\overline{\omega}$, $\underline{\omega}$, and π are subscripted with the policy in question. Thus, $\theta_3 = \overline{\theta}_3$ with probability ρ_3.

The presence of multiple unknown policies gives rise to a large number of possible configurations of policymaker preferences. For example, if policy 3's spatial location, x_3, is sufficiently distant from x_1, then a state choosing between policies 1 and 2 would not need to worry about the realization of $\overline{\theta}_3$. In this case, policy 3 would be that state's last choice regardless of its effectiveness.

To focus on a non-trivial but tractable (and plausible) case, we therefore examine a particular spatial configuration of policies. Let $x_1 < x_2 < x_3$, so that policy 1 is the leftist policy and policy 3 the rightist policy. For any policies i and m, we denote the cutpoints analogous to c and e in VTC by c_{im} and e_{im}. Since it will sometimes be necessary to denote the result of two uncertain policies, we use "$\overline{i}$" and "$\underline{i}$" to denote $r_i = 1$ and $r_i = -1$, respectively, in the policy subscripts. Thus, $c_{\overline{1}2}$ is the cutpoint separating the ideal points of players who would prefer (myopically) policies 1 and 2, given that policy 1 was revealed to be effective. Note that $c_{\overline{1}2}$ would be denoted "$\overline{c}$" in the VTC model.

To focus attention on a non-trivial but tractable set of parameters, we assume that:

$$c_{12} < c_{2\overline{3}} < c_{\overline{1}2} < c_{23}. \tag{1}$$

These assumptions are equivalent to placing bounds on $\overline{\omega}_i$ (or, equivalently, on x_i).[1] Both effectiveness benefits $\overline{\omega}_1$ and $\overline{\omega}_3$ are high enough to allow a sufficiently moderate state to choose any policy, depending on the experimental history. Without this assumption, no state would ever choose between policies 1 and 3, and thus the choice between policies 2 and 3 would simply mirror that between 1 and 2 in VTC. The $\overline{\omega}_i$'s are also low enough to restrict

[1] Without these assumptions the model requires many more subcases. Each has similar properties to those uncovered here. Setting these aside is therefore a simplification for tractability purposes only.

the possible choices of extreme states. For example, a state that prefers the *ex ante* value of policy 1 over policy 2 will also prefer it to a good realization of policy 3.

The "overlap" between policies 1 and 3 implies that the locations of ideal points relative to $c_{\overline{13}}$ become relevant for their period 2 choice when both policies are revealed as effective ($r_1 = r_3 = 1$). A policymaker with ideal point $z_j \in (c_{2\overline{3}}, c_{23})$ prefers policy 3 to policy 2 if $r_3 = 1$, but she may also prefer policy 1 to policy 2 if $r_1 = 1$. Using this expanded set of cutpoints, it is possible to derive an analog to Lemma 1 of VTC for period 2 policy choices.

Lemma 1 Period 2 Policy Choice with Multiple Uncertain Policies.

$$p_j^{2*} = \begin{cases} 1 & \text{if } z_j < c_{\underline{1}2}, \text{or } z_j \in [c_{\underline{1}2}, c_{12}) \text{ and } r_1 > -1, \\ & \quad \text{or } z_j \in [c_{12}, c_{2\overline{3}}) \text{ and } r_1 = 1, \\ & \quad \text{or } z_j \in [c_{2\overline{3}}, c_{\overline{1}2}) \text{ and } r_1 = 1 \text{ and } r_3 < 1, \\ & \quad \text{or } z_j \in [c_{2\overline{3}}, c_{\overline{13}}) \text{ and } r_1 = r_3 = 1 \\ 3 & \text{if } z_j > c_{2\underline{3}}, \text{or } z_j \in (c_{23}, c_{2\underline{3}}] \text{ and } r_3 > -1, \\ & \quad \text{or } z_j \in [c_{\overline{1}2}, c_{23}] \text{ and } r_3 = 1, \\ & \quad \text{or } z_j \in (c_{2\overline{3}}, c_{\overline{1}2}] \text{ and } r_3 = 1 \text{ and } r_1 < 1, \\ & \quad \text{or } z_j \in [c_{\overline{13}}, c_{\overline{1}2}] \text{ and } r_1 = r_3 = 1 \\ 2 & \text{otherwise.} \end{cases}$$

In contrast with Lemma 1 of VTC, the most centrist policymakers now only adopt policy 2 in the second period if neither 1 or 3 has been revealed as effective. If both alternatives have been shown to be effective, policymakers select the "closer" policy, as determined by cutpoint $c_{\overline{13}}$.

Turning to first-period experimentation decisions, it is helpful to consider again a decision-theoretic example, before examining the game-theoretic equilibrium. For a single state S, the experimental cutpoint e_{12} is identical to that in the two-policy decision-theoretic model. It is also straightforward to define the analogous experimental cutpoint e_{23} between policies 2 and 3. Extending the logic of Proposition 1 of VTC, a centrist policymaker ($z \in [e_{12}, e_{23}]$) prefers policy 2 to experimenting with either alternative. The state will experiment with policy 1 if $z < e_{12}$ and with policy 3 if $z > e_{23}$.

A more difficult (and more interesting) case arises when $e_{12} > e_{23}$. Now, for any ideal point $z \in [e_{23}, e_{12}]$, S will prefer experimenting with either policy 1 or policy 3 over policy 2. To characterize the optimal policy choice, note that at $z = e_{23}$, S prefers to experiment with policy 1 because she is indifferent between policies 2 and 3 and favors policy 1 over policy 2. So a policymaker with an ideal point just to the right of e_{23} experiments with 1 and retains it if it is shown to be effective, and otherwise switches to 2. She does not switch to policy 3 because, by (1), an experimental policy without evidence of effectiveness is not preferred to policy 2 in this region. Moreover, because S is the only relevant experimenter in the decision-theoretic model, no evidence of an untried policy's effectiveness could emerge by period 2. A symmetric experimentation strategy with policy 3 holds just to the left of e_{12}.

At some ideal point between e_{23} and e_{12}, S is indifferent between experimenting with policies 1 and 3. This is an experimental cutpoint e_{13} that, for a state with ideal point z, equates the expected utilities from experimenting with each policy:

$$(1+\delta\rho_3\pi_3)\left[u(|z-x_3|)-u(|z-x_2|)\right]+(1+\delta\rho_1\pi_1)\left[u(|z-x_2|)-u(|z-x_1|)\right]= \\ \mu_1(\rho_1)-\mu_3(\rho_3)+\delta\left(\rho_1\pi_1^2\overline{\omega}_1-\rho_3\pi_3^2\overline{\omega}_3\right). \tag{2}$$

Thus equation is analogous to expression (8) of VTC (the two-policy baseline model), now complicated by the third policy option. As with e_{12} and e_{23}, e_{13} is uniquely defined because the left-hand side of (2) is unbounded and strictly increasing in z, while the right-hand side is constant in z. Thus for $z<e_{13}$, S experiments with policy 1 in period 1, and for $z\geq e_{13}$, S experiments with policy 3.[2]

Turning finally to the full game-theoretic model, the calculation of each experimental cutpoint is complicated by two facts. First, Lemma 1 implies that *any* policy might be optimal in period 2, depending on what is revealed in period 1 and on policymaker S_j's preferences. An example illustrates this logic. Suppose that $z_j\in[c_{\overline{1}3},c_{\overline{1}2}]$, so that S_j's period 2 choice would be policy 3 if $r_3=1$, policy 1 if $r_3\neq 1$ and $r_1=1$, and policy 2 otherwise. Compared with the two-policy calculations of expression (10) of VTC, the expected utilities from choosing policies 2 and 3 in the first period now entail an additional condition regarding the anticipated second-period policy choice. To account for the probabilities of finding *either* experimental policy effective, we extend the baseline model's notation to allow k_1 and k_3 to denote the number of other states choosing policies 1 and 3, respectively. The experimental cutpoint $e_{23}(k_1,k_3)$ is the ideal point z_j that satisfies:

$$\left(1+\delta\rho_3\pi_3\left(1-\pi_3\right)^{k_3}\right)\left[u(|z_j-x_2|)-u(|z_j-x_3|)\right]+ \\ \delta\rho_1\left(1-\left(1-\pi_1\right)^{k_1}\right)\rho_3\pi_3\left(1-\pi_3\right)^{k_3}\left[u(|z_j-x_1|)-u(|z_j-x_2|)\right]= \\ \mu_3\left(\rho_3\right)+\delta\rho_3\pi_3^2\left(1-\pi_3\right)^{k_3}\overline{\omega}_3-\delta\rho_1\pi_1\left(1-\left(1-\pi_1\right)^{k_1}\right)\rho_3\pi_3\left(1-\pi_3\right)^{k_3}\overline{\omega}_1. \tag{3}$$

This yields a unique experimental cutpoint, since the left-hand side of (3) is unbounded

[2] The formal derivation of this result is omitted, but closely resembles the proof of Proposition 1 of VTC.

and decreasing in z_j, while the right-hand side is constant in z_j. Analogous cutpoints can be similarly derived for cases in which $z_j \in [c_{2\overline{3}}, c_{\overline{1}3})$, as well as for $e_{12}(k_1, k_3)$.

It can be verified from (3) that $e_{23}(k_1, k_3)$ is increasing in k_1 and k_3; that is, as experimentation with *either* unknown policy increases, the marginal propensity to experiment decreases. This implies that k_3 affects not only the choice between policies 2 and 3, but also the choice between policies 1 and 2. Since higher values of k_3 raise the probability of discovering that policy 3 is effective, and an effective policy 3 is preferred to an unknown policy 1, experimentation with policy 3 will diminish the marginal value of experimenting with policy 1. Thus in addition to free-riding on other states that choose policy 1, states choosing between policies 1 and 2 may also free-ride on those who are experimenting with their "second-best" choice. This raises the possibility that learning-based policy diffusion could drive a state to adopt the policy that would have been its lowest-ranked alternative in the decision-theoretic model. Thus this model presents conditions under which a bandwagon effect could occur not due to irrational actions but through rational calculations that lead to suboptimal outcomes.

The second complication raised by multiple unknown policies is that each state with $z_j \in [c_{2\overline{3}}, c_{\overline{1}2}]$ also needs to consider whether to experiment with policy 1 or 3. A state whose first preference would be for a successful policy 1 may actually experiment with policy 3 if many other states are already going to experiment with policy 1. Extending equation (2) would yield the experimental cutpoint $e_{13}(k_1, k_3)$.[3] In a pure strategy equilibrium, the cutpoints $e_{13}(k_1, k_3)$, $e_{12}(k_1, k_3)$ and $e_{23}(k_1, k_3)$ can fully specify the optimal period 1 policy choice for each state, given any configuration of other experimenters.

Unfortunately, the technique used to construct the pure strategy equilibrium in Proposition 2 of VTC cannot be used under multiple uncertain policies. Unlike $e_{12}(\cdot)$ and $e_{23}(\cdot)$, $e_{13}(\cdot)$ may be non-monotonic in k_1 and k_3. As a result, it is not always possible to partition the states into ideological "regions" of states that choose the same policies. Pure-strategy equilibria exist for many parameter configurations, but we cannot establish their existence in general. Instead, it is possible to show the existence of a mixed-strategy equilibrium. As in any mixed-strategy equilibrium, this requires that some policymakers experiment with the policies 1 and 3 with probabilities that induce others to be indifferent over their policy choices.[4] These policymakers must of course be "moderates" who do not have strong

[3] Analogously to e_{13}, we can show that the uniqueness of $e_{13}(k_1, k_3)$ follows from the increasing ideological returns to choosing policies 1 or 3 as z_j becomes more extreme.

[4] This should not be interpreted as states actively trying to deceive or manipulate the strategies of one another. Rather, the states adopting mixed strategies here are truly indifferent over which policy to adopt.

ideological leanings toward policies 1 and 3.

Proposition 1 Strategic Experimentation with Multiple Uncertain Policies. *There exists a mixed-strategy equilibrium where in period* 1*:*

$$p^{1*} = \begin{cases} 1 & \text{if } z_j \leq c_{12} \\ 3 & \text{if } z_j \geq c_{23}, \end{cases}$$

and policymakers with ideal points $z_j \in (c_{12}, c_{23})$ *mix over the three policies; and period* 2 *strategies are given in Lemma 1.*

Proof. The strategies for $t = 2$ are given by Lemma 1. We therefore restrict attention to strategies at $t = 1$.

It is clear that, given the period 2 strategies in Lemma 1, for any profile of (pure) policy choices s, there exists a uniquely defined expected payoff for each player. Thus consider the normal form game in which payoffs for each s are given by this payoff vector. The existence of a mixed strategy Nash equilibrium strategy profile σ of this game follows from the Nash Theorem. It therefore follows that there exists a perfect Bayesian equilibrium in which the period 2 pure strategies are given by Lemma 1 and period 1 mixed strategies are given by σ.

For $z_j \notin (c_{12}, c_{23})$ at $t = 1$, strategies are derived by an argument identical to that in the proof of Proposition 1 of VTC. All players with $z_j \geq c_{23}$ strictly prefer experimenting with policy 3 to policy 2. By (1), these policymakers never have an incentive to choose policy 1. Thus, $p^{1*} = 3$. Similarly, for all policymakers such that $z_j \leq c_{12}$, $p^{1*} = 1$. ■

The need to rely on a mixed-strategy equilibrium for some configurations of state preferences raises a note a caution about the applicability of the baseline model in more complex environments. However, the main qualitative features from that model carry over to the case of multiple uncertain policies. As before, both the decision-theoretic and game-theoretic models demonstrate similar policy adoptions by similarly positioned states. Once again, consistent with Comment 1 of VTC, this ideological sorting more cleanly separates states in the second period of the game-theoretic model due to identical learning about policy successes across states. And, once more, free-riding on the policy experiments of others takes place only in the game-theoretic model, with a larger range of free-riding when there are more states to free-ride upon. Such consistent findings make us more confident in the empirical implications offered in VTC.

References

Volden, Craig, Michael M. Ting, and Daniel P. Capenter. 2008 (forthcoming). "A Formal Model of Learning and Policy Diffusion." *American Political Science Review*.