JUDICIAL DEFERENCE TO INCONSISTENT AGENCY STATUTORY INTERPRETATIONS*

Yehonatan Givati† & Matthew C. Stephenson‡

ABSTRACT

Although administrative law doctrine requires courts to defer to an agency’s reasonable statutory interpretation, the doctrine is unclear whether an agency gets less deference when it changes its own prior interpretation. We analyze how judicial deference to revised agency interpretations affects those interpretations’ ideological content. We find a non-monotonic relationship between judicial deference to inconsistent agency interpretations and interpretive extremism. This arises because the initial agency must choose between an ideologically extreme interpretation, and a more moderate interpretation that a future agency would not attempt to revise. Greater judicial deference to revised interpretations increases the attractiveness of the latter approach by allowing the initial agency to lock in a more extreme interpretation. Normatively, our results suggest that an interest in responsiveness of interpretive policy to the incumbent administration’s preferences favors maximum deference to revised interpretations, while an interest in ideological moderation favors a less deferential posture to interpretive revisions.

---

* We are grateful to Louis Kaplow, Daryl Levinson, and Eric Posner for helpful comments on earlier drafts.


‡ Assistant Professor, Harvard Law School.
Federal agencies in the United States have a great deal of flexibility in interpreting the statutes they administer. Although an agency may not contravene clear statutory directives, a reviewing court is supposed to defer to an agency’s reasonable construction of an ambiguous statute, even if the court believes a different interpretation is superior. Put another way, the reviewing court’s obligation is not to ascertain the best “point estimate” of statutory meaning, but rather to define the bounds of a “policy space” from which the agency can select its preferred interpretation (Elliott 2005; Stephenson and Vermeule 2009). But what if the agency changes its position? Should courts be any less deferential to an agency interpretation that differs substantially from the agency’s own prior interpretation? Current doctrine is less clear on the answer to this question.

This paper considers one aspect of this issue: whether reducing the degree of judicial deference to an inconsistent agency interpretation will lead an agency’s interpretation to be more aggressive (i.e., closer to the preferred position of the incumbent leadership) or more moderate (i.e., closer to the midpoint between the ideal interpretations of the incumbent party and the opposition party). We consider both the agency’s initial interpretive choice—when the agency is writing on a clean slate, with no prior agency interpretation—and the decision of the agency at a later time, under the control of a different presidential administration with different policy preferences. Our most important finding is that there is a non-monotonic relationship between the degree of judicial deference to revised agency interpretations and the aggressiveness of the agency’s initial interpretation. When courts are very likely to defer to a revised interpretation, or very likely to reject a revised interpretation, the agency’s initial interpretation will strongly favor the interests of the incumbent administration, to the
detriment of the party out of power. When the courts take a more intermediate approach—sometimes upholding revised interpretations but sometimes rejecting them, or increasing the transaction costs to agencies of revising prior interpretations—the agency’s initial interpretation is more likely to be ideologically moderate.

The reason for this is as follows: Although the incumbent administration would like to secure the most favorable interpretation possible, it would also like to lock in its interpretation, reducing the likelihood of future reversal. One way for the agency to protect itself from reversal is to issue a more moderate interpretation, so that a future administration does not find it worthwhile to bear the costs of revising it. If courts are generally unwilling to accept revised agency interpretations, however, the incumbent has little incentive to be moderate, because even an extreme interpretation is protected by the courts from administrative revision. Likewise, if the courts are very deferential to revised interpretations, the incumbent also has little incentive to be moderate, because the concessions that would be needed to preclude a future interpretive revision are too costly.

The moderating strategy only becomes attractive to the incumbent when the courts are somewhat less deferential to inconsistent interpretations. In that case, the incumbent can lock in its interpretation by making modest but meaningful concessions on the ideological content of the interpretation.

The normative and doctrinal ramifications of this finding depend crucially on the specification of the social objective function. Some might argue that the appropriate normative objective is correspondence between the agency’s interpretation and the preferences of the incumbent administration, on the logic that the incumbent’s preferences tend to track the preferences of a majority of voters. If so, one would prefer
that courts defer just as much to a revised agency interpretation as to the original interpretation. On the other hand, one might believe that the median voter’s ideal interpretation is usually somewhere between the ideal interpretations of the two dominant political parties. If so, a judiciary that is somewhat less deferential to a revised agency interpretation may be normatively desirable, because it can induce a more moderate interpretation under administrations of both parties. In short, normative assessment of deference doctrine depends in large part on whether it is more important to foster *responsiveness* to the ideology of the party in power or to encourage *moderation* between the views of the major political competitors.

The paper is organized as follows. Part I provides a brief overview of extant legal doctrine and scholarly commentary on the question whether a reviewing court should be less deferential to an agency interpretation that differs substantially from the agency’s own prior interpretation. Part II develops a stylized formal model to analyze how judicial deference doctrine may affect the aggressiveness of agency interpretations. Part III considers the normative implications of these results. A brief conclusion follows.

I. **AGENCY INTERPRETIVE INCONSISTENCY: DOCTRINE AND COMMENTARY**

The starting point for any discussion of judicial deference to federal administrative agencies’ statutory interpretations is the Supreme Court’s seminal *Chevron v. Natural Resources Defense Council* decision, 467 U.S. 837 (1984). *Chevron* held that a reviewing court should defer to an agency’s reasonable interpretation of a statute the agency administers, even if the court would have interpreted the statute differently. But
the courts have sent mixed signals on whether this strong form of deference applies in the same way when the agency’s interpretation, though reasonable, differs from the agency’s own prior interpretation. *Chevron* itself involved the Environmental Protection Agency’s reversal of its earlier position on the meaning of the term “stationary source” in the Clean Air Act, and the *Chevron* opinion indicated that a change in the agency’s interpretive position is irrelevant to the appropriate level of judicial deference. Indeed, Justice Stevens’ opinion for the Court emphasized that an “initial agency interpretation is not instantly carved in stone,” since interpretive flexibility is an important aspect of agency authority (467 U.S. at 863). Yet only a few years later, in *I.N.S. v. Cardoza-Fonseca*, 480 U.S. 421 (1987), the Court (in an opinion also authored by Justice Stevens) declared that an “agency interpretation of a relevant provision which conflicts with the agency’s earlier interpretation is ‘entitled to considerably less deference’ than a consistently held view” (480 U.S. at 446 n.30).

The tension between the positions taken in *Chevron* itself and in the nearly-contemporaneous *Cardoza-Fonseca* decision caused a great deal of uncertainty about whether interpretive inconsistency mattered under the *Chevron* framework. In *Rust v. Sullivan*, 500 U.S. 173 (1991), the Supreme Court seemed to come down squarely on the side of the original *Chevron* position: *Rust* declared that an agency gets *Chevron* deference even if it reverses an earlier interpretation, emphasizing that *Chevron* itself involved just such a reversal (500 U.S. at 186). Yet dicta in *Pauley v. BethEnergy Mines*, decided the same Term as *Rust*, declared that “the case for judicial deference is less compelling with respect to agency positions that are inconsistent with previously held views,” 501 U.S. 680, 698 (1991), and a couple Terms later, in *Good Samaritan Hospital*
v. Shalala, the Court cited Cardoza-Fonseca for the proposition that “the consistency of an agency’s position is a factor in assessing the weight that position is due,” 508 U.S. 402, 417 (1993). The most recent word from the Supreme Court on this issue is National Cable & Telecommunications Association v. Brand X Internet Services, 545 U.S. 967 (2005). Brand X reaffirmed the irrelevance of agency inconsistency, declaring that such inconsistency “is not a basis for declining to analyze the agency’s interpretation under the Chevron framework” (545 U.S. at 981). Yet the Court’s history of vacillation on this issue would give a savvy litigator reason to question whether an agency interpretation that reversed a prior interpretation would receive the same degree of deference as a more longstanding interpretation.¹

In addition to these conflicting signals regarding the relevance of interpretive inconsistency under the Chevron framework, the Court’s holding in United States v. Mead Corporation, 533 U.S. 218 (2001), made clear that some agency interpretations of statutes—particularly those that appear in less formal guidance documents or interpretive statements, rather than rules or orders—would not be analyzed under Chevron at all, but instead would be reviewed pursuant to the less deferential standard articulated in Skidmore v. Swift & Co., 323 U.S. 134 (1944). Under Skidmore, a reviewing court is

¹ Gossett (1997) reports that notwithstanding the mixed signals from the Supreme Court, in practice federal appeals courts do not seem to defer substantially less to changed agency interpretations. If accurate, this evidence suggests that the position advanced in cases like Chevron, Rust, and Brand X better captures the approach taken by most courts than does the contrary language in cases like Cardoza-Fonseca, Pauley, and Good Samaritan Hospital. Still, using court decisions to measure the level of deference to revised agency interpretation may be an imprecise method, since a lower level of deference may result in more moderate interpretations that will not be brought before a court (Givati 2008).
supposed to “respect” an agency’s view of the correct interpretation, but the weight accorded to the agency’s view depends on a variety of factors, including “consistency with [the agency’s own] earlier and later pronouncements” (323 U.S. at 140). This suggests that even if an agency gets full *Chevron* deference for a changed interpretation if that interpretation is issued in a rulemaking or formal adjudication, if the agency opts to issue the interpretation in a less formal context, the lack of consistency may result in less judicial deference. Just how much this factor matters, however, remains unclear.\(^2\)

The question whether, or under what conditions, courts ought to defer less to agency statutory interpretations that are inconsistent with the agency’s own prior interpretations has provoked debate among administrative law scholars. Many of these scholars conclude that agencies’ need to respond flexibly to changing circumstances militates in favor to deferring just as much to revised agency interpretations as to initial agency interpretations (Diver 1985; Sunstein 1990; Weaver 1992; Gossett 1997; Schuren 2001; Dotan 2005). A powerful additional argument for this position invokes the importance of political accountability: changes in an agency’s interpretive position may reflect changes in the agency’s political priorities—often triggered by a change in the presidential administration—and courts should respect this legitimate rationale for policy change (Pierce 1988; Scalia 1989; Sunstein 1990; Gossett 1997). The principal countervailing consideration noted in the literature is the “rule of law” interest in predictability and consistency in the meaning of law (Sunstein 1990; Merrill 1992; Murphy 2005). Another

\(^2\) In one of the few attempts to assess this issue systematically, Hickman and Krueger (2007) report that although federal appeals court cases applying the *Skidmore* standard consider agency inconsistency, and do sometimes defer less to inconsistent agency interpretations, agency consistency appears less important than the other factors *Skidmore* identifies as relevant.
concern is that a policy-motivated change in interpretation might reflect not a legitimate policy change, but rather an unjustified partisan hostility to an earlier administration’s regulatory programs (Weaver 1992). For these reasons, some scholars have suggested that courts should be less deferential to revised agency interpretations (Sunstein 1990; Merrill 1992), or at least that courts should impose special additional explanatory burdens on agencies when such inconsistency occurs (Diver 1985; Sunstein 1990; Merrill 1992; Weaver 1992). In a similar spirit, Murphy (2005) argues that a revised interpretation should receive substantial deference only when the agency’s initial interpretation was issued in a procedural context that makes it difficult to reverse.

Interestingly, in emphasizing the trade-off between flexibility values and rule-of-law values, the extant administrative law literature has focused less on how the expected level of judicial deference to changed agency administration affects the content—in particular, the ideological slant—of agency interpretations. Much of the literature seems implicitly to presume that the content agency interpretive choices will look essentially the same regardless of the nature of judicial doctrine on deference to revised interpretations. Yet a rich cognate literature in political economy suggests that the substantive choices of agencies or other political actors may depend crucially on strategic calculations regarding the insulation of their decisions from future reversal (e.g., McCubbins et al. 1987, 1989; Moe 1989, 1990; De Figueiredo 2002; Stephenson 2003; Givati 2008). The administrative law literature has similarly noted in other contexts how current administrations attempt to insulate their decisions from future reversal (Beerman 2003; Mendelson 2003; O’Connell 2008). Perhaps, then, judicial doctrine on deference to inconsistent agency interpretations—which, after all, affects agencies’ ability to insulate
their policy choices from future reversal—might have important and non-obvious effects on the substantive content of these interpretations. We explore this issue using the stylized formal model developed in the next section.

II. THE MODEL

A. Primitives

Consider a simple model with two periods, \( t \in \{1, 2\} \), two strategic players—a Conservative Party (C) and a Liberal Party (L)—and one non-strategic player, the Judiciary. In each period, one of the two parties controls the Presidency, and therefore controls the interpretive decisions of an administrative agency charged with implementing some congressional statute. (For simplicity, we take the content of the statute as fixed and exogenous.) The interpretation that prevails in period \( t \) is \( x_t \in \mathbb{R} \).

In the first period, without loss of generality, the Conservative Party is in power and can choose interpretation \( x_1 \). In the second period, the Liberal Party is in power. (In other words, the second period is defined as the point at which the Liberal Party takes over.) The agency, under Liberal control, must decide whether to revise its interpretation of the statute. Formally, the Liberal administration chooses \( x_2 \). If the Liberal agency does not revise the interpretation, then the first-period interpretation selected by the Conservative administration applies in the second period (\( x_2 = x_1 \)). If the Liberal agency does revise its interpretation by selecting some \( x_2 \neq x_1 \), then that interpretation, if upheld, prevails in the second period. If the Liberal agency chooses to revise the prevailing
interpretation, it also pays some fixed cost $k \in [0, \frac{1}{1+\sigma})$, which can be interpreted as including the procedural costs of issuing a new interpretation of the statute, the political cost of reopening a contentious political issue, and the expected litigation costs of defending the revision in court.\(^3\)

The Conservative Party’s ideal interpretation is $x_C \leq 0$, and its utility in period 1 is $u_C(x_1) = -|x_1 - x_C|$. In period 2, the Conservative Party’s utility is $u_C(x_2) = -|x_2 - x_C|$ if the Liberal Party does not revise the Conservative Parties’ interpretation, and is $u_C(x_2) = -|x_2 - x_C| + \beta k$ if the Liberal Party revises the statutory interpretation and bears the cost $k$. The parameter $\beta \in [0,1]$ captures the possibility that the Conservative Party may benefit from the Liberal agency’s diversion of resources to revision of the prior interpretation.\(^4\)

The Liberal Party’s ideal interpretation is $x_L \geq 1$, and its policy utility in period $t$ is $u_L(x_t) = -|x_t - x_L|$. In the first period, both parties discount the expected second-period payoff by the parameter $\delta \in (0,1]$. This parameter reflects how much each party cares about its future utility relative to its present utility, as well as how long each party expects the current period to last.

In each period, the agency’s interpretative decision is subject to judicial review. In the first period the Judiciary, applying something like the *Chevron* doctrine, upholds any

---

\(^3\) The upper bound on $k$ guarantees that there are values of the other parameters such that the Liberal agency will sometimes be willing to revise a previous interpretation.

\(^4\) Although we include the $\beta k$ term in the Conservative Party’s utility function for completeness, our main qualitative results are unaffected by $\beta$, so long as $\beta$ is not too large. While our interpretation of $k$ and $\beta$ suggests the possibility of even more complex strategic behavior, in which the Conservative and Liberal administrations both have to select which among a large set of possible issues they will choose to address, we defer these complications to future research.
$x_1 \in [0,1]$. That is, the [0,1] interval is the “policy space” that the statutory ambiguity opens up for the agency; $x_1=0$ is the most conservative agency decision that the Judiciary would uphold, while $x_1=1$ is the most liberal agency decision that the court would uphold.\(^5\) In the second period, if the Liberal agency does not revise the prevailing interpretation, then there is no judicial review. If, however, the Liberal agency chooses a new interpretation of the statute, $x_2 \neq x_1$, that choice is again subject to judicial review. In the second period as in the first, the Judiciary will invalidate any interpretive choice outside the [0,1] interval. Furthermore, the Judiciary might be less deferential to inconsistent agency interpretations. We model this by assuming that, even if the Liberal agency selects some $x_2 \in [0,1]$, nonetheless if $x_2 \neq x_1$ there is some probability, $\alpha \in [0,1]$, that the Judiciary will reject the agency’s new interpretation. If so, the first-period interpretation prevails in the second period ($x_2 = x_1$).\(^6\) The parameter $\alpha$ therefore captures

\(^5\) If the Judiciary were to reject the agency’s first-period interpretation, the court would select some alternative interpretation $x_1 = q \in [0,1]$. In our model, however, this never occurs in equilibrium, because the Conservative agency is always weakly better off selecting some $x_1 \in [0,1]$.

\(^6\) This paper takes no position on judicial motivations in reviewing agency interpretations of statutes, a topic on which there is a great deal of debate (e.g., Cohen and Spitzer 1994, 1996; Revesz 1997; Cross and Tiller 1998; Stephenson 2004; Miles and Sunstein 2006; Richards et al. 2006; Smith 2007). That said, we do assume that whatever the judges’ underlying motivations, in practice they behave as if there is a defined “zone of discretion” from which the agency may select an initial interpretation. This assumption entails some loss of generality, insofar as it assumes away the possibility that an agency’s effective zone of discretion may depend on the correspondence between judicial and agency preferences (e.g. Spiller 1992; Cohen and Spitzer 1994, 1996; Stephenson 2006b). Nonetheless, we think it is reasonable, at least as a first cut, to model the judiciary as a consistent and faithful adherer of standard deference doctrine, such that the discretionary zone is constant. This assumption can and should be relaxed in future work. Our assumption
the effective legal doctrine regarding judicial deference to revised administrative interpretations. Greater values of $\alpha$ indicate a lower likelihood that the Judiciary will uphold an agency interpretation that is inconsistent with the agency’s prior views. If $\alpha=0$, then inconsistency makes no difference whatsoever—as cases like Rust, Brand X, and Chevron itself suggest ought to be the case. A value of $\alpha=1$ would represent the (unrealistic) extreme case in which the agency’s initial interpretation is indeed “carved in stone” and cannot be revised by the agency, absent congressional amendment of the statute. Values of $\alpha$ between 0 and 1 indicate an intermediate position suggested by cases like Cardoza-Fonseca, Good Samaritan Hospital, and Skidmore: inconsistency increases the chances that the reviewing court will reject the agency’s interpretation, but does not guarantee it.

**B. Equilibrium**

We find the equilibrium of the game by backward induction. For simplicity, and without loss of generality, we assume that $x_C=0$ and $x_L=1$.\(^7\) In the second period, if the Liberal agency decides to incur the cost $k$ and revise the interpretation that was chosen in

---

\(^7\) If either party had more extreme preferences, i.e., if $x_C<0$ or $x_L>1$, then that party would always suffer some additional utility loss even if it achieved its most-favored judicially acceptable interpretation. This would affect that Party’s overall utility level, but not its choice.
the first period ($x_1$), the Liberal agency would choose whatever new interpretation ($x_2$) maximizes its expected second period utility. That is, the Liberal chooses $x_2$ to maximize:

$$EU_{L}^{period\,2\,\text{revise}} = -1 - \alpha \left| 1 - x_2 \right| - \alpha \left| 1 - x_1 \right| - k$$  \hspace{1cm} (1)

From this it follows immediately that a Liberal agency that chooses to revise the interpretation of the statute would choose $x_2^* = 1$. So, the Liberal agency’s expected second period utility from attempting to revise the interpretation of the statute is $-\alpha \left| 1 - x_1 \right| - k$. If, on the other hand, the Liberal agency chooses to accept the prevailing first-period interpretation without attempting any revision, then its second period utility is:

$$U_{L}^{period\,2\,\text{accept}} = -\left| 1 - x_1 \right|$$  \hspace{1cm} (2)

The Liberal agency will choose to revise the existing statutory interpretation only if the expected second period utility from doing so is positive. This calculation will take into account both the likely result—whether the Judiciary upholds the new interpretation—and the cost, $k$, of attempting to effect this interpretive change. Thus, the Liberal agency will promulgate an alternative interpretation of the statute if and only if expression (1) is greater than expression (2). That is, the Liberal agency will attempt to revise the interpretation of the statute if and only if:

$$x_1 < \left( 1 - \frac{k}{1 - \alpha} \right) \equiv x^T.$$  \hspace{1cm} (3)

The value $x^T$ can be interpreted as a threshold value: it is the most conservative first-period interpretation that would not trigger an attempted reversal by the Liberal agency in the second period.
Now consider the Conservative agency’s choice. The Conservative agency can “lock in” its interpretation for both periods if it chooses a first-period interpretation $x_1 \geq x^T$, as we know from (3) that the Liberal agency would not attempt to revise an interpretation in that range. The Conservative agency’s utility from locking in interpretation $x_1$ for both periods is:

$$U_c \text{ lock in } = - 1 + \delta \cdot x_1$$  \hspace{1cm} (4)

This expression is decreasing in $x_1$, so the Conservative agency’s most-preferred $x_1$ in the $x_1 \in [x^T, 1]$ range is $x^T$. This is intuitive: $x^T$ is sufficiently moderate that the Liberal agency would accept it in the second period, and the Conservative agency has no reason to make additional concessions by choosing a more liberal interpretation.

If the Conservative agency were to choose a first-period interpretation $x_1 < x^T$, the Liberal agency would attempt a second-period reversal. Accordingly, the Conservative agency’s expected utility would from this more risky interpretive choice is:

$$EU_c \text{ risky } = - C + \delta \alpha \beta \gamma - \delta \alpha \beta \gamma \delta \beta k$$  \hspace{1cm} (5)

This expression is also decreasing in $x_1$, which implies that if the Conservative agency were to choose an $x_1 < x^T$, it would choose $x_1 = 0$.

We have shown that the Conservative agency’s optimal risky interpretation is $x_1 = 0$, and the optimal interpretation that the Conservative agency can lock in for both periods is $x_1 = x^T$. All that remains is to compare the expected utilities of these two interpretive approaches. Substituting these optimal values into (4) and (5) allows us to establish the condition under which the Conservative agency will (weakly) prefer the lock-in approach to the risky approach. That condition is:
\[ x^T \leq \frac{\delta (\alpha - \beta k)}{1 + \delta} \]  \hspace{1cm} (6)

Intuitively, if \( x^T \) is close enough to the Conservative agency’s ideal point (\( x_C = 0 \)), the Conservative agency will choose \( x_1 = x^T \) over \( x_1 = 0 \). Expression (6) implies that any interpretation that the Conservative agency is both willing and able to lock in for both periods must be weakly less than \( \frac{1}{2} \). Additionally, note that the more the Conservative Party benefits from the Liberal agency’s effort cost (i.e., the greater the value of \( \beta \)), the more likely it is that the Conservative agency will prefer the risky approach to the lock-in approach; if \( \beta k \geq (1 - \alpha) \), then there is no \( x^T \) sufficiently conservative for the Conservative agency to pursue the lock-in approach. For expositional convenience, we can substitute the expression for \( x^T \) from (3) in order to rewrite the condition in (6) as:

\[ Z = 1 - \alpha \left( 1 + \delta \alpha - k \right) + \delta - 1 - \alpha \delta \beta \leq 0 \]  \hspace{1cm} (7)

The Conservative agency’s optimal first-period interpretation, \( x_1^* \), is therefore:

\[ x_1^* = \begin{cases} 0 & \text{if } Z > 0 \\ x^T & \text{if } Z \leq 0 \end{cases} \]  \hspace{1cm} (8)

The expected second-period decision depends on whether the Conservative agency opts for the lock-in approach or the risky approach. If the former, the second-period interpretation is simply \( x^T \). If the latter, the second-period interpretation is either 1 (if the Judiciary accepts the Liberal agency’s revision) or 0 (if the Judiciary rejects the new interpretation on grounds of inconsistency). This means that in equilibrium the expected value of \( x_2 \) is:
\[ E_\mathcal{Z} = \begin{cases} 1 - \alpha & \text{if } Z > 0 \\ x^T & \text{if } Z \leq 0 \end{cases} \]  \hspace{1cm} (9)

\[ C. \text{ Comparative Statics} \]

We are interested principally in how \( \alpha \), the probability that the Judiciary reverses an agency interpretation (which we will call “inconsistency doctrine”), affects the agency’s interpretive choice. Inconsistency doctrine affects interpretive choice in two ways. First, if the Conservative agency chooses the lock-in approach, it follows from (3) that inconsistency doctrine will affect how aggressive the lock-in interpretation will be. Second, from (7) and (8) it is apparent that inconsistency doctrine will influence the Conservative agency’s decision whether to lock in its interpretation or to take a riskier interpretive approach that will provoke a future Liberal agency to attempt reversal.

On the first point, regarding the aggressiveness of the Conservative agency’s interpretation if the agency opts for the lock-in approach, an increase in the probability that the Judiciary will reject an inconsistent agency interpretation makes the optimal lock-in approach more conservative. Formally:

\[ \frac{\partial x^T}{\partial \alpha} = -\frac{k}{-\alpha^2} < 0 \]  \hspace{1cm} (10)

The intuition here is straightforward: As the Judiciary becomes less likely to accept a revised interpretation, attempting such a revision becomes less appealing to the Liberal agency in the second period, because if the Judiciary rejects the attempted revision the Liberal agency will have incurred the cost \( k \) for no benefit. Therefore, the Conservative
agency does not need to concede as much in the first period in order to lock in its interpretation—it can adopt a more aggressively conservative interpretation without provoking the Liberal agency to attempt reversal.

On the second point, regarding the Conservative agency’s choice of interpretative strategy, the lock-in approach becomes more appealing as the probability that the Judiciary rejects an inconsistent interpretation increases. Formally:

$$\frac{\partial Z}{\partial \alpha} = \delta - 2\alpha \delta - 1 - \delta \beta \leq 0$$  \hspace{1cm} (11)

The intuition for this follows immediately from the fact that increasing the Judiciary’s probability of rejecting an inconsistent agency interpretation allows the Conservative agency to lock in an even more aggressive interpretation without risking reversal.

We have established that decreasing judicial deference to inconsistent agency interpretations both increases the appeal of the lock-in approach and makes the lock-in approach more aggressive. The lock-in approach, however, is (weakly) less aggressive than the risky approach. Therefore, decreasing judicial deference to inconsistent agency interpretations may increase or decrease the aggressiveness of the Conservative agency’s first-period interpretation. To illustrate, consider the effect of a change in judicial doctrine from $\alpha_{\text{low}}$ to $\alpha_{\text{high}} > \alpha_{\text{low}}$. Let us further define $\alpha^T$ as the value of $\alpha$ for which the condition in (7) holds with equality (i.e., $\alpha^T$ is the minimum value of $\alpha$ that induces the Conservative agency to choose the lock-in approach). If the Conservative agency would adopt the lock-in approach both before and after the change (i.e., $\alpha_{\text{low}} > \alpha^T$), the increase in $\alpha$ will cause the agency to adopt a more aggressively conservative interpretation in the
first period. If the Conservative agency would adopt the risky approach both before and after the change (i.e., \( \alpha_{\text{high}} < \alpha^T \)), the increase in \( \alpha \) will make no difference to the Conservative agency’s first-period choice. If, however, the increase in \( \alpha \) would induce the Conservative agency to switch from the risky approach to the lock-in approach (i.e., \( \alpha_{\text{high}} > \alpha^T > \alpha_{\text{low}} \)), then the increase in \( \alpha \) will lead the Conservative agency to adopt a less aggressive first-period interpretation, as the Conservative agency’s most-preferred lock-in interpretation is (weakly) more liberal than its most-preferred risky interpretation.

This is the key positive implication of our analysis: The relationship between the level of judicial hostility to inconsistent agency interpretations (\( \alpha \)) and the aggressiveness of the first-period interpretation (\( x_1^* \)) is non-monotonic. When the judiciary is as deferential to a revised interpretation as to an initial interpretation (\( \alpha=0 \)), the Conservative agency selects the most conservative permissible interpretation (\( x_1^*=0 \)) and the Liberal agency always successfully replaces it with the most liberal permissible interpretation (\( x_2^*=1 \)). As \( \alpha \) increases from 0 to \( \alpha^T \), the Conservative agency continues to select the most conservative permissible interpretation (\( x_1^*=0 \)); the Liberal agency continues to attempt to change the interpretation to the most liberal interpretation permissible, but there is now a positive probability that the court will reject that attempt (implying that \( x_2=1 \) with probability \( 1-\alpha \) and \( x_2=0 \) with probability \( \alpha \)). Once \( \alpha=\alpha^T \), the Conservative agency switches from the risky approach to the lock-in approach, which leads to a more moderate first period interpretation (\( x_1^*=(1-(k/(1-\alpha))>0) \); this

---

\(^8\) This assumes that \( \alpha_{\text{low}} < 1- k \); if not, then the Conservative agency can lock in the most conservative possible interpretation when \( \alpha=\alpha_{\text{low}} \), so an increase in \( \alpha \) would have no effect on the Conservative agency’s interpretation.
interpretation persists in the second period as well, as the Liberal agency no longer finds it worthwhile to attempt reversal. Further increases in \( \alpha \) allow the Conservative agency to lock in progressively more conservative interpretations. Once \( \alpha = 1-k \), the Conservative agency again chooses \( x_1^* = 1 \), and this interpretation will stick in both periods. This pattern is illustrated in Figure 1.

Although our main focus concerns changes in \( \alpha \), it is also important to consider the impact of changes in \( k \), the fixed cost to the Liberal agency of attempting an interpretive revision. After all, as noted in Part I, a number of scholars and some judicial opinions have suggested that the correct approach for dealing with revised agency interpretations is not to reduce the probability of deferring to such revisions, but rather to impose additional explanatory or procedural costs on agencies that change their prior interpretive
position. It turns out that the results for changes in $k$ are qualitatively similar to the results above for changes in $\alpha$. Formally:

$$\frac{\partial x^T}{\partial k} = -\frac{1}{1-\alpha} < 0;$$

(12)

$$\frac{\partial Z}{\partial k} = -\left(\epsilon + \delta - \alpha \delta \beta \right) \leq 0.$$  

(13)

This implies that as the fixed cost to the agency of revising a previous interpretive decision increases, the first-period agency can lock in a more aggressive interpretation, and the attractiveness to the first-period agency of this lock-in strategy increases as a result. This leads to an analogous non-monotonicity result: Increasing the cost of revisiting a prior agency interpretation may lead to a more moderate initial interpretation, if this increase causes the first-period agency to switch from the risky approach to the lock-in approach, but further increases in this fixed cost will cause the first-period interpretation to become progressively more extreme.

III. Welfare Implications

The preceding section considered how inconsistency doctrine affects the ideological extremism of agency interpretations. What are the normative implications, if any, of this analysis? A full normative discussion would be a complex enterprise well beyond the scope of this paper. As a preliminary contribution to this larger endeavor, we consider two normative benchmarks: the utility of the two political parties, and the welfare of a hypothetical median voter (whom we will refer to simply as the “voter”).
A. Parties’ Welfare

Decreasing deference to revised agency interpretations always makes the Conservative agency better off.\footnote{This follows from the fact that $EU_c \text{ risky} = -\delta \left( -\alpha + \delta \beta \right)$ and $U_c \text{ lock in} = -\alpha + \delta \left( 1 - \frac{k}{1-\alpha} \right)$ are both increasing in $\alpha$, and the Conservative agency always chooses the higher of the two.} What about the Liberal agency? If the Conservative agency chooses the risky approach in the first period, the Liberal agency will choose $x_2^* = 1$ in the second period, giving the Liberal agency the following total utility:

$$EU_L \text{ risky} = -(1 + \delta \alpha + k)$$ (14)

If, on the other hand, the Conservative agency adopts the lock-in approach in the first period setting $x_1^* = x^r$, the Liberal agency will not alter this interpretation in the second period. This means the Liberal agency’s total utility will be:

$$U_L \text{ lock in} = -(1 + \delta \left( 1 - x^r \right)) = \frac{(1 + \delta)k}{1-\alpha}$$ (15)

Note that the Liberal agency’s expected utility from the Conservative agency’s choice of the risky approach and the lock-in approach decreases as we decrease the deference to revised agency interpretation (i.e. as we increase $\alpha$). But this does not mean that decreasing the deference to revised agency interpretation always decreases the Liberal agency’s utility. To see this formally, define $\tilde{\alpha}^r = \alpha^r - k \left( 1 - \delta \beta \right) / \delta$.\footnote{$\tilde{\alpha}^r$ is defined such that $EU_L \text{ risky, } \tilde{\alpha}^r = U_L \text{ lock in, } \tilde{\alpha}^r$. See appendix.} For any
\( \alpha' \in [\bar{\alpha}^T, \alpha^T] \), we can define \( \alpha'' > \alpha^T \) such that the Liberal agency would prefer any \( \alpha \in [\alpha^T, \alpha'''] \) to \( \alpha' \), although \( \alpha > \alpha' \).\(^{11}\) This result is illustrated in Figure 2.

**Figure 2: The Liberal Agency’s Expected Utility**

In Figure 2 we see two functions. The linear function represents the Liberal agency’s utility when the Conservative agency chooses the risky approach (expression (14)). The curve represents the Liberal agency’s utility when the Conservative agency chooses the lock-in approach (expression (15)). The Liberal agency’s actual expected utility (the bold line in Figure 2) is the linear function for \( \alpha < \alpha^T \), and it shifts to the curve for \( \alpha \geq \alpha^T \), since at \( \alpha^T \) the Conservative agency shifts from the risky approach to the lock-in approach. It is clear from Figure 2 that the agency’s expected utility is higher at \( \alpha \in [\alpha^T, \alpha'''] \) than at \( \alpha' \).

---

\(^{11}\) \( \alpha'' \) is defined such that \( EU_L \) risky, \( \alpha' = U_L \) lock in, \( \alpha'' \). Formally, \( \alpha'' = 1 - \frac{(1 + \delta)k}{1 + \delta \alpha' + k} \). See appendix.
Intuitively, the Conservative and Liberal agencies have zero-sum preferences on the statutory interpretation that applies. At $\alpha^T$ the Conservative agency moves from the risky approach to the lock-in approach, but its utility from both approaches is equal. However, moving to the lock-in approach increases the Liberal agency utility, since it does not have to bear the cost of revising the Conservative agency’s interpretation ($k$). Thus, the Liberal agency would enjoy a decrease in the deference to revised agency interpretation if this decrease results in a shift from the risky approach to the lock-in approach, and the savings on the cost of revising the Conservative agency’s interpretation are greater then its decrease in utility from a more conservative interpretation. This analysis shows that decreasing the deference to revised agency interpretation could be a Pareto improvement for the Conservative and the Liberal parties.

B. Median Voter Welfare

Perhaps what would benefit the political parties is less important than what would benefit the citizenry more generally. We therefore consider, as an alternative normative benchmark, the welfare of the median voter in the electorate. Median voter welfare is problematic as a normative standard, both because the ideal policy of the median voter may not be the policy that maximizes aggregate social welfare (Stiglitz 2000), and because the majoritarianism implicit in catering to the median voter may be insufficiently sensitive to minority interests. And, of course, for multi-dimensional issues, there may not be a “median voter,” nor indeed any coherent way to describe what the “majority prefers” (Arrow 1951; Riker 1982; Shepsle 1992). Nonetheless, since there are some
situations in which the notion of accountability to electoral majorities is both conceptually coherent and normatively relevant, and because much administrative law doctrine proceeds on the assumption that this is often the case, we explore how inconsistency doctrine might affect median voter welfare. Formally, we define the voter’s ideal point in period 1 as \( x_{v1} \), and her ideal point in period 2 as \( x_{v2} \). For simplicity, we assume the voter has the same discount factor, \( \delta \), as the political parties. The voter’s welfare is therefore:

\[
U_v = -|x_{v1} - x_1| - \delta |x_{v2} - x_2|
\]  \hspace{1cm} (16)

We explore two plausible though simplified conjectures about the relationship between the voter’s ideal interpretation and the interpretation favored by the party currently in power. On one interpretation, the voter’s ideal point is equal to the ideal point of the incumbent party (i.e., \( x_{v1}=0 \) and \( x_{v2}=1 \)). The assumption that the Conservative party controls the agency in the first period, on this view, derives from the fact that the voter in the first-period has more conservative preferences. The Liberal party will take over when the voter’s preferences have shifted to the left. Under this assumption, the voter’s expected utility is:

\[
EU_v = \begin{cases} 
-\alpha & \text{if } Z > 0 \\
-x^T - \delta_v x_T & \text{if } Z \leq 0
\end{cases}
\]  \hspace{1cm} (17)

It is apparent from (17) that the voter in this case is best off if the Judiciary always accepts a revised interpretation (\( \alpha=0 \)). That is intuitive, as the voter is assumed to want whichever party is currently in power to implement its most-preferred interpretation. A

\footnote{Note that when \( \alpha=0 \) we get \( Z>0 \).}
judicial doctrine of $\alpha=0$ maximizes the *responsiveness* of agency interpretations to the views of the incumbent administration, which is what the voter would want if the incumbent’s views reliably track the voter’s own.

But we might reasonably suppose that the administration’s ideal point does not perfectly track the voter’s ideal point (cf. Stephenson 2008). Indeed, it is plausible that a great deal of the difference between the preferred interpretations of Liberal and Conservative administrations is not due to shifts in voter preferences on that particular interpretive issue, but rather due to the influence of ideological extremists or parochial interests groups, or perhaps to the leadership’s own policy agenda. This is especially likely since the voter does not elect each agency head independently, but rather elects a single President with a bundle of policy positions (Berry and Gersen 2008). On this view, the voter’s ideal interpretation is likely to lie somewhere in between the ideal interpretations of the two parties. For simplicity, let us suppose that the voter’s ideal interpretation in both periods is midway between the ideal interpretations of the Liberal and Conservative parties (i.e., $x_{v1}=x_{v2}=\frac{1}{2}$). In this case, the voter’s utility if the Conservative agency adopts the lock-in approach is:

$$U_v, lock in = -1 + \delta \left| \frac{1}{2} - x^T \right|.$$  \hspace{1cm} (18)

The voter’s expected utility if the Conservative agency adopts the risky approach is:

$$EU_v,isky \geq -\xi + \delta \left( \frac{1}{2} \right)$$ \hspace{1cm} (19)

It follows that the Voter is weakly better off if the Conservative agency chooses a lock-in approach. The reason is that in this case the voter is indifferent between the Liberal and Conservative parties’ most-favored interpretative approaches; the risky
approach guarantees an extreme interpretation in both periods, which the voter dislikes. The lock-in approach, on the other hand, guarantees a more moderate interpretation in both periods. Therefore, the voter would always prefer $a^T$ to any $a$\<sup>$<$</sup> $a^T$. Furthermore, if $a\geq a^T$, the voter’s expected utility is decreasing in $a$. This follows from the facts that $x^T$ is decreasing in $a$ and that $x^T \leq 1/2$ for any $a\geq a^T$. Therefore, when the voter’s ideal interpretation is at the midpoint between the Conservative and Liberal parties’ ideal interpretations, the voter would most prefer that the Judiciary reject a changed a revised agency interpretation with probability $a^T$, as this judicial doctrine maximizes the moderation of the agency’s interpretation in both periods.

Both of the above normative analyses are clearly unrealistic, even if one restricts the inquiry to the median voter welfare. The former case supposes that the incumbent’s ideal interpretation is perfectly correlated with the median voter’s ideal interpretation, while the latter case supposes that both that the incumbent’s ideal interpretation has no correlation whatsoever with the voter’s ideal interpretation. Yet these simple polar cases illustrate the considerations that would inform a more nuanced and realistic analysis. The tighter the correlation between voter preferences and agency preferences, the more the voter would prefer that courts defer to revised agency interpretations, as this approach maximizes interpretive responsiveness to the incumbent party’s ideology. But, insofar as the agencies are likely to be more ideologically extreme than the median voter, the more attractive some degree of judicial hostility to revised agency interpretations becomes. In our model, this is not because the voter has some intrinsic interest in interpretive consistency—though that might be a real consideration in some circumstances (Stephenson 2006a)—but rather because a reduction in deference to changed agency
interpretations might induce greater interpretive moderation. Appropriate normative analysis, then, must be attentive to how this aspect of judicial doctrine affects the trade-off between responsiveness and moderation in agency interpretation.

CONCLUSION

Modern administrative law doctrine has struggled with the question whether courts ought to be less deferential to an administrative agency’s interpretation of a statute when that interpretation differs from the agency’s own prior construction of the same provision. The appropriate resolution of this thorny doctrinal problem must take into account a number of factors beyond the scope of this paper, including the need to respond to changed circumstances and the “rule of law” interest in administrative consistency. Nonetheless, any complete analysis must also consider how this inconsistency doctrine affects the expected substantive ideological content of agency statutory interpretations. Our analysis has demonstrated that this relationship is more subtle and complicated than it might initially appear.

Most importantly, there is a non-monotonic relationship between the probability that a court will accept a revised interpretation and the ideological extremism of the original interpretation. This occurs because the first agency to interpret a statute must choose between two approaches: it can take a risky approach, advancing an extreme interpretation that it can anticipate its political opponents will try to undo as soon as they take power, or it can “lock in” an interpretation that is sufficiently moderate that its political opponents would not bother trying to revise it, given the inherent costs of
making such an attempt. When courts are less deferential to revised interpretations, the lock-in approach becomes more attractive because the original agency can lock in a more favorable interpretation without fear of attempted reversal. So, while decreasing deference to revised interpretations can sometimes make the original interpretation more extreme, such a decrease can make the original interpretation more moderate if it induces the initial agency to switch from the risky approach to the lock-in approach.

The normative ramifications of this positive observation depend crucially on the social objective function. If the goal is to increase the responsiveness of interpretive choices to the party currently in power—on the logic that electoral victory signifies representativeness of citizen preferences—then one would prefer that judicial deference to revised agency interpretations be no different than judicial deference to initial agency interpretations. If, however, one believes that political parties’ interpretive preferences tend to be extreme relative to a majority of the electorate—that is, if the variance in the ideal points of political leaders over time tends to be much larger than the variance in the ideal point of the median voter over time—or if one is interested in advancing the joint welfare of the party in power and the party out of power, then a moderately lower level of deference to revised interpretations, or somewhat greater fixed costs to an agency of revising a previous interpretation, can be welfare-enhancing.
At $\alpha^*$ the Conservative agency’s utility from the risky approach and the lock-in approach are equal. Therefore:

$$\frac{\zeta + \delta R}{1 - \alpha^*} = 1 + \delta \alpha^* + \delta \beta k$$  \hspace{1cm} (A1)

The Liberal agency’s utility from the conservative agency choosing the lock in approach is defined in expression (15). We can use (A1) to define this utility at $\alpha^*$:

$$U_L \text{lock in, } \alpha^* = \frac{-\zeta + \delta k}{1 - \alpha^*} = -\zeta + \delta \alpha^* + \delta \beta k$$  \hspace{1cm} (A2)

Now, we would like to find $\tilde{\alpha}^*$ such that the Liberal agency’s utility from the Conservative agency’s choice of the risky approach when $\alpha = \tilde{\alpha}^*$ is equal to the Liberal agency’s utility from the Conservative agency’s choice of the lock in approach when $\alpha = \alpha^*$ (i.e. $U_L \text{ risky, } \tilde{\alpha}^* = U_L \text{ lock in, } \alpha^*$). Therefore, using expressions (14) and (A2):

$$-(1 + \delta \tilde{\alpha}^* + k) = -(1 + \delta \alpha^* + \delta \beta k)$$

or

$$\tilde{\alpha}^* = \alpha^* - k \frac{1 - \beta}{\delta}$$  \hspace{1cm} (A3)

At $\alpha' \in [\tilde{\alpha}^*, \alpha^*]$, the Liberal agency’s utility from the Conservative agency choosing the risky approach is:

$$EU_{L \text{ risky, } \alpha'} = -(1 + \delta \alpha' + k)$$  \hspace{1cm} (A4)
We would like to find $\alpha''$ such that the Liberal agency’s utility from the Conservative agency’s choice of the lock-in approach when $\alpha = \alpha''$ is equal to the Liberal agency’s utility from the Conservative agency’s choice of the risky approach when $\alpha = \alpha'$ (i.e., $U_{L, \text{risky}, \alpha'} = U_{L, \text{lock in}, \alpha''}$). Accordingly:

$$-(1 + \delta \alpha' + k) = \frac{(1 + \delta)k}{1 - \alpha''}$$

or

$$\alpha'' = 1 - \frac{(1 + \delta)k}{1 + \delta \alpha' + k} \quad (A5)$$
REFERENCES


