

Flip-flopping and Electoral Concerns

Giovanni Andreottola *

October 2020

*CSEF and Federico II University of Naples,
Department of Economics and Statistics, Via Cintia - Monte Sant'Angelo, 80126 Naples, Italy.
Email: Giovanni.Andreottola@gmail.com Website: www.giovanniandreottola.com

Abstract

I consider a model of political agency in which an incumbent politician takes two decisions in sequence, having private information about the correct decision in each period. I find that policy reversals signal that a politician is incompetent. This suggests that the stigma associated with flip-flopping can be rationalized as a simple heuristic to select competent politicians. Furthermore, the negative impact of policy reversals on reputation gives the incumbent politician an incentive to disregard new information and stick to his initial policy. When reputation concerns are sufficiently strong, these incentives inhibit policy responsiveness, resulting in an inefficient degree of policy persistence. Term-limited politicians, or politicians with particularly safe (or compromised) election prospects, are more likely to reverse decisions. Relatedly, politicians are only willing to reverse decisions after a sufficient amount of time has passed, or when there is sufficient evidence that the state of the world might have changed.

Supplementary material for this article is available in the appendix to the online edition.

Keywords: flip-flopping; elections; political agency; accountability; reputation; persistence; policy change.

The accusation of being a flip-flopper following the reversal of a policy decision has a long history in politics. In a changing environment that requires revising and updating decisions in response to the arrival of new information, however, the stigma faced by politicians proposing policy reversals appears puzzling. Therefore, one might be prone to dismiss the preference for consistent politicians as the result of voters' gut instincts, according to which strong leaders should "stick to their guns".

This paper, on the contrary, shows that this type of voter behavior can be rationalized as a simple heuristic to select competent candidates. Since competent politicians are more likely to be consistent, voters infer incompetence from inconsistency. Flip-flopers are thus penalized not because of an intrinsic aversion to policy reversals, but because of the bad policy-making associated with flip-flopping. The reason behind this is that competent policy-making, being based on accurate information, results in less frequent policy reversals than incompetent policy-making based on noisier information.

Speaking of one of the politicians who have most famously suffered from the accusation of being a flip-flopper in recent American political history, the following remark by journalist Mickey Kaus (2004) effectively reflects the idea at the heart of this paper: *But with Kerry the charge isn't that he's inconstant. It's that in his inconstancy he flips wrong – the far more serious charge of bad judgement.*

Having provided a rationale for the stigma associated with flip-flopping, a further contribution of my paper is to explore its consequences. Given that policy reversals signal incompetence, politicians have an incentive to disregard new information in order to appear competent. This can inhibit policy responsiveness and result in an inefficient degree of policy persistence. In other words, my model has the prima facie paradoxical impli-

cation that even though voters dislike flip-floppers, they would benefit from politicians flip-flopping more.

In particular, politicians have the strongest incentive to avoid policy reversals when their re-election prospects are more sensitive to changes in reputation: conversely, a testable implication of my model is that politicians facing a binding term-limit or – more generally – politicians with very strong (or weak) electoral positions, should be more prone to reverse their previous decisions¹.

Along with displaying a diminished policy responsiveness, politicians who “stick to their guns” for reasons of reputation also hamper the ability of voters to learn about their competence. Hence, flip-flopping avoidance is also detrimental to the selection of competent candidates through elections. The reason is that when politicians facing competitive elections posture consistency, their track-records become void of information on their competence. Given this, an indirect implication of my model is that in order to draw information from the consistency of politicians, it is sometimes necessary for voters to look further back into the history of their decisions. This can explain the fact that sometimes politicians are haunted for a long time by flip-flops which occurred relatively far away in the past.

Having described the consequences of the penalization of politicians who perform policy reversals, it is worth to consider more in depth the issue of why voters use consistency to draw inferences on the competence of candidates. There are two key issues, which also shed light on the main features of the model that I use to deliver my results. The first concerns why track records are informative of a politician’s competence: the answer

1. List and Sturm (2006) show that term-limited governors are more likely to reverse policies that they have previously enacted.

is that politicians are privy to policy relevant information, the quality of which reflects their competence. Not being able to directly observe the information in the hands of politicians, voters look at their policy choices to draw inferences on the politicians' competence. This feature is present in virtually all kinds of policy-making, but it makes the model particularly apt to describe policy areas featuring confidential information, a lack of transparency or a high technical complexity. Foreign policy and financial regulation, for example, fit this description very closely.

The second issue concerns why politicians are indirectly evaluated in terms of their consistency rather than directly in terms of the success or failure of their policies: the reason for this is that voters assign politicians a reputation without being capable (or before being capable) of directly judging their policy choices. This feature is typical of policy areas where the effects of decisions take time to unravel, but it is also consistent with the environments described above, in relation to private information and complexity. As a matter of fact, when policy-relevant information is private or too complex to access, it is also harder for voters to assess whether a politician took the right decision. In this respect, the advantages of the “consistency heuristic” are very clear: compared to information on the merit of a political decision, which can be very difficult to obtain and interpret and which can also be biased or misleading, information on what a politician decided is usually readily available and easily fact-checkable.

My model also speaks to fairly abundant empirical and experimental evidence showing that voters value consistency as a trait in their political representatives. In many instances, inconsistency is negatively evaluated independently of the direction of the policy change and the partisanship of the flip-flopping politician². Among the several

2. See for example Allgeier et al. (1979), Carlson and Dolan (1985), Tomz and Van Houweling (2012b),

reasons reported by survey respondents to motivate the negative evaluation of inconsistent decision-makers, a recurring one is that inconsistent decision-makers are more likely to be incompetent³.

Along with competence being among the reasons directly mentioned to motivate the stigma associated with inconsistency, there is also indirect evidence that a lack of consistency is more damaging to a politician when it signals a lack of competence.

For example, the longer is the time elapsed between a decision and its reversal, the less negative is the evaluation of a flip-flop. The reason for this is that competent politicians do sometimes change their mind, but they are unlikely to do so very shortly after taking a decision⁴. Similarly, policy reversals on complex issues, where new information is more likely to sway the optimal decision, are evaluated less badly than flip-flops on simpler issues⁵. Finally, it has also been shown that the presence of supporting information on the necessity of a policy change improves the evaluation of an inconsistent policy-maker⁶: in other words, if a politician can demonstrate that a policy reversal was the right decision, voters are more forgiving.

In this respect, the contribution of my work is to provide a simple and flexible formalization that captures many of the “stylized facts” highlighted by empirical studies. In particular, my model emphasizes that a fundamental role to understand the consequences

Doherty, Dowling, and Miller (2015), Robison (2017). Please refer to the literature review section for additional details on empirical findings on flip-flopping.

3. The common charges of indecisiveness or weakness can also be related to competence. Incompetence as a reason for the negative evaluation of flip-floppers is reported by Tomz and Van Houweling (2012b), Levendusky and Horowitz (2012) among others.

4. A famous quote by Congressman Jack Flynt goes: *“I can vote conservative or I can vote liberal, and my constituents don’t care. But if vote one way at two o’clock and another way at two-thirty, they’ll think they sent an idiot to Congress”*. See also Doherty, Dowling, and Miller (2015) for an analysis of how time affects the perception of a flip-flop.

5. See again Doherty, Dowling, and Miller (2015) and Tavits (2007) for a related point on principle versus pragmatic issues.

6. See Levendusky and Horowitz (2012) and Robison (2017).

of policy changes is played by the persistence of the state of the world, that is the likelihood of the correct policy approach to remain the same across periods. Persistence can be interpreted as being a function of several key aspects of a policy-making environment, such as i) time between decisions (the shorter the time between decisions, the higher persistence); ii) the type, complexity and similarity of the issues decided upon; iii) the degree of public information on whether a change in the environment occurred. With respect to these elements of the decision-making environment, an implication of my model is that because of electoral incentives, the quality of decision-making and the informativeness of a politician's track-record are non-monotonic in the level of persistence. On top of the positive implications discussed above, this can also have normative, institutional design implications with respect to the timing of decision-making, the types of tasks to assign to reputation concerned decision-makers and the degree of transparency over the environment in which a decision-maker acts.

Literature Review

Literature on Consistency and Flip-flopping

There is a long-standing literature investigating the effects of consistency in shaping an agent's evaluation. Allgeier et al. (1979) is the seminal study introducing the so called *waffle phenomenon*: attitude changes are perceived negatively, even if the change increases the similarity with the evaluator. Carlson and Dolan (1985) reach analogous results in a similar experiment specifically involving political candidates. Despite not finding direct evidence of the waffle phenomenon, Hoffman and Carver (1984) find that agreement with the decision-maker following an inconsistent track-record is not rewarded by voters⁷.

Recent research has further investigated the conditions leading flip-flopping politicians to be punished by voters. For example, Doherty, Dowling, and Miller (2015) present findings consistent with the predictions of my model concerning the effects of time and issue complexity on the evaluation of flip-flops. Relatedly, Tavits (2007) uses election data to show that flip-flops on pragmatic issues (where the state can change) are seen less badly than flip-flops on issues of principle (where by definition it stays fixed). Robison (2017) emphasizes that consistency is more relevant in the absence of direct information on the quality of the action taken by the politician, which mirrors the point my paper makes about consistency as indirect information about the politician's competence. Another important reference is Tomz and Van Houweling (2012b), who find that policy changes are costly not only because of the ideological ambiguity they cause, but also because of

7. Similarly, Croco (2016) finds that alignment with the decision-maker tends to prevail over the judgment of consistency, but she finds an effect of consistency among voters who are neutral on the issue at stake.

the negative inference voters draw about a candidate's character, of which competence constitutes an important component⁸. Sigelman and Sigelman (1986) and Sorek, Haglin, and Geva (2018) also highlight the close link between consistency and the perceived competence of a politician, the latter showing that the more a leader is seen as competent, the better is his reputation following a flip-flop, which reflects the results of my model. With respect to this branch of the literature, my model aims to provide a simple formalization capable of matching several stylized facts.

There are also some theoretical papers broadly related to the idea of flip-flopping: Hummel (2010) explores campaign platform choice in two-stage elections, with an exogenous cost of platform changes. Similarly, in Agranov (2016) candidates undertake costly electoral posturing to appear liberal (in the primary election) or moderate (in the general). These models have therefore a very different focus compared to my work.

Experts and Reputation Concerns

The electoral concerns model that I consider builds on the seminal papers of Canes-Wrone, Herron, and Shotts (2001), Maskin and Tirole (2004) and Prat (2005). In this respect, my contribution is to study a two-period political-agency framework, in order to address the question concerning the penalization of flip-flopping⁹. My work is also related to models of repeated expert behavior, including Prendergast and Stole (1996), Majumdar and Mukand (2004), Li (2007) and Aghion and Jackson (2016). Along with several differences in the model structure, my contribution revolves around the focus on a changing but persistent environment.

8. Related studies by the same authors, with similar findings, include Tomz and Van Houweling (2012a) and Tomz and Van Houweling (2014).

9. Other political-agency models I draw on are Ashworth and Shotts (2010) and Fu and Li (2014).

Audience Costs

Despite being more general and not specifically targeted to international relations, my model is also related to the idea of *audience costs*, i.e. that leaders who turn back on previously made threats suffer a reputation loss (the seminal contribution is Fearon (1994)). Tomz (2007) shows that audience costs are deeply intertwined with considerations about reputation and competence. Schultz (2001) also supports the idea that leaders who back down are perceived as less competent. Along similar lines, a key result in Levendusky and Horowitz (2012) is to show that a leader changing his or her mind is seen as less competent than one who stays coherent, with partisanship playing hardly any role¹⁰.

10. A theoretical model on conflict sharing some insights with my analysis is Smith (1998). In his model, however, reputation concerns make the behavior of the politician more hawkish rather than more persistent as in my model; on top of that, in his model threats are never reversed in equilibrium.

The Model

There are two periods $t \in \{1, 2\}$; an incumbent politician takes a binary action $a_t \in \{0, 1\}$ in each period on behalf of a representative voter. The payoff to the representative voter is 1 if the action a_t matches the state of the world $\omega_t \in \{0, 1\}$, and zero otherwise. The commonly known prior probability that ω_1 takes value 0 or 1 is equal to $\frac{1}{2}$. This simplifies the analysis but it is by no means crucial for the results. The state of the world is persistent, so that $Pr(\omega_2 = \omega_1 | \omega_1) = \gamma > \frac{1}{2}$. The persistence parameter γ , which as we will see is of great importance to the model, can be thought of as representing how informative past signals (and thus past policy choices) are for future policy choices. In this respect, it can also be thought of as measuring the rate of depreciation of the information guiding policies (γ close to $\frac{1}{2}$ represents a high rate of depreciation).

Incumbents can be of two types $\theta \in \{H, L\}$, i.e. competent and incompetent. The incumbent's type is fixed across periods. The probability that an incumbent is competent is $\lambda \in (0, 1)$, which is common knowledge. Incumbents do not observe their type¹¹. In each period, the incumbent observes a realization $s_t \in \{0, 1\}$ of a binary signal about the state of the world: the accuracy of the signal, $Pr(s_t = \omega_t | \omega_t) = q_\theta$, depends on the politician's type: $\frac{1}{2} \leq q_L < q_H \leq 1$. The signal is private information of the incumbent and s_1 and s_2 are independent conditional on the realization of the state of the world ω_t . The incumbent's private signals are the only information available about the state of the world¹².

11. This assumption simplifies the analysis but the results remain qualitatively unchanged under the alternative assumption that incumbents know their type: in fact, previous drafts of the paper were based on the assumption of privately observed competence.

12. In the Online Appendix I consider the scenario in which, in addition to the incumbent's private signals, the state of the world ω_1 is revealed to the public after the first action a_1 has been taken. This can be thought of as public feedback on the policy implemented in the first period, or as the result of a journalistic investigation.

The reputation of the incumbent at the end of $t = 2$ corresponds to the belief held by the representative voter over his competence: I denote it by $r(a_1, a_2) = Pr(\theta = H|a_1, a_2)$, highlighting the fact that it depends on the actions a_1 and a_2 . In addition to utility from the policy actions taken by the incumbent, the representative voter benefits from the ability to select competent politicians, for example through elections: for maximum generality, I model this in the form of a strictly increasing and strictly convex function $\nu(\cdot)$ of the incumbent's reputation¹³. To sum up, the representative voter's utility is:

$$U_v = \sum_{t=1}^2 \mathbb{1}_{a_t=\omega_t} + \delta\nu(r(a_1, a_2)) \quad (1)$$

where δ measures the welfare weight of selection relative to that of the quality of the policies chosen by the incumbent. I therefore refer to $\sum_{t=1}^2 \mathbb{1}_{a_t=\omega_t}$ and to $\delta\nu(r(a_1, a_2))$ as the policy component and the selection component of the voter's welfare respectively.

The utility of the incumbent depends both on the probability of matching the state of the world and on the reputation earned at the end of the two periods¹⁴. The incumbent politician's objective function is hence as follows:

$$U_p = \sum_{t=1}^2 \mathbb{1}_{a_t=\omega_t} + \phi r(a_1, a_2) \quad (2)$$

where ϕ measures the weight of the reputation concern: $\phi = 0$ represents the case of a purely policy-motivated incumbent, whereas $\phi \rightarrow +\infty$ represents the case of a fully

13. If the incumbent is subject to re-election in the presence of a uniformly distributed valence shock, $\nu(\cdot)$ would take the quadratic form. For more details, please refer to the Online Appendix.

14. Just like the utility of the representative voter, this can be shown to be the reduced form of a game ending with an election with a uniformly distributed valence shock. In this respect, the use of a linear function is not without loss of generality, but has been chosen for tractability and follows the vast majority of the career concerns literature.

office-motivated incumbent. As a result of Bayes' Rule, the ex-ante expected value of reputation is equal to the prior belief λ : consequently, social welfare coincides with the welfare of the representative voter.

The incumbent's strategy is, in each time period $t \in \{1, 2\}$, a mapping from the private signals received (i.e. s_1 at $t = 1$ and (s_1, s_2) at $t = 2$) and the actions taken until that period to a probability distribution across the available actions $a_t \in \{0, 1\}$. This also means that s_1 and a_1 matter for the mapping from s_2 to a_2 .

The equilibrium concept I use is Perfect Bayesian Equilibrium (PBE). Notice that at least for sufficiently high values of reputation concerns ϕ , in the set of PBEs there are always uninformative equilibria that are not interesting for the sake of my analysis¹⁵. In order to rule out these equilibria, I focus on equilibria that are robust to the introduction of a fraction $\epsilon \approx 0$ of incumbents who do not value reputation, i.e. for whom $\phi = 0$ in their utility function given by condition (2). Such incumbents, who can be competent or incompetent with the same probability as reputation-concerned incumbents, always act to maximize the probability of matching the state of the world. Henceforth, equilibrium will indicate a PBE at the limit $\epsilon \rightarrow 0$, and incumbent will indicate, unless specified, a reputation-concerned incumbent.

15. An example would be an equilibrium where reputation is zero unless the actions played are $(a_1, a_2) = (0, 0)$ and the incumbent always plays $(0, 0)$. For sufficiently high ϕ , this would be an equilibrium, though based on unreasonable off-equilibrium beliefs. Notice that other candidate equilibria with "pathological" characteristics are ruled out by the fact that incumbents are not fully office motivated.

Results

I start the analysis of the model with the information available to the incumbent. After receiving the first signal realization s_1 , the incumbent assigns the probability $Pr(\omega_1 = s_1|s_1) = \rho_1$ to the state of the world being equal to the signal¹⁶. It can immediately be seen that since $q_H > q_L \geq \frac{1}{2}$, it holds that $\rho_1 > \frac{1}{2}$: in particular, $\rho_1 = \lambda q_H + (1 - \lambda)q_L$, that is the average incumbent's signal accuracy. Notice that, given the symmetric prior, ρ_1 is the same for both $s_1 = 0$ and $s_1 = 1$. Analogously, the posterior probability that the state of the world is equal to the second realization s_2 only depends on whether $s_2 = s_1$ or $s_2 \neq s_1$ and not on the realizations of s_1 and s_2 . In other words, what matters is whether the signal realizations observed by the incumbent are consistent (subscript c), or whether they flipped (subscript f). In light of this, I denote the posteriors $Pr(\omega_2 = s_2|s_1, s_2 = s_1)$ and $Pr(\omega_2 = s_2|s_1, s_2 \neq s_1)$ as $\rho_{2,c}$ and $\rho_{2,f}$ respectively¹⁷.

16. The posterior ρ_1 can be written as follows:

$$\rho_1 = \frac{\frac{1}{2}[\lambda q_H + (1 - \lambda)q_L]}{\frac{1}{2}[\lambda q_H + (1 - \lambda)q_L] + \frac{1}{2}[\lambda(1 - q_H) + (1 - \lambda)(1 - q_L)]} = \lambda q_H + (1 - \lambda)q_L, \quad (3)$$

where $\lambda q_H + (1 - \lambda)q_L = Pr(s_1|\omega_1 = s_1)$ and analogously $\lambda(1 - q_H) + (1 - \lambda)(1 - q_L) = Pr(s_1|\omega_1 \neq s_1)$.

17. The following holds:

$$\rho_{2,c} = \frac{\frac{1}{2}A_c}{\frac{1}{2}A_c + \frac{1}{2}B_c} \quad (4)$$

where $A_c \equiv Pr(s_2 = s_1|\omega_2 = s_2, s_1)$ and $B_c \equiv Pr(s_2 = s_1|\omega_2 \neq s_2, s_1)$ which can in turn be written out as follows:

$$\begin{aligned} A_c &= \lambda[\gamma q_H^2 + (1 - \gamma)q_H(1 - q_H)] + (1 - \lambda)[\gamma q_L^2 + (1 - \gamma)q_L(1 - q_L)] \\ B_c &= \lambda[\gamma(1 - q_H)^2 + (1 - \gamma)q_H(1 - q_H)] + (1 - \lambda)[\gamma(1 - q_L)^2 + (1 - \gamma)q_L(1 - q_L)] \end{aligned}$$

The derivation of $\rho_{2,f}$ is analogous:

$$\rho_{2,f} = \frac{\frac{1}{2}A_f}{\frac{1}{2}A_f + \frac{1}{2}B_f} \quad (5)$$

where again I use $A_f \equiv Pr(s_2 \neq s_1|\omega_2 = s_2, s_1)$ and $B_f \equiv Pr(s_2 \neq s_1|\omega_2 = s_2, s_1)$ which can be written out analogously to A_c and B_c :

$$\begin{aligned} A_f &= \lambda[(1 - \gamma)q_H^2 + \gamma q_H(1 - q_H)] + (1 - \lambda)[(1 - \gamma)q_L^2 + \gamma q_L(1 - q_L)] \\ B_f &= \lambda[(1 - \gamma)(1 - q_H)^2 + \gamma q_H(1 - q_H)] + (1 - \lambda)[(1 - \gamma)(1 - q_L)^2 + \gamma q_L(1 - q_L)]. \end{aligned}$$

Notice that $\rho_{2,c} \geq \rho_{2,f} \geq \frac{1}{2}$, with strict inequalities for $\gamma \in (0.5, 1)$. In particular, this means that even if the two signal realizations s_1 and s_2 are different, the most likely state of the world is the one coinciding with s_2 . Therefore, the probability of matching the action to the state of the world is maximized by choosing $a_t = s_t$. Along with guiding political decisions, the private signals received by the incumbent also represent the information from which the voter could update her beliefs on the incumbent's competence. Since the voter cannot directly observe the private signals, however, she can only base her updating on the actions taken by the politician. As a result, the actions of the politician are most informative when they truthfully report the private signals received by the incumbent. That is to say, $a_t = s_t$ maximizes the expected value of the selection component of the representative voter's utility as per condition (1). Since the representative voter's utility coincides with social welfare, it follows that social welfare is maximized by the candidate equilibrium in which the incumbent's optimal strategy for all s_t in both periods $t \in \{1, 2\}$ is $a_t = s_t$. I denote this equilibrium as the truthful equilibrium.

If the incumbent's incentives were solely policy-related, $a_t = s_t$ would be the dominant strategy and the truthful equilibrium would always be the outcome of the game. However, reputation concerns can distort the incumbent's behavior away from the truthful equilibrium benchmark. In order to understand how this can happen, consider the four possible track records $(a_1, a_2) \in \{(0, 0), (0, 1), (1, 0), (1, 1)\}$. I denote $(0, 0)$ and $(1, 1)$ as consistent track-records; $(1, 0)$ and $(0, 1)$ are instead flip-flopping track-records, since the incumbent changes his policy stance from period 1 to period 2. For each realization of

(a_1, a_2) , the incumbent's reputation is calculated as follows:

$$r(a_1, a_2) = Pr(\theta = H|(a_1, a_2)) = \frac{\lambda Pr((a_1, a_2)|\theta = H)}{\lambda Pr((a_1, a_2)|\theta = H) + (1 - \lambda)Pr((a_1, a_2)|\theta = L)} \quad (6)$$

As a result of condition (6), a track-record delivers a reputation higher than the prior λ as long as the probability to observe that track-record conditional on the incumbent being a high type is higher than the probability conditional on the incumbent being a low type, that is $\frac{Pr((a_1, a_2)|\theta=H)}{Pr((a_1, a_2)|\theta=L)} \geq 1$. The opposite holds for a reputation lower than the prior. Notice that given the symmetric initial prior on the state of the world ω_1 , for each type θ the probability of obtaining each of the two consistent and flip-flopping signal sequences is the same: therefore, we can simply focus on the probability of having a consistent track-record, resulting in a reputation level that I denote as r_c , versus a flip-flopping one, which yields reputation r_f .

Since the incumbent is free to choose a policy action independently of the signal he received, the probabilities $Pr((a_1, a_2)|\theta)$, and therefore r_c and r_f , crucially depend on the equilibrium strategy chosen by the incumbent. To fix ideas, I first consider the benchmark case of a truthful equilibrium, in which $(a_1, a_2) = (s_1, s_2)$, giving rise to reputations that I denote as r_c^T and r_f^T . As it will become clear later on, it is useful to work with the conditional probability of a flip-flopping signal sequence given an initial signal realization s_1 and the type of incumbent θ , which I denote as:

$$Pr(c|\theta) \equiv Pr(s_2 = s_1|s_1, \theta) = \gamma[q_\theta^2 + (1 - q_\theta)^2] + (1 - \gamma)2q_\theta(1 - q_\theta) \quad (7)$$

$$Pr(f|\theta) \equiv Pr(s_2 \neq s_1|s_1, \theta) = (1 - \gamma)[q_\theta^2 + (1 - q_\theta)^2] + \gamma 2q_\theta(1 - q_\theta) = 1 - Pr(c|\theta) \quad (8)$$

It can easily be checked that $Pr(c|\theta)$ is increasing in q_θ and, conversely, $Pr(f|\theta)$ is decreasing in q_θ . In other words, the more precise is the incumbent's signal, the less it flips. Expressions (7) and (8) allow me to rewrite condition (6), for $j \in \{c, f\}$, as:

$$r_j^T = \frac{\lambda Pr(j|H)}{\lambda Pr(j|H) + (1 - \lambda) Pr(j|L)} \quad (9)$$

Since $\frac{Pr(c|H)}{Pr(c|L)} > 1$, condition (9) implies that in the truthful equilibrium, the reputation r_c^T after a consistent track record is higher than the prior. By the same argument, the reputation r_f^T is lower than the prior, so that $r_f^T < \lambda < r_c^T$.

The intuition behind the bad reputation associated with flip-flopping is that since the state of the world is persistent ($\gamma > \frac{1}{2}$), an accurate signal has to be persistent, too. At the extremes, a perfect signal flips with probability equal to the state of the world changing, i.e. $1 - \gamma$, whereas a completely noisy signal flips with probability $\frac{1}{2}$, and $\frac{1}{2} > 1 - \gamma$. This also means that r_f^T decreases as γ increases: when the state is fully persistent, i.e. $\gamma = 1$, a flipping signal conveys that the policy choice was wrong in one of the two periods.

So far I have assumed a truthful behavior by the incumbent. This, of course, cannot be taken for granted: in order to prevent the reputation loss from a flip-flopping track-record, the incumbent can ignore the signal received in the second period and act consistently. This of course comes at a cost, both for the politician distorting his action and for the representative voter. Consider an incumbent about to take the second policy action, with the second signal realization s_2 suggesting to change the policy compared to what he had chosen in the first period. Following the signal delivers a higher probability of matching the state of the world, equal to $\rho_{2,f} > \frac{1}{2}$ versus $1 - \rho_{2,f} < \frac{1}{2}$ by acting against the signal. At the same time, however, reversing policy as suggested by the signal makes

the incumbent appear incompetent, which costs him $\phi(r_c - r_f)$ in terms of reputation loss. Hence, the incumbent has an incentive to follow his signal as long as the following holds:

$$2\rho_{2,f} - 1 \geq \phi(r_c - r_f) \quad (10)$$

For high enough reputation concerns, as a result, the incumbent has an incentive to avoid the flip-flop and posture consistency, ignoring his private information. Therefore, a truthful equilibrium is only sustainable if, everything else equal, reputation concerns ϕ are sufficiently low. Using $r_c = r_c^T$ and $r_f = r_f^T$ and letting condition (10) hold as an equality yields a threshold on ϕ , which I denote as $\phi_1 \equiv \frac{2\rho_{2,f}-1}{r_C^T - r_F^T}$.

Notice that so far I have focused my attention on the second action a_2 . The reason is that in the first period, the incumbent always follows his signal s_1 : on top of being the optimal choice to match the state of the world, following the first signal is also optimal from the point of view of reputation. The reason is that given the persistence of the state of the world, the second signal is more likely to coincide with the first than to flip: hence, by choosing $a_1 = s_1$ the incumbent is more likely not to face a trade-off between informed policy-making and reputation in the second period¹⁸.

Having established that a truthful equilibrium is not always sustainable, the natural question to ask now concerns what happens when reputation concerns exceed ϕ_1 and the incumbent does not have an incentive to always follow his signal s_2 if $s_2 \neq a_1$. To this end, denote by $\sigma \equiv Pr(a_2 = s_2 | s_2 \neq s_1, a_1 = s_1)$, that is the probability of following the signal s_2 when it suggests to change policy with respect to the first period. Notice that

18. If the prior probability of the state of the world was unbalanced in favor of one state, the first action would also affect the incumbent's reputation, hence there would be an additional trade-off in the first period when receiving the less likely signal. The aversion to flip-flopping, however, would discipline the incumbent towards following the signal in the first period.

in the truthful equilibrium, $\sigma = 1$. For a given σ , following condition (9), the reputation from a consistent track-record r_c takes the following value:

$$r_c = \frac{\lambda[1 - Pr(f|H)\sigma]}{\lambda[1 - Pr(f|H)\sigma] + (1 - \lambda)[1 - Pr(f|L)\sigma]} \quad (11)$$

Given that $Pr(c|H) > Pr(c|L)$ and $Pr(f|L) > Pr(f|H)$, condition (11) is clearly increasing in σ . Unlike r_c , the reputation from a flip-flopping track-record r_f does not depend on σ , which cancels out of the expression:

$$r_f = \frac{\lambda Pr(f|H)\sigma}{\lambda Pr(f|H)\sigma + (1 - \lambda)Pr(f|L)\sigma} = r_f^T \quad (12)$$

Notice that at the extreme value of $\sigma = 0$, the reputation of a consistent policy record reduces to $r_c = \lambda$: that is, it contains no information on the incumbent's competence, since the incumbent always acts consistently.

Suppose now that the truthful equilibrium is not sustainable but that the incumbent follows a mixed strategy after observing $s_2 \neq s_1$: that is, σ takes a value in the interior of the interval $(0, 1)$. Denote this value of σ by σ^* . Clearly, σ^* has to make the incumbent indifferent between the two possible actions, which, following condition (10), requires:

$$2\rho_{2,f} - 1 = \phi(r_c - r_f^T) \quad (13)$$

Denoting by $Pr(f) = \lambda Pr(f|H) + (1 - \lambda)Pr(f|L)$ the total probability of the signal flipping conditional on s_1 and by $k = \frac{2\rho_{2,f}-1}{\phi}$ the cost of not following the signal normalized

by the weight of reputation ϕ , the indifference condition (13) can be rewritten as:

$$\frac{1 - \sigma^* Pr(f|H)}{1 - \sigma^* Pr(f)} = \frac{k + r_f^T}{\lambda} \quad (14)$$

which finally allows me to pin down the value of σ^* after some algebraic manipulations¹⁹.

In an equilibrium with $\sigma = \sigma^* \in (0, 1)$, the mixing done by the incumbent decreases r_c , lowering the reputation premium from a consistent track-record just enough to make the incumbent indifferent between the two actions. Notice that the left-hand side of condition (14) is strictly larger than 1 as long as $\sigma > 0$: therefore, a necessary condition for the existence of a mixed equilibrium is that $k > \lambda - r_f^T$. This means that the cost k of not following the signal has to be sufficiently high, such that if $\sigma = 0$, i.e. if $r_c = \lambda$, the incumbent prefers to follow his signal. For sufficiently high values of ϕ , however, $k \leq \lambda - r_f^T$ and there is no interior mixing probability that makes the incumbent indifferent between following the signal and receiving a consistent action record. The only possible outcome is then an equilibrium with $\sigma = 0$, that is with complete flip-flop avoidance. The threshold above which this happens is denoted by ϕ_2 , which takes the following value: $\phi_2 = \frac{2\rho_{2,f}-1}{\lambda-r_f^T}$.

Equilibrium

The game has a unique equilibrium, with the following properties²⁰:

- For $\phi \leq \phi_1 = \frac{2\rho_{2,f}-1}{r_c^T - r_f^T}$, the unique equilibrium is the truthful equilibrium, that is

19. Further denoting $\chi = \frac{k+r_f^T}{\lambda}$, (14) can be solved to yield:

$$\sigma^* = \frac{\chi - 1}{\chi Pr(f) - Pr(f|H)} \quad (15)$$

20. Notice that the term “equilibrium” follows the definition provided in the Model section. For formal existence and uniqueness proofs, please refer to the Online Appendix.

$$\sigma = 1.$$

- For $\phi_1 < \phi \leq \phi_2$, the unique equilibrium displays partial flip-flopping avoidance, with $\sigma = \sigma^* \in (0, 1)$ and σ^* defined by condition (14).
- For $\phi > \phi_2 = \frac{2\rho_{2,f}-1}{\lambda-r_f^T}$, the unique equilibrium displays complete flip-flopping avoidance, that is $\sigma = 0$.

The equilibrium value of reputation following a flip-flopping track-record is always equal to r_f^T defined in condition (12); reputation following a consistent track-record is instead, given the appropriate value of σ , equal to r_c defined in condition (11). For all values of ϕ , $Pr(a_1 = s_1 | s_1) = 1$ and $Pr(a_2 = s_2 | s_2 = s_1) = 1$.

These results have several implications. The first is that even in the absence of any direct information on the quality of the policy chosen by the incumbent, his track-record provides information on his competence. Consistent incumbents are on average more competent than flip-flopping ones. Therefore, my model provides a rationale for the stigma associated with flip-flopping, pointing to the fact that the preference for consistent politicians can be thought of as a simple heuristic used by voters to select competent incumbents. This interpretation is also consistent with the empirical evidence pointing out that flip-floppers are negatively evaluated independently of partisanship and even in environments in which the optimal policy is expected to change over time²¹.

With respect to the incentives arising from the negative perception of policy changes, when reputation concerns are lower than ϕ_1 , the equilibrium outcome is equivalent to what would occur with a fully policy-motivated incumbent; when instead reputation

21. See for example Carlson and Dolan (1985) and Tomz and Van Houweling (2012b). For more details refer to the literature review.

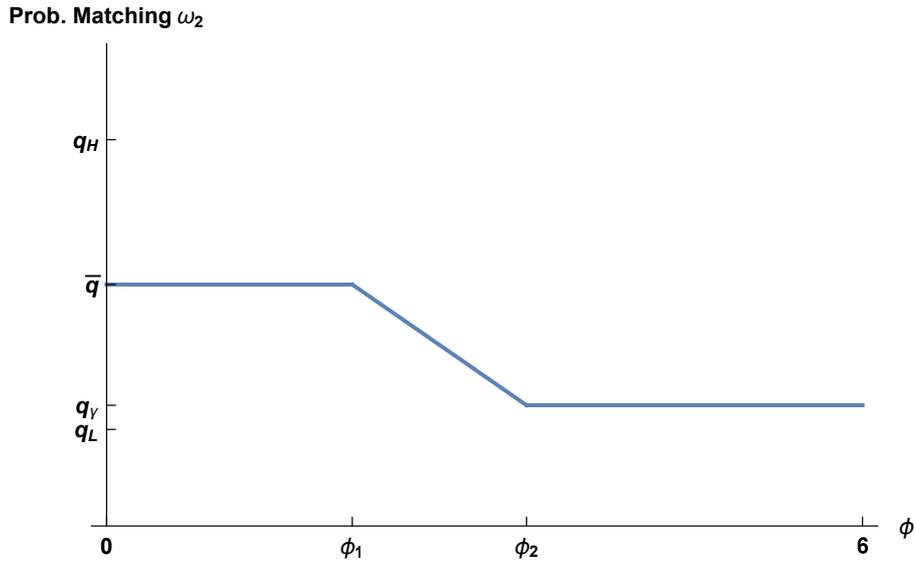
concerns are higher than ϕ_2 , the distortion is maximal and the incumbent plays as if he was fully office-motivated. For intermediate values of ϕ , the value of reputation makes the incumbent indifferent between pursuing the best policy and maximizing the value of his reputation.

When reputation concerns are higher than ϕ_1 , therefore, strategic politicians have the incentive to avoid flip-flops in order to protect their reputation. This leads to an insufficient amount of flip-flopping in equilibrium and an excessive degree of policy persistence. Despite “disliking” flip-floppers for their lower competence, the voter would be better off if more flip-flopping were to take place. In equilibrium, the more flip-flopping avoidance takes place, the smaller is the reputation premium for consistency, which disappears when flip-flopping avoidance is complete.

Welfare and Comparative Statics

The consequences of flip-flopping avoidance for welfare are two-fold: first of all, the second policy action is distorted, as the incumbent sometimes acts disregarding his informative signal. This decreases the probability of matching the state of the world ω_2 , as Figure 1 shows graphically. Moreover, the more the incumbent avoids flip-flopping, the closer r_c moves to λ , that is the incumbent’s track record becomes less and less informative. This decreases the selection component of social welfare: to see why, notice that by Bayes’ Rule, $\mathbb{E}r = \lambda$. Since $\nu(\cdot)$ in expression (1) is strictly convex, $\mathbb{E}\nu(r)$ decreases as r_c decreases from its maximum value of r_c^T (notice that meanwhile, r_f remains fixed at r_f^T). In this model, therefore, reputation concerns such as electoral incentives can only have a negative effect on welfare, by increasing flip-flopping avoidance.

Figure 1: Effect of ϕ on Policy Quality



Probability of $a_2 = \omega_2$ as function of ϕ ; parameters are set at $\lambda = 0.5$, $\gamma = 0.75$, $q_L = 0.6$ and $q_H = 0.9$. Notice that $\bar{q} = \lambda q_H + (1 - \lambda)q_L$ and $q_V = \gamma \bar{q} + (1 - \gamma)(1 - \bar{q})$.

In the analysis undertaken so far I have characterized the equilibrium based on the reputation concerns parameter ϕ , keeping all other parameters fixed. However, the model offers additional interesting comparative statics. In particular, a crucial role is played by the parameter γ , which measures the persistence of the state of the world between $t = 1$ and $t = 2$. The higher γ , the more informative the first signal (and therefore the first decision) is for the decision to be taken in the second period. As a consequence, the higher is γ the more informative flip-flopping is about the incumbent's incompetence: that is, r_f^T is decreasing in γ . Similarly, as γ increases consistency becomes a stronger signal of competence, hence r_c^T is increasing in γ .

Therefore, analogously to what happens with ϕ , as γ increases the incentives to avoid flip-flops become larger. This means that the thresholds ϕ_1 and ϕ_2 decrease in γ . In other words, the equilibrium can be characterized in terms of γ in a similar way as it was characterized in terms of ϕ . In particular, there exists a value of γ that I denote as

γ_1 , such that $\phi_1(\gamma_1) = \phi$, above which the truthful equilibrium stops being sustainable²². Finally, there is another threshold γ_2 above which no interior mixing probability can make the incumbent indifferent when $s_2 \neq s_1$: therefore, the incumbent completely avoids flip-flops²³.

Proposition 1. *The thresholds ϕ_1 and ϕ_2 characterizing the equilibrium are decreasing in γ . Hence, the equilibrium value of σ is also decreasing in persistence γ .*

This has two effects on welfare: the first is that the quality of the second policy decision is non-monotonic in persistence γ . Since γ does not change the probability of matching the state of the world when $a_t = s_t$, the policy component of welfare starts off constant in γ , whereas it then decreases between γ_1 and γ_2 . Finally, when γ is above γ_2 and $\sigma = 0$, the policy component of welfare increases again, thanks to the fact that as $a_2 = s_1$ and γ grows, the probability that $s_1 = \omega_2$ increases. As a matter of fact, when $\gamma = 1$, i.e. the state of the world is fixed, the probability of matching ω_2 under $\sigma = 0$ is the same as under $\sigma = 1$ or as under any other value that σ can take²⁴.

The second effect on welfare has to do with the selection component. For $\gamma \in (\frac{1}{2}, \gamma_1]$, the selection component of welfare increases thanks to the increasing informativeness of consistency as a signal of the incumbent's competence. However, for values above γ_1 the ability of the representative voter to learn is hindered by the flip-flopping avoidance. Finally, for values above γ_2 no learning takes place in equilibrium. Summing up the two components of welfare, as a consequence of Proposition 1 we obtain that $\gamma = \gamma_1$ maximizes social welfare. This value of γ corresponds to the highest value of γ for which

22. $\phi_1(\gamma)$ reflects the fact that ϕ_1 is a function of γ .

23. Analogously to γ_1 , the threshold γ_2 is defined as the value of γ that solves $\phi_2(\gamma_2) = \phi$.

24. The reason for this is that when $\gamma = 1$, choosing the second policy according to the first signal causes no distortion, since when the signal flips, the signal received in the first period is just as good as that received in the second to determine the second period policy.

the truthful equilibrium is sustainable. Therefore, the quality of policy-making is strictly maximal, thanks to the selection component of welfare.

Proposition 2. *The value of γ that maximizes social welfare, both in terms of policy quality (weakly) and in terms of selection, is equal to γ_1 and strictly smaller than 1.*

In the “Discussion and Implications” section I present, among others, some politico-economic implications of the results in Proposition 1 and Proposition 2.

Concerning parameters other than persistence γ , it is worth mentioning some interesting and slightly counterintuitive non-monotonicities. Take for example the fraction of competent politicians λ : the incumbent is more likely to act in an undistorted way when λ takes values close to the extremes. The intuition is that if the representative voter is very confident about the level of competence of the incumbent, the effect of the track-record on reputation is small. As a consequence, when λ increases from small to intermediate values, it is possible for social welfare to decrease. The reason is that the benefit of an increased proportion of competent incumbents is more than offset by the increased distortions in their behavior.

A similar effect can also occur with respect to q_H : as q_H increases, the difference in signal accuracy between high and low types increases and hence flip-flops become more telling of the incumbent being a low type. As a result, given an average quality of politicians $\bar{q} \equiv \lambda q_H + (1 - \lambda)q_L$, flip-flopping avoidance is always (weakly) higher when the ability of politicians is more unequally distributed. Therefore, at least from the perspective of policy quality, it is better to have many relatively good politicians rather than a few stars among a majority of incompetent politicians.

Discussion and Implications

Having presented the main results, this section provides a broader discussion of the implications of the model.

Testable Implications from Variation in Electoral Concerns

- *Behavior of Term-Limited Politicians*: The parameter ϕ measures the strength of electoral concerns. Therefore, it is natural to assume that term-limited politicians have a lower ϕ , since they are not up for re-election. As a result, according to my model we should expect term-limited politicians to reverse their previous decisions more often than those up for re-election²⁵.

By suggesting that term-limited politicians are more willing to challenge the status quo and enact welfare improving policy changes, my model provides an argument in favor of term-limits, to be weighed against the cost of not being able to re-appoint incumbents who are believed to be competent, as well as against the possibility of incumbents acting against the public interest when not subject to electoral accountability.

Finally, another related implication of the effect of a variation in ϕ is that policy changes should be more likely when the underlying issue is not under strong public scrutiny.

- *The Importance of “Genuine Consistency”*: as I discussed in the previous paragraph, an implication of my model is that politicians are more willing to reverse

25. Somewhat consistently with my model, List and Sturm (2006) find that US governors are more likely to reverse environmentally friendly policies when term-limited. Lopes da Fonseca (2020) finds that term-limited mayors act in a more fiscally conservative way.

their previous decisions when reputation concerns ϕ are low. This implies that consistent behavior taking place under strong electoral concerns will be discounted by voters, who anticipate the possibility of politicians posturing consistency. On the contrary, a display of “genuine” consistency at a point in a politician’s career in which electoral concerns were lower should be rewarded more by voters.

This also means that voters and political commentators will have an incentive to look back into the history of decisions by a politician to extract useful information on competence: this can include old decisions taken when not running for an important office, or previously little-known decisions on issues generally outside the public scrutiny. My model can therefore shed light on the observation of politicians being haunted by flip-flops taken relatively far away in the past, which suddenly re-emerge, for example in the midst of an electoral campaign.

Effects and Interpretation of Persistence

- *Timing of Decisions:* So far, the persistence parameter γ has been interpreted as measuring the changeability of the policy-making environment; fixing the time between decisions, the parameter γ measures the speed at which the state of the world changes. Alternatively, keeping the speed of change fixed, the persistence parameter γ represents the time interval between decisions. The longer the interval, the smaller the correlation between ω_1 and ω_2 . Therefore, γ can be thought as being a decreasing function of the time passed after the previous decision: according to this interpretation, γ is close to $\frac{1}{2}$ for two decisions far away in time from each other and close to 1 for decisions very close to each other. This has several interesting

implications: first of all, also in light of Proposition 1, it fits with the evidence (see Doherty, Dowling, and Miller (2015)) that the closer a policy reversal is to the previous decision, the more negatively it is evaluated by voters. Moreover, it suggests that the equilibrium characterization result can be interpreted as defining how responsive the incumbent politician is as a function of the time passed from his previous decision: because of reputation concerns, a minimum amount of time has to pass in order for the politician to again become responsive to information. An interesting extension for future research would be to build on this and make the timing of decisions endogenous²⁶.

- *Task Allocation:* along with persistence of the state of the world over time, γ can be more generally interpreted as measuring the degree of similarity of the two tasks addressed by the incumbent politician in his term in office. The more similar two issues are, the more likely the appropriate policy response to both problems is to be the same. In this respect, following Proposition 1, my model predicts that the more related two issues are, the more negatively a policy reversal is judged. Similarly, the result in Proposition 2 can be interpreted as suggesting that limits to specialization or rotations across sufficiently differentiated tasks can be welfare improving when agents have strong reputation concerns.

Along similar lines, γ can also be interpreted as an index of issue complexity/simplicity: when an issue is very complex, such that there are several factors that might warrant a policy reversal, it is intuitive to expect γ to be close to $\frac{1}{2}$ and hence expect a small effect of flip-flopping on reputation. Conversely, for simpler issues γ is close

26. More details are available in the Online Appendix.

to 1 and flip-flopping is more detrimental. This is consistent with the empirical evidence reported in Doherty, Dowling, and Miller (2015) and Tavits (2007).

- *Public Information:* it is often the case that voters have a relatively good sense of the fact that conditions changed, but much less so of what is the appropriate policy conduct conditional on a given set of circumstances. In light of this fact, the model I describe in this paper is equivalent to the subgame of a larger model, in which after the first policy decision has been taken, players observe the realization of a public signal $\tilde{s} \in \{c, f\}$ informing them on whether the state of the world remained constant (c) or changed (f). The parameter γ can be then interpreted as the accuracy of the signal \tilde{s} , that is $\gamma = Pr(\tilde{s} = c | \omega_2 = \omega_1) = Pr(\tilde{s} = f | \omega_2 \neq \omega_1)$.

Given $\gamma > \frac{1}{2}$, the subgame following the realization of $\tilde{s} = c$ is equivalent to my model, whereas the only difference in the subgame following $\tilde{s} = f$ is that persistence would be equal to $1 - \gamma$, hence lower than $\frac{1}{2}$ unlike what I assume in my baseline model. As a result, in that subgame the voter updates upwards instead of downwards following a flip-flop, which gives the incumbent the incentive to flip-flop, leading to excessive policy volatility. The comparative statics with respect to γ can then be interpreted as stating that the stronger the evidence that the state of the world did not change (respectively, that it changed), the more negatively a flip-flop (respectively, a failure to change the policy) are evaluated. This is consistent with the empirical evidence on the importance of information supporting the decision of the incumbent to reverse policy for the evaluation of a policy change²⁷.

Importantly, following this interpretation of the model, the welfare maximizing

27. See Robison (2017) and Levendusky and Horowitz (2012).

value of γ derived in Proposition 2 corresponds to the socially optimal value of the accuracy of the signal on the change of the state of the world. My model can therefore be interpreted as suggesting that excessive public information on whether the state of the world changed is detrimental to social welfare.

Consistency as Indirect Information on Competence

My model shows that consistency affects the reputation of a politician because it provides indirect information on competence. In particular, in a persistent environment, consistency is a signal of competence, whereas flip-flopping is a signal of incompetence. The persistence of the environment is therefore the key condition under which consistency makes a good heuristic for voters to select competent candidates. At the same time, my model also highlights that – given the fact that policy-relevant information is private – when electoral concerns are high politicians have the incentive to posture consistency, decreasing the value of consistency as a source of information on competence.

The other key ingredient for consistency to play an important role in the evaluation of politicians is the absence of sufficiently accurate direct information on competence that could act as a substitute²⁸. This points to the fact that consistency should play a larger role in environments in which evaluating the quality of a policy decision is difficult: this might be due to timing issues (elections take place before a policy can be proved successful), to the technical character of the policy decision on which the politician is evaluated, or to the lack of government transparency.

28. In the Online Appendix I extend the model to allow for the state of the world ω_1 to be revealed before the second action is taken. I show that when the feedback on the state of the world shows that the incumbent has taken the right decision in the first period, consistency becomes an even stronger signal of competence than in the baseline model. On the other hand, when the first decision of the incumbent turns out to be wrong the voter rewards consistency with respect to the state of the world ω_1 .

In this respect, my model also sheds light on the fact that the consistency heuristic can be a useful alternative to potentially biased or misleading information on the quality of a politician's decisions. Being simple, readily available and easily fact-checkable, it can be particularly useful as communication tool in environments where the credibility of other messages is heavily discounted: this provides another argument to explain why flip-flopping is often used to attack political opponents.

Conclusion

In this paper I present a political agency model of sequential policy-making, which shows how reversing previous policy decisions can be detrimental to a politician's reputation. Even in the absence of any direct information on the quality of the decisions taken by an incumbent politician, flip-flopping is a signal of incompetence. The first contribution of this paper is therefore to provide a rationale for the stigma faced by flip-flopping politicians: this need not be the consequence of an instinctive psychological preference for leaders "sticking to their guns", but it can be a simple and rational heuristic to help voters select competent candidates.

As a result of the reputation damage associated with policy reversals, politicians have an incentive to avoid policy changes and appear consistent. When reputation concerns are strong, this leads to an excessive degree of policy persistence and also hampers the ability of voters to select competent representatives. The unwillingness of candidates to reverse policy decisions when reputation concerns are strong can also explain why politicians are sometimes haunted by flip-flops done a long time in the past: when recent consistency is not informative, as a matter of fact, voters resort to decisions taken at times of lower electoral pressure in order to draw inferences on the competence of candidates.

My model is consistent with several stylized facts concerning the effects of policy changes on reputation: first of all, empirical evidence suggests that voters reward consistency – often independently of partisanship – and that incompetence is one of the reasons behind the negative evaluations of flip-floppers. Moreover, flip-flops on complex issues are better received than flip-flops on simpler ones. Similarly, there is evidence that policy changes close to each other are more damaging, whereas additional information

supporting the decision to reverse the policy has the opposite effect. In this respect, a contribution of my model is to highlight the role of “persistence” as a common denominator of these stylized facts.

My model also has several additional implications, many of which could either be tested empirically or be the object of further theoretical investigations. For example, roll-call voting records – together with approval ratings or primary and general election data – could be used to test whether representatives voting in an inconsistent way are more likely to face a decline in approval ratings, a primary election challenge or a decrease in the shares of votes won in elections²⁹. Another directly testable implication of the model is that term-limited policy-makers should be more willing to revise previous policy choices than policy makers up for re-election. Similarly, politicians should be more willing to change their mind on issues that are away from public scrutiny.

On the theory side there are several interesting open questions. For example, an avenue of future research would be to build on the results of this paper to further explore the substitutability relation between direct and indirect information about the incumbent’s competence and the use of flip-flopping as a accusation in electoral campaigns. Along similar lines, another relevant question to answer is whether there are other environmental and personal characteristics – in addition to those highlighted in this model – that affect the evaluation of a politician’s consistency.

29. I thank an anonymous referee for suggesting this possibility.

Acknowledgements

I would like to thank the Editor Tom S. Clark and three anonymous referees for excellent comments and suggestions. I am grateful to my Ph.D. advisors Andrea Mattozzi and David Levine, and to Larry Samuelson for hosting me as a postdoc at the Cowles Foundation in Yale. The paper also benefitted from discussions with Piero Gottardi, Andrea Galeotti, Ronnie Razin, Alessandro Riboni and several others. I would also like to thank audiences at Yale University, the University of Bologna, the University of Naples, the Max Planck Institute in Munich, the European University Institute, the NICEP Workshop and the SIOE Conference for their questions and comments.

References

- Aghion, Philippe, and Matthew O. Jackson. 2016. "Inducing Leaders to Take Risky Decisions: Dismissal, Tenure, and Term Limits." *American Economic Journal: Microeconomics* 8, no. 3 (August): 1–38.
- Agranov, Marina. 2016. "Flip-Flopping, Primary Visibility, and the Selection of Candidates." *American Economic Journal: Microeconomics* 8, no. 2 (May): 61–85.
- Allgeier, A.R., Donn Byrne, Barbara Brooks, and Diane Revnes. 1979. "The Waffle Phenomenon: Negative Evaluations of Those Who Shift Attitudinally." *Journal of Applied Social Psychology* 9 (2): 170–182.
- Ashworth, Scott, and Kenneth W. Shotts. 2010. "Does Informative Media Commentary Reduce Politicians' Incentives to Pander?" *Journal of Public Economics* 94 (11-12): 838–847.

- Canes-Wrone, Brandice, Michael C. Herron, and Kenneth W. Shotts. 2001. "Leadership and Pandering: A Theory of Executive Policymaking." *American Journal of Political Science* 45 (3): 532–550.
- Carlson, James M, and Kathleen Dolan. 1985. "The Waffle Phenomenon and Candidates' Image." *Psychological Reports* 57 (3): 795–798.
- Croco, Sarah E. 2016. "The Flipside of Flip-flopping: Leader Inconsistency, Citizen Preferences, and the War in Iraq." *Foreign Policy Analysis* 12 (3): 237–257.
- Doherty, David, Conor M. Dowling, and Michael G. Miller. 2015. "When is Changing Policy Positions Costly for Politicians? Experimental Evidence." *Political Behavior* 38 (2): 455–484.
- Fearon, James D. 1994. "Domestic Political Audiences and the Escalation of International Disputes." *American Political Science Review* 88 (3): 577–592.
- Fu, Qiang, and Ming Li. 2014. "Reputation-concerned Policy Makers and Institutional Status Quo Bias." *Journal of Public Economics* 110:15–25.
- Hoffman, Hillary S, and Charles S Carver. 1984. "Political Waffling: Its Effects on the Evaluations of Observers." *Journal of Applied Social Psychology* 14 (4): 375–385.
- Hummel, Patrick. 2010. "Flip-flopping from Primaries to General Elections." *Journal of Public Economics* 94 (11–12): 1020–1027.
- Kaus, Mickey. 2004. "Why Kerry's Flip-flops Matter." *Slate*. doi:www.slate.com/articles/news_and_politics/kausfiles/2004/03/unified_kerry_theory_ii.html.

- Levendusky, Matthew S., and Michael C. Horowitz. 2012. "When Backing Down Is the Right Decision: Partisanship, New Information, and Audience Costs." *The Journal of Politics* 74 (02): 323–338.
- Li, Wei. 2007. "Changing One's Mind When the Facts Change: Incentives of Experts and the Design of Reporting Protocols." *The Review of Economic Studies* 74 (4): 1175–1194.
- List, John A, and Daniel M Sturm. 2006. "How Elections Matter: Theory and Evidence from Environmental Policy." *The Quarterly Journal of Economics* 121 (4): 1249–1281.
- Lopes da Fonseca, Mariana. 2020. "Lame Ducks and Local Fiscal Policy: Quasi-Experimental Evidence from Portugal." *The Economic Journal* 130 (626): 511–533.
- Majumdar, Sumon, and Sharun W. Mukand. 2004. "Policy Gambles." *American Economic Review* 94, no. 4 (September): 1207–1222.
- Maskin, Eric S., and Jean Tirole. 2004. "The Politician and the Judge: Accountability in Government." *American Economic Review* 94 (4): 1034–1054.
- Prat, Andrea. 2005. "The Wrong Kind of Transparency." *The American Economic Review* 95 (3): 862–877.
- Prendergast, Candice, and Lars Stole. 1996. "Impetuous Youngsters and Jaded Old-Timers: Acquiring a Reputation for Learning." *The Journal of Political Economy* 104, no. 6 (December): 1105–1134.

- Robison, Joshua. 2017. "The Role of Elite Accounts in Mitigating the Negative Effects of Repositioning." *Political Behavior* 39 (3): 609–628.
- Schultz, Kenneth A. 2001. *Democracy and Coercive Diplomacy*. Vol. 76. Cambridge University Press.
- Sigelman, Lee, and Carol K. Sigelman. 1986. "Shattered Expectations: Public Responses to "Out-of-Character" Presidential Actions." *Political Behavior* 8 (3): 262–286.
- Smith, Alastair. 1998. "International Crises and Domestic Politics." *American Political Science Review* 92 (3): 623–638.
- Sorek, Ayala Yarkoney, Kathryn Haglin, and Nehemia Geva. 2018. "In Capable Hands: An Experimental Study of the Effects of Competence and Consistency on Leadership Approval." *Political Behavior* 40 (3): 659–679.
- Tavits, Margit. 2007. "Principle vs. Pragmatism: Policy Shifts and Political Competition." *American Journal of Political Science* 51 (1): 151–165.
- Tomz, Michael. 2007. "Domestic Audience Costs in International Relations: An Experimental Approach." *International Organization* 61 (4): 821–840.
- Tomz, Michael, and Robert P. Van Houweling. 2012a. "Candidate Repositioning." *Unpublished Manuscript, Stanford University*: 1–43.
- . 2012b. "Political Repositioning." *Unpublished Manuscript, Stanford University*: 1–34.
- . 2014. "Political Repositioning: A Conjoint Analysis." *Unpublished manuscript, Stanford University*.

Biographical Sketch

Giovanni Andreottola is a Postdoctoral Research Fellow at CSEF and the Federico II University of Naples, Italy. He has previously been a Postdoctoral Associate at the Cowles Foundation at Yale University. He holds a Ph.D. in Economics from the European University Institute in Florence.