Politicization of Intelligence Reporting:
Evidence from the Cold War

Oliver Latham*
University of Cambridge

Abstract

We examine whether there is systematic evidence that the US intelligence services pandered to their political masters when constructing intelligence estimates during the Cold War. We construct a model which shows how career concerns on the part of intelligence analysts could lead them to distort reports towards their President’s prior beliefs. We then take the model’s prediction that errors in intelligence reports should be correlated with Presidential ideology, to the data by constructing a unique measure of intelligence failures that compares CIA/ORE reports on the Soviet strategic, nuclear arsenal to credible, post-Cold War estimates of the Soviet Union’s actual nuclear capabilities. We find that report errors are systematically, positively correlated with both a conventional measure of presidential ideology and a unique, text-based measure of Presidential “hawkishness”. This result is robust to controlling for a number of endogeneity issues and alternative mechanisms such as reverse causality, variation in inter-Superpower relations, collusion between politicians and the intelligence agencies, and turnover in agency staff. Finally, there is evidence that longer-term forecasts are more sensitive to ideology in a manner that is consistent with our model. (JEL: H56, L82, N42)

1 Introduction

To what extent can intelligence agencies be relied upon to provide balanced, accurate reporting to their political masters? In recent years this question has gained greater prominence as the United State’s Central Intelligence Agency (CIA) and the United Kingdom’s Secret Intelligence Service (MI6) have been accused of tailoring information regarding Iraq’s pursuit of Weapons of Mass Destruction to support their respective governments’ preferred policy of regime change in Iraq. (see Barstow et al. (2004) and Norton-Taylor (2011))

If these events are symptomatic of a wider tendency of intelligence agencies to pander to their “consumers” then there are important implications for the reliability of intelligence reporting and hence issues of national security. In this paper we construct a model that builds on the existing economics literature on media bias to argue that career concerns on the part of intelligence analysts could lead them to twist intelligence information to conform to the sitting president’s beliefs. We then test this theory by constructing a unique measure of intelligence failures that compares data from US intelligence reports on the size of the Soviet strategic nuclear arsenal during the Cold War to credible, post-Cold War estimates of the actual number of Soviet strategic weapons. Using this measure as our dependent variable we find robust evidence that more “hawkish” presidents received systematically, upwardly-biased reports. We also find evidence that longer-term predictions are more sensitive to presidential ideology in a way that is also consistent with our model.

*I am grateful to Toke Aidt, Kenneth Benoit, Howard Cobb, Tom Crossley, Bob Evans, Jane Cooley-Fruehwirth, Matthew Gentzkow, James Mahon, John McLaren, Sharun Mukand, Andrea Prat, Christoph Vanberg and participants at the Cambridge Empirical Microeconomics Workshop, CAGE Easter School, North American meeting of the Econometric Society and Silvaplana Political Economy Workshop for helpful comments and suggestions. In addition, I am grateful to Christopher Andrew and participants at the University of Nottingham’s Landscapes of Secrecy conference for helping provide the historical context to the paper.
We measure presidential hawkishness using existing measures of presidential ideology from the political science literature and a new measure based on text analysis. This second measure is constructed using the Wordscores algorithm of Laver et al. (2003) and compares the words spoken by each president in defence-related speeches to those spoken by the most liberal and most conservative deciles of congressmen in contemporaneous congressional debates on defence. We exclude the possibility of reverse causality between the material in the intelligence reports and presidential ideology by basing our text-based measure only on speeches made before each president received security clearance.

Institutional evidence also allows us to discount some alternative explanations. Because we use Top Secret reports to which only the President himself, the Joint Chiefs of Staff and the National Security Council had access, it is unlikely that the estimated, positive effect of presidential ideology is due to analysts colluding with the sitting president to produce information to convince political opponents of the president’s preferred policies.

Another potential mechanism we have to account for is the possibility that there could be turnover in Agency staff that is correlated with new presidents taking office. If, for example, hawkish presidents systematically replace senior analysts with like-minded people, we might estimate a positive effect of ideology on bias, but this would reflect changes in the personnel making the intelligence reports rather than changes in their behaviour due to reputational incentives. We provide evidence against this turnover effect by restricting our sample to years in which historians agree senior appointments to the intelligence services were not politicized and find that the positive effect of ideology persists.

One might also be concerned that the estimated effect of presidential ideology could be picking up the effect of changes in the state of relations between the superpower. The first reason this would be a concern is if there is a selection effect: if more hawkish presidents are more likely to be elected in times of greater paranoia about the threat posed by the Soviet Union and this paranoia also causes the intelligence services to be excessively pessimistic when constructing their reports, then we could pick up a positive correlation between presidential ideology and intelligence forecast errors even when no causal effect exists.

The second concern is that hawkish presidents may be more aggressive in their interactions with the Soviets and it is the resulting worsening in relations that causes an increase in errors in intelligence forecasts of Soviet military strength. We provide evidence that our results are robust to these concerns by showing that the positive effect of ideology is robust to controlling for proxies for inter-Superpower relations. While it is impossible to exclude these effects completely, we argue that the correlations we find are still of interest: whether the positive effect of presidential ideology results from pandering by analysts or an excessive response to changes in the geopolitical situation, the implication is still that the intelligence services made systematic errors when constructing estimates of Soviet nuclear strength.

These results, while interesting in themselves, could well be symptomatic of a wider problem in intelligence reporting. The fact that we find evidence of pandering even in reports which were quantitative in nature (and so could be easily contradicted by future events), were about an issue of ongoing importance (so analysts will have expected to be called up on any discrepancies between reports from one year to the next), and were about issues with extremely high stakes (which one might hope would induce analysts to sacrifice career concerns for the national good), suggests that this phenomenon may play an even greater role in reports where these constraints are absent.

In terms of the existing economics literature, this paper is most closely related to the literature examining bias in media reporting. In the theoretical literature Mullainathan and Shleifer (2005) show how bias can arise when consumers have behavioural preferences. Of more direct relevance to our analysis, Gentzkow and Shapiro (2006) show how reputational concerns on the part of reporters can lead to bias when consumers are fully rational, but have strong prior beliefs as to the true state of the world. Empirical studies of the determinants of bias include Knight (2011), Groseclose and Milyo (2005), DellaVigna and Kaplan (2007) and Latham (2012). Our model also shares similarities with other models of cheap talk with reputational concerns such as Ottaviani and Sorensen (2006) and models of “yes men” such as Prendergast (1993).

This paper also uses techniques from the empirical literature on media bias. In particular, we construct a measure of presidential ideology by using text-analysis methods to compare presidential speeches to speeches by Members of Congress. The concept of using Members of Congress as a reference group is similar to
Gentzkow and Shapiro (2010) who constructed a measure of newspaper slant by examining the frequency with which newspapers used phrases which, when spoken in Congress, are highly correlated with Congressmen’s party affiliation. However, the actual method used to calculate our text based measure is more closely related to the Wordscores algorithm of Laver et al. (2003).

A second, related, literature studies the motivations for and effects of, foreign interventions by intelligence agencies, like the CIA. However, it should be noted that, while this literature is similar in that it studies the actions of intelligence agencies, it differs in that it is concerned with their secondary role as a paramilitary force whereas we focus on their primary role of intelligence gathering.\(^1\)

In the theoretical literature on foreign intervention Aidt and Albornoz (2011) demonstrate how the foreign interests of domestic firms in developed countries might cause them to lobby their governments to suppress democratic movements around the world. In the empirical literature Berger et al. (2012) provide evidence that CIA interventions result in US goods taking up a greater share of trade in the countries in which the interventions took place, while Easterly et al. (2008) provide evidence that CIA interventions have persistent, negative effects on democracy and that these effects are comparable to those caused by interventions by the Soviet MGB and KGB. Finally, Dube et al. (2011) provide further evidence that foreign interventions were motivated by economic factors by showing that top secret presidential authorizations for covert actions in foreign countries affect the stock prices of domestic firms with interests in these countries before the interventions actually take place.

The paper is structured as follows. Section 2 discusses some anecdotal evidence of politicization in intelligence reports, Section 3 describes our model and presents the theoretical results that motivate our empirical analysis. Sections 4 and 5 outline the data and our empirical approach. Sections 6 and 7 present our results and some additional robustness checks. Section 8 concludes.

2 Anecdotal Evidence

In this section we briefly outline a number of cases in which the US intelligence agencies’ reports allegedly became politicized. The main aim of this section is to provide historical evidence that career concerns really did cause analysts to distort intelligence reports and to make clear the distinction between two forms of politicization: the first, which we argue took place during the Johnson administration, results when analysts with career concerns feel unable to speak truth to power. The second, more sinister, form involves analysts deliberately manufacturing material that can be used by their political masters to advance their policy agenda, an alleged case of which occurred in the run up to the 2003 Iraq War.

2.1 President Johnson and the Vietnamese Order-of-Battle

By the mid 1960s American strategy in Vietnam had become a war of attrition. It was believed that, as long as US forces could inflict casualties on Communist forces that were sufficiently in excess of their ability to recruit new members, the war would be won. In order to assess the progress of the war, it was therefore vital to accurately estimate the size of the Communist forces: its “Order-of-Battle”. In the case of the regular North Vietnamese Army this was relatively straightforward. In contrast, estimating the number of guerrilla fighters was much more difficult and responsibility for this fell on the Defense Intelligence Agency (DIA) with oversight from the CIA.

Unfortunately the DIA persistently underestimated the number of guerrilla fighters. “American forces and firepower, it was believed, must be defeating Asian peasant soldiers and guerillas. Order-of-battle (OB) intelligence must therefore reflect that supposed reality,” (Andrew (1995) page 327). This problem was exacerbated by the tendency of the Johnson administration, and the President in particular, to interpret reports which conflicted with their existing beliefs as evidence of disloyalty on the part of the authors (Andrew (1995) page 328).

\(^1\)Primary in the sense of their original mission statements. In fact some historians (e.g. Weiner (2007)) argue that covert operations became the central concern of the CIA with negative implications for its intelligence gathering capabilities.
Even once the CIA began to exercise increased oversight over the reports, conflicting information was suppressed. When a subordinate CIA analyst, Sam Adams, wrote a memo querying the fact that the DIA’s estimates of guerrilla numbers in one province conflicted with captured Vietnamese documents by a factor of ten, this information was suppressed by senior members of the CIA “...After a week Adams went to the seventh floor and found his memo in a folder marked “indefinite hold”...the official in the next-door office told him that battle-order intelligence was...no concern of the CIA”.

However, by 1969, two newly-appointed CIA analysts decided to brief the President of the alternative estimates. They did so despite the fact that they considered it career suicide. One of the analysts, George Carver, recalled that, as he relayed the information, the President “became darker and darker of visage...various courtiers turned white because you didn’t talk to him like that”. Afterwards he concluded “There’s a nice, promising career shot to hell”. In reality the response was very different. Johnson, “thanked (him)...for...services to the Republic” (Andrew (1995) page 346) and, four days later, President Johnson announced on television the cessation of air raids over North Vietnam and his decision not to seek a second term as President.

This episode clearly illustrates the fact that career concerns can cause intelligence officials to suppress information that conflicts with intelligence consumer’s prior beliefs. Importantly this is true even when the analysts have no strong, intrinsic preferences over policy and the consumer is amenable to changing his mind. It is also important to note that there is no evidence that the Johnson administration was explicitly demanding intelligence reports that supported its priors so as to win over its political opponents. Instead there was a genuine desire for accurate information to conduct policy, but this was accompanied by a dangerous tendency to think less of analysts who produced unwelcome news.

2.2 Iraq and Weapons of Mass Destruction

Perhaps the most prominent example of intelligence failure in recent years is the CIA’s support for the Bush administration’s claim that Saddam Hussein’s Iraq possessed weapons of mass destruction (WMDs). In the years since the 2003 invasion it has been widely alleged that the CIA came under, and succumbed to pressure to manufacture evidence that favoured the Bush administration’s desire to go to war.

In particular, it is alleged that the CIA’s then director, George Tenet’s “all-consuming desire to please his superiors” lead to the production of a report “Iraq’s Continuing Programs for Weapons of Mass Destruction” that has since been described as “the worst body of work in its (the CIA’s) long history”. (Weiner (2007), page 563). In particular the agency made a large number of confidently-worded and alarming claims about Iraq’s military capabilities despite a lack of credible sources within the Iraqi regime.

The damning assessment by the journalist Timothy Weiner was that “The analysts accepted whatever supported the case for war. They swallowed secondhand and thirdhand hearsay that conformed to the president’s plans...The CIA as an institution, desperately sought the White House’s attention and approval. It did so by telling the president what he wanted to hear” (Weiner (2007) page 567).

It is important to note that, while the distortion in the CIA’s reports to support government policy was similar to what occurred under President Johnson, the exact mechanism could have been rather different. Rather than simply having strong prior beliefs about whether or not Iraq had WMD, the Bush administration may have simply chosen the existence of these weapons as a suitable pretext to make their desired policy, war with Iraq, politically feasible. The CIA reports were then a useful tool to convince the general public or members of congress. However, evidence against this being the explanation for the CIA’s actions comes from the bipartisan Iraq Intelligence Committee’s conclusion that the reports which were restricted to being “president’s eyes only” were actually “even more misleading...more alarmist and less nuanced” (Silberman et al. (2005), page 14) than those issued to the general public. This suggests that pandering to the prior beliefs of the administration also played a role in driving the CIA’s intelligence failures.

2 The consensus among historians is that, by the mid 1960s, the CIA did not have a strong, intrinsic desire to continue the war in Vietnam unless the evidence supported doing so. In contrast, some have argued that the US military and the DIA were deliberately trying to deceive the President about Vietnamese military strength so as to prolong the war. (Weiner (2007))

3 A more cynical explanation would be that members of the Bush administration valued their future reputations and foresaw that, if no weapons were subsequently found in Iraq, they would have to launch an inquiry. Knowing this, they had an incentive
3 The Model

In this section we present a model of intelligence reporting which formalises a mechanism by which analysts could choose to distort information that conflicts with their president's prior beliefs. In the model, presidents with strong prior beliefs as to the threat posed by the Soviet Union rationally believe that intelligence directors whose reports match their priors are more likely to be of high quality and so are less likely to remove them from office. As a result, career concerns on the part of directors can then cause them to bias their reports towards their president's priors.

The primary goal of the model is to generate testable predictions. As a result, the key contribution of the model is that it can be used to make comparative static predictions on the expected error in intelligence reports in a given time period. Because our data allows us to construct a measure of report errors, these predictions can be tested directly in our empirical analysis. In the main text we focus on the intuition behind the results. A more formal analysis can be found in the appendix.

3.1 The Setup

We consider a doubly infinite time horizon, \( t = (-\infty, \ldots, 0, \ldots, \infty) \). In each period an agency director produces reports for a sitting President about an unobserved state of the world. The sitting President then uses the director's reports to inform him when making a decision about an action with a state-dependent payoff.

A new President is born each period and lives for two periods. The first period of a President's life is his "office period" in which he receives intelligence reports and takes actions. The second is his "legacy period" in which he is out of office, takes no action, but still has preferences over the action chosen by his successor. To simplify the analysis we make the extreme assumption that incoming presidents do not observe the prior history of the game. To do otherwise would quickly make the game intractable and would not shed much new light on the phenomenon we are interested in.

Agency directors, in contrast, are potentially infinitely-lived: their survival from one period to the next is decided by the sitting President. In the baseline model we assume that agency directors care only about remaining in office and capture this by having them receive an endowment, \( A \), at the beginning of each period they are in office.

The director's role is to make reports about a binary state variable, \( S \in \{ L, H \} \). In the context of our empirical analysis this can be interpreted as the military posture of the Soviet Union. We assume that \( H > L \) so that periods where \( S = H \) can be thought of as those in which the Soviet Union poses a greater military threat to the United States. We assume that the state variable is drawn independently across time periods and is unobservable. Agency directors share the common prior that \( P(S = H) = .5 \).

Unlike agency directors, Presidents have heterogeneous priors about the state. An incoming President's prior belief that the state of world is \( H \) is given by the probability, \( \theta \), which is independently drawn from the distribution \( f \) at the beginning of the period. The President's belief is observable to the agency director and we assume that \( f \) is continuous, has support \((0, 1)\) and is symmetric around .5. We refer to all Presidents who have \( \theta \geq .5 \) and so believe the state \( H \) to be more likely as "hawks" and those with \( \theta < .5 \) as "doves".

Although the state itself is unobserved, the agency director has access to additional information in the form of a binary signal, \( s \in \{ L, H \} \) which he then uses to make a binary report \( r \in \{ L, H \} \). The accuracy of the director's signal depends on his type. With prior probability \( \gamma \) the director is "Good" in which case he receives a perfect signal of the state of the world \((s = S \text{ with probability one})\) and always reports his signal truthfully. Otherwise the director is "Normal" in which case his signal is correct with probability \( \pi \in (\frac{1}{2}, 1) \) and he is free to make a report that differs from his signal.\(^7\)

\(^4\)The doubly infinite time horizon is a technical assumption that allows equilibrium play to be stationary for all periods \( t = (0, \ldots, \infty) \).
\(^5\)We relax this assumption in the appendix by allowing for the director to also have preferences over the President's action.
\(^6\)Of course in reality Presidents are elected rather than being chosen at random. While we ignore this aspect in the theoretical analysis, when it comes to the empirics we attempt to control for endogenity issues that might arise if voters systematically elected more hawkish Presidents when they were more concerned about the threat posed by the Soviet Union.
\(^7\)One concern that might arise from this set up is that the "good" type differs from the normal type in two distinct ways: not
After receiving the director’s report, the sitting President takes a binary action, \( a \in \{L, H\} \). The payoff from this action is state dependent: he receives a payoff of one if his action matches the state and zero otherwise. In addition, the previous President receives a payoff of one if his successor takes the correct action in his legacy period and zero otherwise. We assume that the President’s payoffs from the two actions made in his lifetime are received at the end of his legacy period. To do otherwise would be problematic as payoffs from the action perfectly reveal what the state was in any given period.

After the President takes his action, the true state is revealed with some probability, \( \mu \). To match up with our empirical analysis, we allow this probability to vary across time periods. At the beginning of the period \( \mu \) equals \( \mu_h \) with probability \( q \) and \( \mu_l < \mu_h \) with probability \( 1 - q \). The value of \( \mu \) is observed by both the sitting president and the director. We can then interpret periods in which \( \mu = \mu_h \) as those in which the agency director is more likely to be held accountable for the content of his reports by having an external source (such as a Soviet defector or allied intelligence agency) reveal the true state of the world.

After receiving the report and potentially finding out the true state, the incumbent President decides whether to reappoint the agency director to serve his successor in the next period. Intuitively, one would expect that the fact that Presidents care about the action taken in their legacy period would encourage them to reappoint directors who they believe to be good. This follows because a good director will provide more accurate information and so will make it more likely that the correct action will be chosen in future.\(^8\)

We also allow for the President’s reappointment decisions to be affected by other factors that are independent of his expected payoff from the action in his legacy period. For example, the incumbent director may need to be removed due to some exogenous reason such as ill-health or an embarrassing failure in some orthogonal policy dimension.\(^9\) Formally, denoting the President’s expected legacy payoff if he reappoints the director by, \( V_{\text{Reappoint}} \), and if he removes him by \( V_{\text{Remove}} \), then his reappointment decision is governed by the rule:

\[
V_{\text{reappoint}} \geq V_{\text{remove}} + \epsilon
\]  

(1)

Where all of the potential orthogonal factors are captured by shock \( \epsilon \) which is independently drawn each period from a distribution with continuous cdf \( G \). From a modeling perspective this assumption ensures that, in equilibria in which reputation matters, the director’s probability of reappointment is continuous and increasing in the President’s belief that he is a good type.

In summary, the timings of the game in a given period are as follows:

1. Nature independently determines the state of the world, \( S \in \{L, H\} \), the incoming President’s prior belief, \( \theta \in (0, 1) \) and the probability that the state will be revealed, \( \mu \in \{\mu_l, \mu_h\} \)

2. The agency director receives his ego rent, \( A \), and a signal, \( s \in \{L, H\} \)

3. The agency director makes a report \( r \in \{L, H\} \)

4. The incumbent President chooses his action \( a \in \{L, H\} \)

5. With probability \( \mu \) the true state is revealed

6. \( \epsilon \) is drawn and the President decides whether to reappoint the agency director

7. The previous period’s President receives his payoff from the actions made during his office and legacy periods and leaves the game

8. The incumbent President leaves office and enters his legacy period

only does he receive a higher quality signal than the normal type, but he is also more trustworthy. Only the first assumption is critical for our analysis. All of the equilibria we study remain equilibria when we allow the good type to lie. The assumption that the good type receives a perfect signal is also noncritical: qualitatively identical results can be obtained if we assume that the good type receives a superior, but imperfect signal (i.e. that he observes the state with some probability greater than \( \pi \), but less than one.)

\(^8\)In fact the incentive to reappoint good directors is not quite so clear-cut for reasons we will see shortly

\(^9\)Examples of such exogenous reasons in the history of the CIA would be DCI William Colby’s death in a boating accident and DCI Allen Dulles’ removal for perceived failures during the planning of the infamous Bay of Pigs incident.
3.2 Equilibrium

In this section we informally describe our equilibrium concept and state the main results. A more formal statement of players’ strategies and beliefs, the full derivations of the main results and a number of extensions to the baseline model can be found in the appendix.

Throughout the analysis we focus on symmetric, stationary Markov Perfect Equilibrium (SSMPE). These are equilibria which are stationary: no player’s actions or beliefs depends on the calendar time of the game, symmetric: all types of president have the same prior belief as to the incumbent director’s quality, and Markov: the agency director does not use a history dependent strategy. Under these assumptions we can write the normal director’s reporting strategy as, $\sigma^\theta_s(r, \mu)$, the probability that he reports $r$ after signal $s$ given that the President is of type $\theta$ and the probability of external verification is given by $\mu$.

The President’s posterior belief that the director is a good type (the director’s “reputation”) will depend on both the director’s report and whether or not the true state is revealed. The President will either find out that the true state was $L$, that it was $H$ or he will receive no additional information about the state. Denoting the last case by $\emptyset$, the three possible forms of “feedback” the President could receive about the state are given by $F \in \{\emptyset, L, H\}$. We then denote a President of type $\theta$’s posterior belief about the director’s quality after report $r$ and feedback $F$ by, $\lambda^\theta(r, F, \mu)$.

Rather than assume an exogenous benefit from reputation, our model derives the value of reputation endogenously through the President’s reappointment decision. Intuitively, the incumbent President wants to reappoint high quality directors because doing so will mean his successor will receive more accurate reports which, provided they are not ignored, will make it more likely that he takes the correct action. This translates into a higher expected legacy payoff for the incumbent President. As a result, the President’s probability of reappointing the director will be increasing in the director’s reputation. Finally, a desire to increase their probability of being reappointed will lead normal directors to choose their reporting strategy so as to maximise their reputation.

In an informative equilibrium (one in which intelligence reports change Presidents’ beliefs about the state of the world) reporting is determined by the interaction between two effects. The first is a pandering effect: if the state is not revealed, the director will typically be better off if he reported the state the President already believed to be more likely.

The intuition for this is most easily understood if we focus on the case where the incumbent President is a hawk and believes that normal types will report their signals truthfully. In this case he expects good types to always report the state correctly and so report $H$ with probability $\theta > .5$. Normal types, in contrast, will report $H$ either when the true state is $H$ and they get the correct signal or the true state is $L$ and they get the wrong signal. This occurs with probability $\theta r + (1 - \theta)(1 - \pi) < \theta$. This implies that the President believes good types are more likely to report $H$ than $L$ so that $H$ reports are more likely to come from good types. This implies that the director’s reputation will be higher after $H$ signals than $L$ signals and gives the normal director an incentive to misreport $L$ signals.

However, the more often normal types misreport $L$ signals, the smaller his reputational advantage becomes because misreporting by normal types reduces the difference in reporting probabilities between normal and good types. As a result, the size of the reputational incentive to pander diminishes with the degree of pandering. In the case of a dovish President the opposite occurs: when the state is not revealed and at low levels of misreporting the President will interpret $L$ reports as being more likely to have come from high quality directors and so the director will have an incentive to pander by misreporting $H$ signals.

The second effect comes from the possibility that the state is revealed. Whenever the state is revealed
and fails to match the director’s report he reveals himself to be a normal type for sure (because good types always report the state correctly). As a result, a failure to report the state correctly when it is revealed will cause the director’s reputation to drop to zero. This risk of being found out encourages the director to report his signal truthfully, whether or not it conforms with the President’s priors.\textsuperscript{11}

When the director’s signal matches the President’s prior beliefs these two effects work in the same direction: both encourage him to report truthfully. As a result, he will always report such signals correctly. In contrast, when the director’s signal contradicts the President’s prior, the two effects work in opposite directions: the first encourages him to lie, while the second encourages him to tell the truth.

If the second effect always dominates, then the director will report contradictory signals truthfully. Otherwise, the equilibrium probability of misreporting is determined so as to exactly offset these two effects and leave him indifferent between reporting $L$ or $H$ after signals that contradict the President’s priors. Because the reputational benefit from pandering falls as the probability of lying increases until it eventually becomes negative, this probability will be unique and strictly less than one.

Proposition 1 states this equilibrium formally:

**Proposition 1.** The game has a SSMPE with the following properties:

1. If the President is a hawk, then normal agency directors always report $H$ signals truthfully, but distort $L$ signals with the unique probability $\sigma_h^*(\theta, \mu) \in [0, 1)$. In the range where $\sigma_h^*(\theta, \mu)$ is strictly positive it is increasing in $\theta$.

2. If the President is a dove, then normal agency directors always report $L$ signals truthfully, but distort $H$ signals with the unique probability $\sigma_d^*(\theta, \mu) \in [0, 1)$. In the range where $\sigma_d^*(\theta, \mu)$ is strictly positive it is decreasing in $\theta$.

3. For each type of President there exists $\mu_0^*$ such that if $\mu_h, \mu_l \geq \mu_0^*$ the director always reports his signal truthfully to a President of type $\theta$, if $\mu_h > \mu_0^* > \mu_l$ there will be no misreporting to presidents of type $\theta$ when $\mu = \mu_h$ and positive probability of misreporting when $\mu = \mu_l$. Finally, if $\mu_h, \mu_l < \mu_0^*$ there will always be positive probability of misreporting to a President of type $\theta$.

All informative SSMPEs take this form and have the same comparative statics.

The key results in terms of our empirical analysis are the comparative statics on $\theta$. In the case of a hawkish president, increasing $\theta$ means that, all else being equal, the director’s reputation will be higher after $H$ reports relative to $L$ reports. This in turn increases the incentive to pander to the President’s priors by reporting $H$ after an $L$ signal. In order to preserve equilibrium the probability of misreporting $L$ signals must increase so as to reduce the incentive to pander and, once again, leave the normal director indifferent between reporting $L$ and reporting $H$ after an $L$ signal.

When the President is a dove the opposite occurs: increasing $\theta$ reduces the director’s probability of reappointment after reporting $L$, both when the state is revealed and when it is not. Hence it unambiguously decreases the director’s incentive to misreport $H$ signals and so, in order to preserve equilibrium, the probability of misreporting must fall.

After examining the effect of changes in $\theta$ the next question is how the incentive to pander to the President’s prior beliefs and hence the level of misreporting in equilibrium, changes with the probability of external verification of the state, $\mu$. This is summarised in Proposition 2:

**Proposition 2.** Normal directors pander more when $\mu = \mu_l$ and the probability the true state is revealed is low. Formally, $\sigma_h^*(\theta, \mu_h) \leq \sigma_h^*(\theta, \mu_l)$ and $\sigma_d^*(\theta, \mu_h) \leq \sigma_d^*(\theta, \mu_l)$. These inequalities hold strictly when respectively $\sigma_h^*(\theta, \mu_l) > 0$ and $\sigma_d^*(\theta, \mu_l) > 0$.

\textsuperscript{11}In addition to the risk of being caught out, there is also a benefit to being contrary when the state is revealed: when normal types are only misreporting signals that contradict the President’s priors, successfully matching the state has a greater reputational benefit when the state contradicts the President’s priors (i.e. for a hawkish President $X^g(L, L, \mu) > X^g(H, H, \mu)$.)
Whatever the ideology of the President, increasing the probability of external verification reduces the degree of pandering. The intuition is identical to that in Gentzkow and Shapiro (2006): when $\mu$ is large the director puts greater weight on what happens to his reputation when the state is revealed. It then becomes important to avoid mismatching the state (as this leads to him revealing his type for sure) and so encourages the director to report his signal truthfully. This implies that hawkish presidents will receive $L$ reports more often when $\mu = \mu_h$ while dovish presidents will receive $L$ reports less often.

### 3.3 Empirical Implications

The results in Propositions 1 and 2 tell us how the probability with which a normal director misreports signals that contradict his President’s prior beliefs changes with the president’s ideology and the probability the state is revealed. While this helps formalise the channel by which reputation concerns can affect intelligence reporting, none of these results can be tested with the data we have available because, rather than observe the director’s reporting strategy, we only observe the reports he made along with the true underlying state of the world.

In this section we close the gap between our model and the data by using the equilibrium values of $\sigma^*_h(\theta, \mu)$ and $\sigma^*_d(\theta, \mu)$ to calculate the expected report error in a given period as a function of the parameters of the model. Because this is observable given our data, any comparative static results on it can be tested directly.

As a first step we write the unconditional expectation of the director’s report as follows:

$$E(\text{report}) = \lambda \frac{H + L}{2} + (1 - \lambda) \Sigma P(\mu) \left[ \int_{\frac{1}{2}}^{1} \frac{L}{2} (1 + \sigma^*_d(\theta \mu)) + \frac{H}{2} (1 - \sigma^*_d(\theta \mu)) f(\theta) d\theta + \int_{\frac{1}{2}}^{1} \frac{L}{2} (1 - \sigma^*_h(\theta \mu)) + \frac{H}{2} (1 + \sigma^*_h(\theta \mu)) f(\theta) d\theta \right]$$

With probability $\lambda$ the director is a good type. With equal probability he receives an $H$ signal or an $L$ signal to give his expected report as $\frac{H + L}{2}$. If instead the director is a normal type then his expected report depends on both the President’s type and the probability of external verification. The first integral inside the square brackets gives the expected report of a normal director when faced with a dovish president. He receives signal $H$ with probability one half, but, because of his reputational concerns he only reports $H$ with probability $\frac{1}{2}(1 - \sigma^*_h(\theta, \mu))$. Similarly the second term accounts for the case of a hawkish president: a normal director only reports $L$ to a hawkish president of type $\theta$ with probability $\frac{1}{2}(1 - \sigma^*_h(\theta, \mu))$. Rearranging gives the expected report as:

$$E(\text{report}) = \frac{H + L}{2} + (1 - \lambda) \Sigma P(\mu) \left[ \int_{\frac{1}{2}}^{1} \sigma^*_h(\theta, \mu) f(\theta) d\theta - \int_{\frac{1}{2}}^{1} \sigma^*_d(\theta, \mu) f(\theta) d\theta \right] (\frac{H - L}{2})$$

Using the fact that $\frac{H + L}{2}$ is simply the expected state of the world allows us to express the report error as a function of the deep parameters of the model, $\theta$, $\mu$ and $\pi$:

$$E(\text{report} - \text{state}) = \Sigma P(\mu) \kappa \left[ \int_{\frac{1}{2}}^{1} \sigma^*_h(\theta, \mu, \pi) f(\theta) d\theta - \int_{\frac{1}{2}}^{1} \sigma^*_d(\theta, \mu, \pi) f(\theta) d\theta \right]$$

Where $\kappa = (1 - \lambda) \left( \frac{H - L}{2} \right)$, a positive constant.

The first prediction of our model is that the expected report error should be strictly increasing in the President’s hawkishness. To see this one can examine the change in the conditional expectation of the report error as a function of the deep parameters of the model, $\theta$, $\mu$, and $\pi$. For these calculations we use the director’s prior belief over the state.
error when we change the President’s ideology from $\theta$ to $\theta' > \theta$ and see that it is positive whether the President is a dove or a hawk:

$$E(\text{report-state} \mid \theta', \mu) - E(\text{report-state} \mid \theta, \mu) = \begin{cases} \kappa (\sigma^*_h(\theta', \mu) - \sigma^*_h(\theta, \mu)) \geq 0 & \text{if } \theta > .5 \\ -\kappa (\sigma^*_d(\theta', \mu) - \sigma^*_d(\theta, \mu)) \geq 0 & \text{if } \theta < .5 \end{cases}$$

Intuitively, if the incumbent president is a dove, then increasing $\theta$ will reduce the incentive for the director to report $L$ and so will increase the expected report error (i.e. make it less negative). Alternatively, when the president is a hawk, increasing $\theta$ will increase the incentive for the director to report $H$ and so will again increase the expected error. As a result there is a monotonic, positive effect of presidential ideology on the level of the report error:

**Prediction 1.** The error in intelligence reports about Soviet military strength should be increasing in Presidential hawkishness, $\theta$

In contrast, the model does not predict a monotonic effect of moving from $\mu_h$ to $\mu_l$ on the expected report error. Intuitively, moving from $\mu_h$ to $\mu_l$ will increase the level of pandering to all types of president. This will lead to an increase in the expected report error when the President is a hawk and a decrease when he is a dove. In the appendix we show that, when the distribution of Presidential types is symmetric, these two effects are of equivalent magnitude and so the net effect on the expected report error is equal to zero. The model therefore predicts that changes in $\mu$ should have no direct effect on the size of the report error:

**Prediction 2.** There should be no direct effect of changes in $\mu$ on the error in intelligence reports about Soviet military strength

While there should be no direct effect of changes in $\mu$, there could be an interactive effect between ideology and the probability of external verification. This will be the case if reports are more or less sensitive to changes in ideology when the probability of state verification is $\mu = \mu_h$ rather than $\mu_l$. Unfortunately it is not possible to determine the sign of this effect without further specifying the shape of $G$ (the function that maps a director’s reputation to the president’s reappointment decision). However, we show in the appendix that, as long as the marginal effect of making the president’s beliefs more “extreme” (i.e. of increasing $\theta$ for a hawk or reducing it for a dove) on the incentive to pander is larger when $\mu = \mu_l$, we would expect a negative interactive effect between $\theta$ and $\mu$:

**Prediction 3.** Suppose that $\frac{\partial \sigma^*_h(\theta, \mu)}{\partial \theta} > \frac{\partial \sigma^*_h(\theta, \mu)}{\partial \theta}$ and $-\frac{\partial \sigma^*_d(\theta, \mu)}{\partial \theta} > -\frac{\partial \sigma^*_d(\theta, \mu)}{\partial \theta}$ so that the marginal effect of making a President “more extreme” on the probability of misreporting is greater when $\mu = \mu_l$ than when $\mu = \mu_h$. Then there should be a negative, interactive effect between $\mu$ and $\theta$ on the report error (i.e. the marginal effect of increasing $\theta$ on the report error should be smaller when $\mu = \mu_h$).

Intuitively, one might expect that reducing the probability of external verification would make intelligence reports more sensitive to Presidential ideology. Prediction 3 states that, if this is the case, we should expect a reduction in the probability of external verification to amplify the effect of changes in Presidential hawkishness on the expected report error. It is this amplification effect that we test for when we include an interaction term between our proxies for $\theta$ and $\mu$ in our empirical specifications.

With the theoretical mechanism and its predictions established we now turn to finding empirical counterparts for the various components of equation (2) before taking these predictions to the data.
4 The Data

4.1 Measuring Intelligence Errors

An Overview of the Soviet Nuclear Arsenal 1949-1990

Before describing the data in detail we will first provide a brief categorization of the nuclear weapons developed by the Soviet Union during the Cold War. The first Soviet atomic bomb was detonated in 1949, to the surprise of Western intelligence agencies who had not anticipated the Soviets to develop a bomb until the 1950’s (Goodman (2007)). In the early stages of the Cold War there was little distinction between different atomic weapons and the main objective of both the US and USSR was simply to amass as large a stockpile as possible. However, as technology developed, the variety of weapons and delivery systems increased and their roles became more specialised.

The first key distinction the kind of targets a weapon is designed to attack. There are three broad categories. The first is strategic arms, these are long-range weapons with the capability of striking the continental United States. Second are weapons for peripheral attack. These are intermediate-range weapons which, while potentially capable of striking the continental United States, are primarily designed to attack European targets or the non-continental United States. Finally there are tactical nuclear weapons. These are “small”, short-range weapons designed for use against nearby targets in support of conventional military operations.

Our empirical analysis focuses entirely on strategic arms. This is partly from necessity: the CIA did not make systematic forecasts of the size of the USSR’s stockpile of tactical or peripheral weapons for a long enough period of time and credible post Cold War data of the true size of these stockpiles are not readily available. However, even if such data were available one could make the case for focusing on strategic arms. As the only weapons capable of inflicting significant casualties on American citizens, the size of the USSR’s strategic arsenal was clearly of paramount importance to the President.

The second element of categorization is the delivery vehicle upon which a given nuclear weapon was mounted. Broadly speaking there were three means of delivering strategic nuclear weapons to targets in the United States. Firstly, warheads could be loaded on to long-range Heavy Bombers such as the TU-4 “Bear” or M-4 “Bison”. Secondly they could be mounted on Intercontinental Ballistic Missiles (ICBMs) such as the SS-19 “Stiletto”. Finally they could be mounted on Submarine Launched Ballistic Missiles (SLBMs) like the SS-N-6 “Serb”, which are themselves carried by ballistic missile submarines (SSBs) or nuclear-powered ballistic missile submarines (SSBNs) like the “Golf” or “Typhoon”.

These delivery mechanisms were developed at different points during the Cold War. The first Soviet bombers capable of attacking the continental United States while still having enough fuel to return to a friendly airbase, were developed in 1952. The Soviet Union’s first ballistic missile submarine was launched in 1958 and its first ICBM was deployed in 1960.

Until the mid 1970s there was no distinction between the number of deployed Soviet missiles and the number of deployed Soviet nuclear warheads on those missiles. However, after the invention of Multiple Independently-Targetable Reentry Vehicles (MIRVs) one missile could carry multiple warheads, each capable of attacking a different city or military target. Because MIRVed missiles look outwardly identical to missiles with a single warhead, estimates of the number of MIRVs on Soviet ICBMs are an example of reports with a very low probability of external verification. If the amplification effect in Prediction 3 is correct, one would expect the error in CIA reports on the number of MIRVs to be more sensitive to presidential ideology than reports on more externally-verifiable statistics like the number of Soviet ICBM launchers. Unfortunately we cannot test this prediction as the detailed break down of the Soviet arsenal has not yet been declassified for a sufficiently long time-series.

The Intelligence Reports

US fears of a potential conflict with the Soviet Union existed even before the end of the Second World War and intelligence assessments of the USSR’s ability to attack the US predate the creation of the CIA. After the formation of the CIA in 1947 its head, the Director of Central Intelligence (DCI) was charged with
coordinating the 16 various US intelligence agencies to produce annual reports on Soviet capabilities and intentions. The actual production of the reports was, from 1950 to 1973, conducted by the Office of Reports and Estimates (ORE). The responsibility then transferred to the National Intelligence Officers and finally, in 1979, to the National Intelligence Council. Throughout the Cold War the reports were passed through the CIA and presented to the President by the DCI.\(^{13}\)

In each year the reports contain estimates of the number of Soviet SLBMs, ICBMs and Heavy Bombers (“Bears”, “Bison” and “Blackjacks”).\(^{14}\) In some cases the reports have been redacted and so we have missing observations. Where possible we fill in the gaps with data from other Top Secret reports released by the CIA such as estimates of Soviet naval strength which typically contain estimates of the number of SLBMs. However, there are still some missing observations.

It should be noted that these missing observations appear to result from random variation in the way the reports were constructed. By this we mean that the data in our sample is not itself classified, but it is sometimes redacted because it appears next to other, more sensitive, data. For example, in some reports the number of ICBM launchers is referred to as being in a table that also contains sensitive, technical information on Soviet weaponry and the entire table has been removed.

Crucially each report gives us more than the CIA’s estimate of the current status of the Soviet arsenal. The reports also give forecasts of future Soviet capabilities. From each report we are therefore able to record the CIA’s estimate of Soviet nuclear strength in the corresponding year as well as for a varying number of future years.

When making forward predictions the reports do not always give a single point estimate. Instead they sometimes present a number of different estimates of future Soviet forces based on different geopolitical scenarios. In particular, during the 70’s and 80’s the forecasts differ depending on the outcome of negotiations between the US and USSR regarding arms limitation treaties such as the Strategic Arms Limitation Treaty (SALT I) and its successor SALT II. To ensure the maximum level of robustness for our results we use the CIA predictions that correspond to the actual outcome of the talks: so we use the SALT I scenarios for the years 1974 to 1979. The SALT II agreement represents a greater difficulty as, although the treaty was signed by Jimmy Carter in June 1979, it was never ratified by the senate. However, historians agree that both superpowers adhered to its terms until 1986 when Ronald Reagan withdrew the US from the treaty (Andrew (1995); Weiner (2007)). For this reason we use the SALT II scenarios for the years 1979-86 and the non-SALT scenarios afterwards.\(^{15}\)

As well as conditioning its predictions on the outcome of negotiations between the superpowers, the reports also occasionally present multiple levels of Soviet Forces, typically “Low” “Medium” or “High”. In most cases the document’s authors make clear which scenario they consider most likely and this is the set of numbers we use. In almost all cases this leads to the medium scenario being used.

One scenario the CIA did not account for was the rapid collapse of the Soviet Union after the fall of the Berlin Wall in 1990. The subsequent end of the Cold War and economic restructuring of the former Soviet Union lead to a dramatic reduction in the number of “Soviet” strategic weapons. For this reason we exclude all predictions made for years after 1990. To do otherwise would drastically exaggerate the degree of CIA bias during the Reagan administration. As President Reagan is widely perceived to have been more hawkish than average, doing otherwise would only make it more likely that the results would support our hypothesis.

The final result is the variable reported\(_{itt'}\). The observations are indexed by \(itt'\) where \(i \in \{\text{Bomber, ICBM, SLBM}\}\) records the weapon system, \(t\) the year the report was published and \(t'\) the year to which the estimate in question corresponds. In other words, reported\(_{itt'}\) records “the number of weapon system \(i\) the report in year \(t\) predicted that the Soviets would have in year \(t'\)”. The overlapping nature of the reports and the fact that we have data on multiple weapon systems means that we have around 300 observations despite the fact that the Cold War lasted for less than half a century.

\(^{13}\)For this reason we sometimes refer to the “CIA reports” or “CIA data”. Strictly speaking this is incorrect, but using this short-hand terminology is common even among historians.

\(^{14}\)We exclude any Bears which the CIA believed the Soviet Union was using for anti-submarine warfare rather than strategic bombing. Note that we refer to weapon systems by their NATO reporting names rather than the Soviet equivalents.

\(^{15}\)Note that, if we never used the SALT II estimates, the result would be to exaggerate the report error under President Reagan (a hawk) so would actually reinforce our results.
The Post Cold War Data

Clearly it is not possible to look for systematic variation in the accuracy of these intelligence reports without a reference point to compare to. Increases in the estimated size of the Soviet arsenal during a hawkish President’s administration could reflect biased reporting by the CIA or a genuine increase in the actual number of Soviet weapons. For this reason we need credible estimates of the actual size of the Soviet arsenal during the period in question.\footnote{In the language of the model we need to know the true state of the world.}

We obtained this data from the Bulletin of the Atomic Scientists (BoAS) who constructed the data in collaboration with the National Resources Defense Council. This data gives us time series of the number of strategic bombers, SLBMs and ICBM for the years 1956 to 1989.\footnote{In fact the data actually goes even further and breaks down these categories to the individual weapon classes (i.e. the number of bombers is disaggregated to give the number of Bear bombers, Bison bombers etc.) Unfortunately it is not possible to construct equally disaggregated data from the intelligence reports.}

Crucially these sources are independent of government and the data was compiled in 2003, after the end of the Cold War. This means that any bias in the numbers should not be systematically correlated with biases in the CIA data. The BoAS data is also based on more accurate information: the level of secrecy around the Soviet nuclear program has greatly diminished since the end of the Cold War. Much of the relevant data has been declassified and, even when it has not, many pieces of information that were once conjecture (such as the number of SLBMs carried on each Typhoon SSBN or whether the Backfire bomber was to be used for strategic or peripheral purposes) are now common knowledge. In addition, a large number of pieces of Soviet equipment have been withdrawn from service. As the total production runs of each system are now known to a high degree of accuracy, historians have been able to estimate the size of the Soviet arsenal through the Cold War.

Throughout the analysis we take the BOAS estimates to be accurate reflections of actual Soviet strength and refer to any deviation of the CIA reports from these estimates as an “error” in the reports. Figure 1 presents the intelligence estimate of the current number of Soviet ICBM launchers in each year of the Cold War alongside the BoAS estimates.

As can be seen there is substantial variation in the direction and size of the mistakes. Under Eisenhower the estimates suggested a rapidly growing force of ICBMs when in fact no such weapons existed. In contrast, under President Johnson, the intelligence reports actually significantly underestimated Soviet strength. This gap was closed during the first years of the Nixon presidency after which the reports vastly overestimated the number of launchers. This tendency to overestimate broadly continued until the end of the Cold War.

Having established that there is some variation in the accuracy of the intelligence reports we construct the raw forecast error as our dependent variable:

\[
\text{reported}_{it} - \text{actual}_{it}
\]

4.2 Measuring Presidential Ideology

The second element of the model for which we need to find an empirical counterpart is \( \theta \), the President’s “hawkishness”. This poses a number of problems. The first is one of measurement error: the bulk of measures of presidential ideology from the political science literature measure a president’s position on a liberal/conservative dimension and, while one would expect this to be correlated with a president’s hawkishness it will not be perfectly so.

A second issue is one of reverse causality: all measures which are based on the ex-post actions of a president after he took office are somewhat suspect, as there will be a degree of feedback from the information in the reports. For example, presidents who received exaggerated estimates of Soviet military strength will have been more likely to have approved increased expenditure on defence. We will discuss these issues in more detail in the next section, but will first outline the two ideology measures we use and how they are constructed.
Nominate Score

The first measure of president’s ideology we use is their DW nominate score. This score places each President on a left-right dimension. It is constructed by first estimating ideological scores for every roll call vote held in the House of Representatives using a probabilistic voting model of roll call voting by congressmen. Presidential scores are constructed by examining roll call votes in which the President publicly declared himself to be in favour or against the bill in question. The DW nominate algorithm uses congressmen whose term of office spanned multiple congresses to construct a measure that is comparable over time (Poole (2005)).

A Text Based Measure

The second approach we use is to construct a text-based measure of presidential hawkishness using the Wordscore algorithm of Laver et al. (2003). This constructs a value-neutral ideology measure for a “virgin text” by mapping it to a set of “reference texts” whose ideology is either known or can be assumed by the researcher. This is achieved by comparing the relative frequency of words in the different reference texts to construct a score for each word. The virgin text is then given a score by constructing a weighted sum of the scores of all words in the text.

In our case we map speeches on defence made by each president to congressional speeches made during the same time period in defence-related debates. Crucially we use only pre-presidential speeches made before each president received security clearance so that there can be no possibility of reverse causality between the material in the reports and the measure we construct.

Here we describe how we calculate ideology scores conditional on having the virgin and reference texts. A full description of how these texts were constructed is in the appendix. For each president we separately run the Wordscore algorithm using the pooled text of a set of defense-related speeches as the virgin text. To avoid problems of reverse causality we ensure that all speeches were made prior to the president receiving security clearance.

For each president we then construct two reference texts by pooling together the words spoken in defence-related debates by respectively the most liberal ten percent of congressmen and the most conservative ten-percent of congressmen as measured by their DW nominate score. When constructing these reference texts we use the congressional debates from the year that corresponds to the time the presidential speeches were made. A full description of the speeches used for each president and the corresponding congresses is in Table 1.

Once we have constructed the reference texts the Wordscore algorithm works by first identifying every unique word that appears in the reference texts. Indexing these words by \( i = 1, \ldots, N \) we then calculate the frequencies with which each word appears in the liberal reference text and the conservative reference text as \( \{f_L^i\}_{i=1}^N \) and \( \{f_C^i\}_{i=1}^N \). Each word is then given a score equal to the difference in these two frequencies:

\[
    w_i = f_C^i - f_L^i \tag{3}
\]

We then calculate the corresponding frequencies from the presidential speeches as \( \{f_P^i\}_{i=1}^N \) a presidential hawkishness score is then calculated as a weighted sum of the wordscores calculated in equation (3) where the weights correspond to the frequency with which the president spoke a given word:

\[
    TS_t = \frac{1}{N} \sum_{i=1}^{N} f_P^i w_i \tag{4}
\]

So that presidents who, when talking about defence, persistently uses words that are mostly spoken by conservative congressmen during defence-related debates will be considered more hawkish. Finally, for ease of exposition we rescale the above measure so that it has the same standard deviation as the presidential DW nominate scores. Denoting the sample average text score by \( \bar{TS} \) and the sample standard deviations of
the presidential DW nominate score and text score by $SD_d$ and $SD_t$ we calculate the transformed scores as:

$$text\ score_t = TS + (TS_t - \overline{T}) \frac{SD_d}{SD_t}$$

This ensures that both ideology measures are on the same scale so that their estimated coefficients are directly comparable.

**Discussion**

All of the discussion about measurement error in the presidential DW nominate scores would be academic if our text based measure delivered an identical ranking of presidents or one that strongly contradicted the stylised historical facts. To ensure this is not the case Table 2 gives the implied ideological ranking from liberal/dovish to conservative/hawkish when we use the nominate or text score measure.

Turning first to the ranking by nominate score we see that it gives an intuitive ranking of the Presidents from liberal to conservative: Carter and Johnson are the most liberal, while Reagan and Nixon are the most conservative. Unfortunately, this ranking does not correspond as closely to an intuitive ranking of Presidential hawkishness. In particular Kennedy, while being relatively liberal when it came to domestic matters was relatively hawkish in his approach to the arms race and towards preventing the spread of Communism in Latin America. In contrast, Richard Nixon, while socially conservative, was the first President to open diplomatic relations with Communist China, scaled back the war in Vietnam and implemented the first arms limitation agreement between the superpowers.

Turning to the ranking induced by the Wordscore algorithm we see that these anomalies are rectified: Kennedy moves significantly higher in the rankings while Nixon moves further down. Importantly for the credibility of our measure President Reagan retains his position as the most hawkish President while Presidents Johnson and Carter remain at the low end of the rankings.

While the resulting ranking is consistent with stylized historical facts some of the assumptions made in its construction are worthy of discussion. Firstly, one might ask why we use only the extremes of the distribution rather than mapping the presidential speeches to the entire set of congressmen. This has been the approach of Gentzkow and Shapiro (2010) when constructing a measure of the ideological slant of newspapers.

The reason is that previous studies have been interested in making intratemporal rather than intertemporal comparisons. Because we are trying to compare presidents across time we need to be careful that the calculated differences between presidents are not driven by intertemporal changes in the congress we are mapping to.

To take one example of how this could occur, suppose that Republicans are, on average, more hawkish than Democrats and that we are interested in comparing two presidents who, in reality, are as hawkish as the average republican congressman. Now suppose that there are two cases: one in which the president faces a majority Democrat congress and one in which he faces a majority Republican congress. In the first case the President’s speeches will come across as relatively hawkish: he will sound like the average Republican who, because there are few Republicans in Congress will be in the right hand tail of the distribution of congressmen. In the second case he will seem relatively moderate: he will sound like an average Republican who, because Republicans dominate the house, will be a relatively centrist congressman. We argue that using only the extremes of the distribution mitigates this problem as these are likely to be less sensitive to electoral changes.

A second issue is how our measure handles changes in the language used over time? One might be concerned that the accuracy of the measure could be reduced if there are large changes in language over the time period of our study. We argue that this shouldn’t be a problem as long as changes in language affect the entire distribution. If this is the case then our measure will still be informative: it tells us that the president is relatively dovish/hawkish by the standards of his time.

Having calculated our two measures of presidential ideology we now describe our set of control variables before discussing endogeneity issues and presenting our results.
4.3 Other Controls

External Verification and Quality of Monitoring Technology

According to our theoretical model we need to include controls for $\pi$ and $\mu$: the quality of the CIA’s information and the likelihood of outside verification. To do this we firstly include weapon fixed effects: we include the controls bomber and icbm leaving SLBMs as the excluded dummy. In the context of the model one would expect that bombers, because they operate outside the Soviet Union would be both easily detectable and open to external verification by the US Air Force and Civilian authorities. In contrast, ICBMs were stationed within the Soviet Union and so could only be observed by spy planes or satellites, both of which were under the purview of the intelligence agencies. Finally, submarine-based weapons, once operational, are difficult to observe. However, the fact that they operate in international waters means that they are somewhat open to external verification. In addition, while it might be difficult to observe submarines once they are launched, it may be less difficult to determine the number under construction at any one time.

In addition to weapon fixed effects we also need to control for the fact that the CIA’s ability to monitor the Soviet Union increased over time as better technology became available. Unfortunately, a detailed description of the monitoring technology available to the CIA has not been declassified. One exception is that it is now known that the United State’s most advanced spy satellite, Keyhole, was launched in 1976 and we control for this by including a dummy variable for years after 1976. To capture other changes in technology we include a linear time trend in our model and examine the model’s robustness by allowing for a quadratic trend.\footnote{We have tried to obtain alternative proxies for the quality of the intelligence services’ monitoring technology. Unfortunately, the secrecy around the intelligence community has made this impossible. For example, the CIA’s budget is still classified and has only been officially revealed for two years of its existence: 1997 and 1998 when it was, respectively $26.6bn and $26.7bn (CIA (2012)). Even if the total operating budget were known it might not be of much use: a large proportion of the CIA’s budget was spend on foreign interventions and other clandestine operations rather than intelligence gatherings (Weiner (2007)). Hence, changes in the operating budget would more likely reflect changes in the CIA’s ability to conduct special operations rather than gather intelligence.}

We also control for the length of the prediction being made, $t’ - t$. Intuitively, the fact that long-term forecasts are inherently more subjective means it is harder for a consumer to verify when an erroneous prediction is due to bias or simply an honest mistake. This implies that the value of $\mu$ is lower for longer-term predictions and so these predictions may be more sensitive to presidential ideology. To capture this effect we also generate the dummy variable forward that equals one if $t’ > t$ and zero otherwise and use its interaction with the ideology measures as an additional regressor. One can think of forward predictions as being less subject to external verification (in the language of the model they correspond to periods with $\mu = \mu_L$). As a result, if the conditions underlying Prediction 3 hold, our theory would predict that the interaction between forward and either of our ideology measures should be positive.

Superpower Relations

To try and rule out a number of alternative explanations for why presidential ideology might be correlated with intelligence errors we also test to see whether our results are robust to controlling for relations between the two superpowers. We do this by including the reading on the “Doomsday Clock” as an additional control variable. This measure was produced throughout the Cold War by the Bureau of the Atomic Scientists and measures in “minutes to midnight” the Bureau’s subjective view of how close the world was to Armageddon. High values therefore correspond to periods of relative calm and low values periods of relative tension. The resulting series, minutes to midnight, is plotted in Figure 4

While, in the post Cold War world, the Bureau has started to expand the measure’s definition to include issues like nuclear proliferation, terrorism and global warming, the only criterion during the Cold War was the proximity of nuclear war. As a result the key driver in the reports was the level of animosity between the NATO and Warsaw Pact countries and so we argue it is a good proxy for the state of inter-Superpower relations (BoAS (2007))).
Endogeneity Issues and Empirical Approach

Estimation

The baseline equations we wish to estimate is derived by approximating equation (2) with a linear functional form:

$$\text{reported}_{it} - \text{actual}_{it} = \beta_0 + \theta_t \beta \theta + (t' - t) \beta_1 + x'_t \beta + \delta_i + \epsilon_{it}$$  \hspace{1cm} (5)

Where $\theta_t$ is our measure of ideology (either nominate or text score), $x_t$ is a vector of controls which may contain a proxy for the state of inter-superpower relations and $\delta_i$ is a weapon fixed effect. Our theory predicts that $\beta_0 > 0$: that there is a causal, positive relationship between the degree of overestimation of Soviet capabilities in intelligence reporting and presidential ideology.

Endogeneity Issues

In this section we discuss the potential endogeneity issues that might prevent OLS from delivering consistent estimates of $\beta_0$. Perhaps the most obvious source of endogeneity when using the president’s nominate score as a proxy for ideology is measurement error: rather than observing each President’s true belief of the threat posed by the Soviet Union we instead observe their position on a liberal-conservative scale. To the extent that conservativeness does correspond to hawkishness our estimates will be biased towards zero. The goal of our text-based measure of hawkishness is to mitigate this problem, but it will again suffer from measurement error if presidential speeches do not perfectly reflect a president’s underlying views.

A second issue is one of reverse causality. Any measure of Presidential ideology which is based on examining a president’s actions while in office will suffer from simultaneity bias because a president could have been influenced by the information he received from the intelligence services. This is true of the nominate scores which are based on the congressional roll call votes on which each President publicly took a position. However, because the variable text score is based on speeches made before each president got security clearance, we are able to discount this concern as there is no possibility of the content of the reports having any direct effect on the measure.

Another serious concern is the potential for a selection effect that erroneously induces a positive correlation between CIA forecast errors and presidential ideology. The argument for how this might come about is as follows: suppose that hawkish presidents are more likely to be elected at times of greater public unease about the threat posed by the Soviet Union. In addition, suppose that this unease is also felt by intelligence analysts and causes them to excessively scale up their estimates of Soviet military strength. Then we would pick up an erroneous effect of presidential ideology: essentially at times of paranoia we would expect to see hawkish presidents winning elections and CIA analysts overestimating Soviet military strength.

A similar story, also running through the level of public paranoia, is that hawkish presidents act to worsen the state of relations between the superpowers and that analysts again excessively increase their forecasts of Soviet military strength. To ensure our results are not being driven by changes in the level of tension between the superpowers we add the variable minutes to midnight described above as a control variable.

Of course, finding that the correlation with presidential ideology is robust to this does not guarantee that these two mechanisms are not a factor. However, we would argue that the presence of either mechanism would be interesting in itself. Because we are using the forecast error as a dependent variable, finding that changes in minutes to midnight had a statistically significant effect would imply that the CIA was making a systematic mistake by giving too much weight to changes in the overall climate rather than focusing on the actual data of interest.

Even if one could prove the existence of a positive, causal effect of presidential ideology on report errors, one might still worry about whether the mechanism behind this effect was actually the one described in our model. One obvious alternative hypothesis would be that, rather than reflecting the CIA pandering to the president by skewing their reports, it results from hawkish Presidents having an agenda that they wanted to pursue (such as increased defence spending) and ordering the CIA to give evidence supporting this agenda.

The nature of the reports suggests this wasn’t the case. All of our data is taken from Top Secret documents to which only the President himself, the Joint Chiefs of Staff and the National Security Council had access.
There is therefore no obvious reason why each President would deliberately distort the information in the reports as they were not intended to be seen by anyone outside the administration.

Even if we allow for the possibility of legacy concerns which meant that presidents were unwilling to have the data in their personal reports differ markedly from that given to people outside the administration it is still difficult to see how the effect could have come about: without the end of the Cold War these reports would almost certainly still be classified and, even given the early date of their release, most of the Presidents in our sample were dead by the time the documents were in the public domain.

One might also be concerned that any positive effect results from turnover in agency staff rather than pandering. If, whenever a new president entered office, senior analysts were replaced with people who shared the new president’s beliefs, we might observe an effect of ideology on report errors. However, this would not be because the analysts’ incentives have changed as in our model, but simply because a different set of people are writing the reports.

To deal with this issue we rely on historical evidence to construct a simple test. The broad consensus among historians is that appointments to the CIA prior to 1977 were relatively apolitical (Andrew (1995)). For example, DCI Richard Helms served in both the Johnson and Nixon administrations. In addition, prior to 1977, there was not a single case of a DCI being replaced immediately after an election. However, this changed when, upon taking office, Jimmy Carter dismissed George H.W. Bush from his position on party political grounds and despite him having been in the post for only a few months. Subsequently it has became customary for new Presidents to make changes to the intelligence agencies’ senior hierarchy and in particular, to replace the DCI. As a result, if it were really the case that an estimated, positive effect of ideology were driven by staff turnover we would expect it to disappear or at least diminish when we restrict our attention to just the pre-Carter years.

A final concern might be that the intelligence community could itself have had an agenda (such as higher defence spending). If this were the case then intelligence analysts might give more upwardly biased reports to relatively dovish presidents so as to influence their choice of policy. While we cannot exclude this possibility, this mechanism would act against us finding an effect and so, to the extent this effect existed, our results can be considered a lower bound.

Standard Errors

Because both our measures of presidential ideology are fixed over time and hence are constant within a presidency, we cluster our standard errors at the president level. Imposing this error structure implies that we effectively have only seven independent observations: one for each president in our sample. Naturally the result is to increase the size of our standard errors and so substantially reduce the power of our estimates. This in turn makes it significantly less likely for us to find evidence in support of our hypothesis. We argue that the fact that we still find a statistically significant, positive effect of ideology on intelligence reporting under these stringent assumptions is strong evidence for our hypothesis.

6 Results

Graphical Evidence

Before presenting our regression results we first present some simple, graphical evidence. Because of the fact that we have so few presidents in our sample relative to the number of observations it is important to ensure that any positive effect of ideology we find is also visible when we collapse our data to the presidency level.

19The results remain significant if we alternatively cluster at the report-year level or dispense with clustering and use het-
eroskedasticity robust standard errors. When calculating test statistics and confidence intervals we use a t-distribution with degrees of freedom equal to $P - 1$ where $P$ is the number of Presidents in our sample. Unfortunately we are unable to use the biased reduced linearization method of Bell and McCaffrey (2002) to correct for the small number of clusters because we are interested in a variable, Presidential ideology, which is fixed at the presidency level. Similarly, we cannot use the degrees of freedom correction of Donald and Lang (2007) because we have insufficient Presidents in our sample.

18
This is done in Figures 2 and 3. The vertical axis measures the average raw error made across all weapon systems and report years for each president, while the horizontal axis measures presidential ideology as measured by, respectively, \textit{nominate} and \textit{text score}.

In both cases there is clear evidence of a positive correlation between presidential ideology and intelligence errors. While the extremely large average error under Eisenhower represents something of an outlier, this first pass examination of the data is broadly supportive of our hypothesis.

**Baseline Results**

The first two columns of Table 3 give our baseline results when we regress the report error on presidential ideology as measured by \textit{nominate} and \textit{text score} respectively. In both cases the coefficient on presidential ideology is positive and significant and so is consistent with Prediction 1 of our model. The point estimates are also economically significant: they indicate that a one standard deviation increase in presidential ideology will increase the report error for a given weapon system by sixty units.

We also find that the coefficient on prediction length is statistically insignificant. Although the standard errors on this estimate are large it is consistent with Prediction 2 of our model that states that changes in the probability of external verification should not have a level effect.

Turning to the other coefficients we see that the weapon dummies are never individually significant at the ten percent level. The coefficient on the time trend is negative in both cases and statistically significant in the first specification. This is consistent with an improvement in the intelligence agencies’ monitoring technology over time. One surprising result is that the coefficient on \textit{key hole} is actually positive. This probably results from the fact that the dummy is effectively controlling for all unobserved intercept changes that occurred after 1974 and so should not be interpreted as a causal effect of the introduction of the satellite system.

**Interaction Between Prediction Length and Hawkishness**

The third and fourth columns of Table 3 examine whether there is an interaction effect that is consistent with Prediction 3 of our model.

Intuitively, longer-term predictions are inherently more noisy and subjective and hence it is less likely that information will be revealed that contradict the analysis behind them. In the language of our model, long-term reports have a lower probability of external verification (a lower $\mu$). This implies that, if the condition in Prediction 3 holds, we should see changes in ideology having a larger effect on the report error in forward predictions than in in-year prediction.

To capture this empirically we interact the Presidential ideology measures with a dummy variable, \textit{forward}, that equals one for observations with $t' > t$ and zero otherwise. Because we are still controlling for the prediction length, a positive coefficient on this interaction would indicate that the effect of ideology on reporting is larger when the agency makes longer-term predictions and so would be consistent with the amplification effect outlined in Prediction 3.

The results in columns 3 and 4 are consistent with this story: in both cases the interaction term is positive and significant at the ten percent level. In the specification using the \textit{text score} measure of presidential ideology, the coefficient on \textit{text score} actually loses significance, suggesting that all of the ideology effect runs through longer term reports (although an alternative explanation is collinearity between text score and its interaction.) In both cases text score and its interaction are jointly significant at the five percent level. In summary, there is clear evidence of a positive interaction between ideology and the forward dummy which is consistent with our theoretical model.

**Agency Turnover**

As discussed in section 5, a potential alternative to our pandering story is that new presidents install like-minded staff in the intelligence agencies and this induces a correlation between report errors and presidential ideology. We check the sensitivity of our results to this mechanism by examining whether the results persists
when we restrict our sample to the pre-Carter years where the convention was for presidents to refrain from making changes to the CIA’s senior hierarchy upon taking office.

These results are displayed in columns (1) and (2) of Table 4. In both cases the coefficients on the ideology measure remain significant at the ten percent level and the point estimates actually increase. This is difficult to reconcile with the idea that the baseline results are driven by staff turnover.

Controlling For US/Soviet Relations

As explained in section 5, the positive effect of presidential ideology could potentially result from alternative mechanisms running through the state of relations between the superpowers. For this reason we examine the robustness of our results to controlling for our proxy for superpower relations, minutes to midnight. The unconditional correlation between minutes to midnight at a president’s election date and the president’s ideology is approximately −0.5 when we use nominate and −0.4 when we use text score. This is consistent with the selection story outlined in section 5 that more hawkish president’s at times of increased tension (and the doomsday clock records there being fewer minutes to midnight).

Ideally we would like to see that the positive, significant effect of presidential ideology persists even when we control for the state of relations between the superpowers. However, finding such a result wouldn’t tell us very much if minutes to midnight had very little predictive power over intelligence errors when we did not control for presidential ideology. If this were the case it could simply be that minutes to midnight was a bad proxy and measurement error was driving the coefficient estimate to zero.

Turning to the results in column (3) of Table 4 we see that this is not the case: when we do not control for presidential ideology, minutes to midnight has a negative and highly statistically significant effect on intelligence errors. This is consistent with the idea that, in times of reduced tension (i.e. times with a higher reading on the doomsday clock) the US intelligence community produced less pessimistic reports.

However, once we include the ideology scores the results change markedly. When we measure presidential ideology by nominate the coefficient on ideology remains positive and significant while the coefficient on minutes to midnight loses significance. Similarly, when we measure presidential ideology by text score, the coefficient on ideology remains statistically significant. However, the coefficient on minutes to midnight also remains negative and significant, so we cannot exclude the possibility that the level of international tension contributes to the content of the reports.

In summary, the positive effect of presidential ideology is robust to controlling for the relations between the superpowers. However, in the text score specification, it is not possible to exclude the possibility that changes in superpower relations did affect the CIA’s report errors. In our defence, we would argue that this, somewhat ambiguous, result is still of interest: if the correlation between superpower relations and report errors reflects a causal effect it implies that the intelligence community was systematically overcompensating for changes in the overall world situation when constructing their reports. This effect has important implications for the quality of intelligence reporting.

Evolution of the Ideology Effect Over Time

Table 5 explores whether there is any variation in the ideology effect over time. We do this by restricting the data to each president’s first term in office and then interacting each president’s ideology measure with a set of year dummies. In addition, we control for the number of years to the next election with the variable elec time. The results from using nominate are in column (1) and those from text score are in column (2).

In both cases we see evidence of a phase-in effect: the effect of ideology is positive in all cases, but is statistically insignificant in the president’s first year of office. This can be reconciled with our model if it takes some time for intelligence analysts work out the beliefs of a new president. Alternatively, it could be that there is a reputational benefit to ensuring consistency with previous reports: intelligence analysts might not make abrupt changes in their reports when a new president takes office so as to protect themselves from accusations of pandering. Finally, it could simply reflect a time lag in producing reports: the reports given to a new president may be based in part on analysis conducted during their predecessor’s time in office.
One mechanism that might be of interest is the possibility that the degree of sensitivity of analysts’ reports to presidential ideology could depend on the president’s electoral prospects. Our analysis, available on request, finds little evidence of this: no robust results are found when we interact presidential ideology with presidential approval ratings, the time until the next election etc.

7 Robustness

In this section we demonstrate that our baseline results are robust to changes in the specification. In particular, we allow for lagged Soviet strength to affect the intelligence reports in a given year and allow for a quadratic rather than linear trend. In addition, we examine whether the results change when we use the percentage error in the reports rather than the raw error suggested by our model.

Including Lagged Soviet Strength and a Quadratic Trend

Specifications (1) and (2) in Table 6 present the results when we include the lagged actual number of Soviet weapons as an additional control. This should control for the possibility that past Soviet actions could affect analysts’ interpretation of new evidence. For example, if the Soviet arsenal was known to be large in the previous period analysts may make more pessimistic interpretations of the available evidence. As can be seen the results are qualitatively identical and quantitatively similar to the baseline results and the estimated effect of presidential ideology is positive in all four cases. This remains true if we instead include the lag of the reported number of Soviet weapons or lags of their growth rate.

Specifications (3) and (4) present the results when we fit a quadratic trend in place of the linear trend used in the baseline specifications. Again the coefficients on the ideology measures remain positive and actually increase in absolute value relative to the baseline case. However, the size of the standard errors increase and both are only significant at the ten percent level. Finally, comparing the $R^2$ to those in the baseline specification suggests that the quadratic term adds little in terms of predictive power.

Results with a Percentage Error Specification

In the above analysis we have used the raw report error as our dependent variable as this corresponds with the predictions of our theoretical model. As can be seen in Figures 2 and 3 the absolute errors made by the intelligence services were of a similar order of magnitude throughout the Cold War. However, for completeness we repeat our analysis using the report error as a proportion of the true stock of Soviet weapons as our dependent variable:

$$\frac{\text{reported}_{it} - \text{actual}_{it}}{\text{actual}_{it}}$$

Given our previous observation that the average error in the reports were of a similar order of magnitude, the effect of this approach is to significantly amplify the level of bias in the early stages of the Cold War when the stock of Soviet weapons was relatively small.

The counterpart to the results in Table 3 are displayed in Table 7. In all cases the coefficient on the ideology measures are positive. However, the estimates using text score measure are not statistically significant at any reasonable level of significance. In addition, the point estimates suggest an ideology effect that is implausibly large. We will argue in the next subsection that this is due to the excessive weight given to the reports made during the Eisenhower presidency.

Table 8 checks the robustness to removing the pre-Carter years and controlling for minutes to midnight. The specifications in columns (1) and (2) show that, once again, removing the post-Carter years does not substantively change the results. However, the results from controlling for minutes to midnight are less encouraging; in both cases the ideology measure loses significance, although it remains positive when we measure presidential ideology by nominate.

Clearly the results under the percentage error specification are less clear-cut and this is particularly true when we use the text score measure. Even when we do find a positive and significant effect, the magnitudes are
implausibly large. However, we argue that this anomaly results primarily from the way the percentage error specification weights observations under President Eisenhower and show that removing these observations delivers results that are much more in line with our theory.

The reason for why this would be the case can be seen in Figures 2 and 3. In both cases President Eisenhower is a notable outlier: the average raw report error was higher under his presidency than any other. Because, Eisenhower was president during the early stage of the Cold War when the Soviet arsenal was at its smallest the discrepancy is even larger when we use percentage report errors.

This is less of an issue when we measure presidential ideology by *nominate*: as can be seen from Table 2 this measure records Eisenhower as having been relatively right-wing. In contrast, our *text score* measure records Eisenhower as having been relatively dovish and hence the Eisenhower years become even more anomalous.

The results in Table 9 illustrate the extent to which the Eisenhower years may be confounding our analysis when we estimate the percentage error specification. Columns (1) and (2) report the baseline results, column (3) interacts *text score* with a forward dummy and column (4) includes *minutes to midnight* as an additional control.

The first point to note is that the results are much more in line with our theory: the coefficient on text score is positive and significant in all cases, as is its interaction with *forward*. While *nominate* is now marginally insignificant at the ten percent level it remains positive. Finally, the point estimates are much more plausible: they suggest that a one-standard deviation change in ideology leads to a seven to ten percentage point increase in the forecast error rather than the sixty to seventy percent effect suggested by the baseline specification.

It is possible to make a case for excluding the Eisenhower years. Firstly, they represented the zenith of anti-Soviet paranoia in the US. In addition, the US’s ability to monitor the Soviet Union was very much in its infancy: the first overflight of the Soviet Union by a US spy plane occurred in 1956, half-way through the Eisenhower presidency (Andrew (1995)).

8 Conclusion

This paper conducts the first formal, quantitative analysis of intelligence failures. We extend existing theories of media bias to predict that career concerns on the part of intelligence analysts could lead to them tailoring intelligence reports to pander to a sitting president’s prior beliefs. We then tested this theory by constructing a unique measure of intelligence failures by comparing US intelligence reports on the size of the Soviet strategic nuclear arsenal during the Cold War to credible, post Cold War estimates of the actual number of Soviet strategic weapons.

We find robust evidence of a positive effect of presidential “hawkishness” as measured by existing ideology measures and a new measure based on text analysis on intelligence errors that is consistent with our model. We also find evidence that errors in longer-term reports are more sensitive to presidential ideology in a manner consistent with our theory. Institutional evidence draws doubt on alternative explanations based on collusion between presidents and the intelligence services or turnover in agency staff following presidential elections. Finally the results are robust to controlling for alternative mechanisms working through changes in inter-Superpower relations.

In conclusion, the paper produces empirical evidence of the existence of a mechanism by which persistent intelligence failures can and occur. As such it has important implications for issues of foreign policy and national security. The fact that we find evidence of pandering in intelligence reports which are both quantitative in nature (and so are more open to external verification), and relate to such high-stakes issues (which we might hope would induce analysts to sacrifice their career concerns for the national good), suggests that this mechanism may play an even larger role in intelligence reports where these constraints do not exist.
References


Appendix A

Strategies, Beliefs and Equilibrium Definition

Beliefs: Because of the dynamic nature of the game all strategies and beliefs could, in principle, depend on the calendar time of the game, t. Because we are focussing on stationary equilibria and to economise on notation we supress this dependence where possible.

The incumbent President has beliefs over both the state of the world and the quality of the incumbent director. The President’s prior belief as to the state of the world is simply his type, θ. We then denote his posterior belief after he receives report r by \( \hat{\theta}(r, \theta, \mu) \).

We denote the President’s prior belief that the director is a good type by \( \lambda_0(\theta) \). As well as calendar time this could also depend on his prior belief as to the state, \( \theta \), (in which case presidents of different types would hold different prior beliefs as to the director’s quality and play would not be symmetric). The President’s posterior belief as to the quality of the director will depend on both the report, \( r \), and any feedback received as to the true state. As in the main text we denote the possible forms of feedback by \( F \in \{\emptyset, L, H\} \) and the President’s posterior belief by \( \lambda^\theta(r, F, \mu) \).

The director also has beliefs over the state of the world. Because all directors have the prior belief that the states are equally likely, the normal director’s posterior belief that the state is \( H \) is \( \pi \) if \( s = H \) and \( 1 - \pi \) if \( s = L \). Good directors receive a perfect signal and so their beliefs are that \( s = S \) with probability one. We denote the normal director’s beliefs by \( \theta_D(s) \).

Strategies: Turning to the players’ strategies, we begin with the agency director. Because good types follow a behavioral rule, we can focus our attention on the strategy of a normal director. As in the text, we denote
a normal director’s strategy by \( \sigma^\theta(r, \mu) \) which gives the probability that the director makes report \( r \) to a president of type \( \theta \) after observing signal \( s \) when the probability the state is revealed is given by \( \mu \). Without loss of generality we restrict our attention to the case where \( \sigma^\theta_H(H, \mu) \geq \sigma^\theta_L(H, \mu) \) for all values of \( \theta \) and \( \mu \).²⁰

The President’s strategy determines his action, \( a \), and his reappointment decision, which we denote by the binary variable \( R \in \{0, 1\} \). Given our assumption that the President receives a positive payoff if his action matches the state and zero otherwise, he will choose the action which corresponds to the state he believes to be more likely after he has received the director’s report:

\[
a^\theta(r, \mu) = \begin{cases} H & \text{if } \hat{\theta}(r, \theta, \mu) \geq 0.5 \\ L & \text{otherwise} \end{cases} \tag{6}
\]

The President’s reappointment decision will depend on his continuation value in his legacy period, which in turn depends on the next President’s type and beliefs, the strategy of a normal type, \( \sigma \) and his belief that the incumbent director is good. Denoting the (unknown) type of his successor by \( \theta' \) the President’s expected legacy payoff from reappointing the incumbent director is given by \( Ev(\lambda^\theta(r, F, \mu), \sigma, \theta') \). If, on the other hand, he chooses to replace the director, the replacement will be a good type with probability, \( \gamma \) and his expected payoff will be given by \( Ev(\gamma, \sigma, \theta') \).

From equation (1) it follows that the president’s equilibrium reappointment rule is given by:

\[
R = \begin{cases} 1 & \epsilon \leq Ev(\lambda^\theta(r, F, \mu), \sigma, \theta') - Ev(\gamma, \sigma, \theta') \\ 0 & \text{otherwise} \end{cases} \tag{7}
\]

Before \( \epsilon \) is known the incumbent director calculates his probability of reappointment as a function of his report, the probability of the state being revealed and the director’s beliefs as:

\[
G(Ev(\lambda^\theta(r, F, \mu), \sigma, \theta') - Ev(\gamma, \sigma, \theta'))
\]

Given this reappointment probability the agency director chooses his reporting strategy to maximise his expected utility. In other words the agency director’s strategy solves the following value function:

\[
V(\theta, s, \mu) = \max_{\sigma_s^\theta} A + G(Ev(\lambda^\theta(r, F, \mu), \sigma, \theta') - Ev(\gamma, \sigma, \theta'))EV(\theta', s', \mu') \tag{8}
\]

A symmetric, stationary MPE of the game is then defined as follows:

**Definition 1.** A symmetric, stationary MPE is a collection of President and agency director strategies: \( (a^\theta(\cdot), R(\cdot), \sigma_s^\theta(\cdot)) \) and President and agency director beliefs: \( (\hat{\theta}(\cdot), \lambda_0(\cdot), \lambda^\theta(\cdot), \hat{\theta}_D(\cdot)) \), such that:

1. All players’ strategies and beliefs are stationary (i.e. they are all independent of calendar time)
2. Presidents have symmetric prior beliefs over the director’s quality (i.e. \( \lambda_0(\theta) = \lambda_0 \) for all \( \theta \)
3. The reporting strategy of a normal agency director maximises (8)
4. Each type of President chooses his action and reappointment decision optimally according to (6) and (7)
5. Presidents’ beliefs are correct and updated using Bayes’ rule where possible

In order to prove Proposition 1 it is first necessary to prove two lemmas. The first determines the condition under which the director’s reputation will matter for his reappointment probability while the second characterises how the director’s reputation depends on his reports.

²⁰This is without loss of generality because to do otherwise amounts to relabeling the two states.
The Reappointment Decision

**Lemma 1.** The director’s reappointment probability is strictly increasing in the President’s posterior belief over quality if and only if, in equilibrium, there exists some range of \( \theta \) and value of \( \mu \) for which \( \theta(L, \theta, \mu) \geq .5 > \theta(L, \theta, \mu) \). If such a range does not exist then the director’s survival probability is given by \( G(0) \) in every period.

**Proof.** If we define the indicator function, \( I(\theta, \mu) \), which equals one if a president of type \( \theta \)’s equilibrium actions are responsive to reports when the probability of external verification of the state is given by \( \mu \), the function \( K(\theta) \), which gives the expected payoff to the current president given that his successor is of type \( \theta \) and does not condition his action on the director’s reports and the probability \( p(\theta, \mu) < 1 \) which gives the equilibrium probability that a normal director gives the correct report as a function of \( \theta \) and \( \mu \). We can calculate the sitting president’s expected legacy payoff if he reappoints the director as a function of his belief that the incumbent director is good, \( \lambda_1 \).21

\[
Ev(\lambda_1) = \sum_\mu P(\mu' = \mu) \int_0^1 \left[ (\lambda_1 + (1 - \lambda_1)p(\theta', \mu'))I(\theta', \mu') + (1 - I(\theta', \mu'))K(\theta') \right] f(\theta') d\theta'
\]

In the cases where \( I(\theta, \mu) = 1 \) and the president’s successor listens to the report, his expected payoff when he reappoints the director is strictly increasing in his belief that the director is a good type. This is because, when the president’s successor listens to reports, reappointing a good type ensures that the correct action is taken. In the cases where \( I(\theta, \mu) = 0 \) the incumbent president’s legacy payoff is independent of the director’s quality because any reports he makes will simply be ignored. It therefore follows that, as long as there is some positive probability that the President’s successor listens to reports (and \( I(\theta, \mu) = 1 \) for some \( \mu \) and \( \theta \) ), reappointing a good type will increase the probability that the correct action is taken and the President’s expected payoff from reappointing the director will be strictly increasing in \( \lambda_1 \).

The President’s expected legacy payoff if he removes the director is calculated by setting \( \lambda_1 = \gamma \), the ex-ante probability that a director is high quality:

\[
Ev(\gamma) = \sum_\mu P(\mu' = \mu) \int_0^1 \left[ (\gamma + (1 - \gamma)p(\theta', \mu'))I(\theta', \mu') + (1 - I(\theta', \mu'))K \right] f(\theta') d\theta'
\]

From equation (7) the director’s reappointment probability will be given by \( G(v(\lambda_1) - v(\gamma)) \). As long as \( I(\theta, \mu) = 1 \) for some range of \( \theta \) and \( \mu \) this will be strictly increasing in \( \lambda_1 \). Otherwise, the expected legacy payoffs will be the same whether the director is reappointed or removed and the probability the director is reappointed is given by \( G(0) \). \( \square \)

Reporting Histories and Reputation

**Lemma 2.** Suppose the incumbent President’s prior belief over quality is given by \( \lambda_0 \). Then a President of type \( \theta \)’s beliefs after report, \( r \in \{L, H\} \), feedback \( F \in \{L, H, \emptyset\} \) and probability of verification, \( \mu \) have the following properties:

1. \( \lambda^0(H, \emptyset, \mu) \) and \( \lambda^0(H, H, \mu) \) are strictly increasing in \( \theta \) and decreasing in \( \sigma_H^0(H, \mu) \) and \( \sigma_L^0(H, \mu) \)

2. \( \lambda^0(L, \emptyset, \mu) \) and \( \lambda^0(L, L, \mu) \) are strictly decreasing in \( \theta \) and increasing in \( \sigma_H^0(H, \mu) \) and \( \sigma_L^0(H, \mu) \)

3. Suppose \( \sigma_H^0(H, \mu) = 1 \) and \( \sigma_L^0(H, \mu) = 0 \). Then \( \lambda^0(H, \emptyset, \mu) \geq \lambda^0(L, \emptyset, \mu) \) when the President is a hawk and \( \lambda^0(H, \emptyset, \mu) < \lambda^0(L, \emptyset, \mu) \) when he is a dove

4. \( \lambda^0(L, H, \mu) = \lambda^0(H, L, \mu) = 0 \)

---

21 It must always be the case that \( p(\theta) < 1 \) because, even in the best possible outcome where the normal type truthfully reports his signal, the probability that the President takes the correct action is \( \pi < 1 \)
Proof. A President of type $\theta$’s beliefs after each report when the state isn’t revealed are given by:

$$
\lambda^\theta(H, \emptyset, \mu) = \frac{P(\text{Good} \mid H, \emptyset) P(\text{Good})}{P(H, \emptyset)} = \frac{\lambda \theta_0}{\lambda \theta_0 + (1 - \lambda \theta)(\theta \pi + (1 - \theta)(1 - \pi))(\theta \pi \theta^\theta_H(H, \mu) + (\theta(1 - \pi) + (1 - \theta)\pi)(1 - \sigma^\theta_L(H, \mu)))}
$$

and:

$$
\lambda^\theta(L, \emptyset, \mu) = \frac{(1 - \theta)\lambda_0}{\lambda \theta_0(1 - \theta) + (1 - \lambda_0)((\theta \pi + (1 - \theta)(1 - \pi))(1 - \sigma^\theta_H(H, \mu)) + (\theta(1 - \pi) + (1 - \theta)\pi)(1 - \sigma^\theta_L(H, \mu)))}
$$

While in the cases where the state is revealed they are given by:

$$
\lambda^\theta(H, H, \mu) = \frac{\theta \lambda_0}{\lambda_0 \theta + (1 - \lambda_0)((\theta \pi \sigma^\theta_H(H, \mu) + \theta(1 - \pi)\sigma^\theta_L(H, \mu))}
$$

$$
\lambda^\theta(L, L, \mu) = \frac{(1 - \theta)\lambda_0}{\lambda_0(1 - \theta) + (1 - \lambda_0)(1 - \sigma^\theta_H(H, \mu)) + \pi(1 - \sigma^\theta_L(H, \mu))}
$$

$$
\lambda^\theta(H, L, \mu) = \lambda^\theta(L, H, \mu) = 0
$$

Differentiation shows that $\lambda^\theta(H, \emptyset, \mu)$ and $\lambda^\theta(H, H, \mu)$ are strictly increasing in $\theta$ and strictly decreasing in $\sigma^\theta_L(H, \mu)$ and $\sigma^\theta_H(H, \mu)$ while $\lambda^\theta(L, \emptyset, \mu)$ and $\lambda^\theta(L, L, \mu)$ are strictly decreasing in $\theta$ and strictly increasing in $\sigma^\theta_H(H, \mu)$ and $\sigma^\theta_L(H, \mu)$. The final property in the Lemma is established by comparing $\lambda^\theta(H, \emptyset, \mu)$ and $\lambda^\theta(L, \emptyset, \mu)$ at $\sigma^\theta_H(H, \mu) = 1$ and $\sigma^\theta_L(H, \mu) = 0$.

Proof of Proposition 1

With the results in Lemmas 1 and 2 established we can prove Proposition 1. To do this we initially assume that all players’ strategies and beliefs are stationary and that all types of President share the symmetric prior belief over quality, $\lambda_0$. In addition, we assume that there is indeed some range of types of Presidents for whom $\theta(\theta, H, \mu) > .5 > \theta(L, H, \mu)$ so that reputation matters in equilibrium.

We then derive the unique outcome under these assumptions and confirm that 1) under the resulting reporting strategies there is indeed a range of Presidents who condition their action on the director’s reports and 2) A symmetric, stationary prior, $\lambda_0 \in (0, 1)$, exists which supports this behavior as an equilibrium.

In the analysis that follows we abuse notation and write the probability of reappointment, $G(v(\lambda_1) - v(\gamma))$, as $G(\lambda_1)$.

Step 1: Characterise Reporting Behaviour Under These Assumptions

Case a) President is a Hawk

Given that $\theta \geq .5$ the agency director chooses his report to maximise his expected payoff. If we denote the director’s continuation payoff if he is reappointed by $\bar{V} = EV(\theta', s', \mu')$ then, if he receives a signal $s = H$ and truthfully reports $r = H$ when the probability of external verification is $\mu$, his expected payoff is given by:

$$
V(\theta, s = H, r = H, \mu) = A + [(1 - \mu)G(\lambda^\theta(H, \emptyset, \mu)) + \mu G(\lambda^\theta(H, H, \mu)) + (1 - \pi)G(0)] \bar{V}
$$

He receives the ego rent $A$ in the current period. If the state is not revealed then the President’s believes he is a good type with probability, $\lambda^\theta(H, \emptyset, \mu)$ and he is reappointed with probability $G(\lambda^\theta(H, \emptyset, \mu))$. If the state is revealed then, with probability $\pi$ his initial signal was correct, the true state is indeed $H$ and he is
so will hold the same equilibrium beliefs when the director deviates. However, with probability $1-\pi$ the true state is actually $L$, the director reveals himself to be a normal type and he is reappointed with probability $G(0)$.

Compare this to his expected payoff if he reports $L$ when his signal is $H$:

$$V(\theta, s = H, r = L, \mu) = A + \left[ (1-\mu)G(\lambda^0(H, H, \mu)) + \mu((1-\pi)G(\lambda^0(L, L, \mu)) + \pi G(0)) \right] \bar{V}$$

Because he is reporting against his signal, his probability of matching the state is only $1-\pi$. Comparing these two payoffs we can write the differential gain to reporting $H$ rather than $L$ after receiving an $H$ signal as a function of $\mu$:

$$\Delta^H_H(H, \mu) = [(1-\mu)G(\lambda^0(H, H, \mu)) - G(\lambda^0(H, \emptyset, \mu))] + \mu \left[ (\pi G(\lambda^0(H, H, \mu)) + (1-\pi)G(0)) - ((1-\pi)G(\lambda^0(L, L, \mu)) + \pi G(0)) \right] \bar{V}$$

(9)

Repeating this analysis for when the director received an $L$ signal gives the differential gain of reporting $H$ as:

$$\Delta^L_L(H, \mu) = [(1-\mu)G(\lambda^0(L, H, \mu)) - G(\lambda^0(L, \emptyset, \mu))] + \mu \left[ (\pi G(\lambda^0(H, H, \mu)) + (1-\pi)G(0)) - ((1-\pi)G(\lambda^0(L, L, \mu)) + (1-\pi)G(0)) \right] \bar{V}$$

(10)

Examining these two expressions we see that the first term in each case is identical. However, because $\pi > .5$ the second term in $\Delta^H_H(H, \mu)$ is strictly larger and so $\Delta^H_H(H, \mu) > \Delta^L_L(H, \mu)$. This implies that, if a normal director ever reports $L$ to a hawkish president after receiving signal $H$, he must always report $L$ after receiving signal $L$.

Suppose that this were the case in equilibrium. Then one can use Lemma 2 to show that it implies that $\lambda^0(H, \emptyset, \mu) > \lambda^0(L, \emptyset, \mu)$ and $\lambda^0(H, H, \mu) > \lambda^0(L, L, \mu)$. However, if this were the case then $\Delta^H_H(H, \mu) > 0$ and so the normal director would always report $H$ to a hawkish president after observing signal $H$, a contradiction. This implies that we will never see misreporting of $H$ signals to hawkish presidents and so any bias will come from misreporting of $L$ signals.

The director’s strategy in the case where $\theta > .5$ can therefore be summarised by the probability of distorting $L$ signals, $\sigma^0_L(H, \mu)$. We first show that, in equilibrium, $\sigma^0_L(H, \mu) \neq 1$ and the director does not always misreport $L$ signals. Suppose this were the case: then $\lambda^0(L, \emptyset, \mu) = \lambda^0(L, L, \mu) = 1$ as only good types would ever report $L$. In order to show that this implies that the director has a profitable deviation we must first account for the off-equilibrium path history $(L, H)$ and do so by allowing the president’s belief to take the arbitrary value $\hat{\lambda} \in [0, 1]$

The expected payoff to the director from reporting $L$ after observing signal $L$ is then $A + \left[ (1-\mu)G(1) + \mu(\pi G(1) + (1-\pi)G(0)) \right] \bar{V}$ which is strictly lower for all values of $\hat{\lambda}$, hence it cannot be an equilibrium to always misreport $L$ signals.22

We now prove that, given our assumptions of a symmetric prior belief over quality, $\lambda_0$ and a type of president who conditions his action on the director’s reports, there exists a unique equilibrium probability of misreporting. To do this we note that $\Delta^0_H(L, \mu)$ is strictly decreasing in $\sigma^0_L(H, \mu)$. This follows because, from Lemma 2, $\lambda^0(H, \emptyset, \mu)$ and $\lambda^0(H, H, \mu)$ are strictly decreasing in $\sigma^0_L(H, \mu)$ while $\lambda^0(H, \emptyset, \mu)$ and $\lambda^0(L, L, \mu)$ are strictly increasing. Hence if $\Delta^0_H(L, \mu) \leq 0$ at $\sigma^0_L(H, \mu) = 0$ then this is the unique equilibrium outcome. Otherwise there exists a unique $\sigma^0_L(H, \mu) \in (0, 1)$ such that $\Delta^0_H(L, \mu) = 0$. To record the fact that this is the equilibrium strategy when the president is a hawk and that it depends on the president’s type, $\theta$ and the probability of external verification, $\mu$ we write the equilibrium probability of misreporting as $\sigma^0_L(\theta, \mu)$.

To prove that there exists a threshold level of the probability of external verification $\mu^*_\theta$ that will determine whether there is positive bias in the reports given to a President of type $\theta$ we use Lemma 2 to note that, when evaluated at $\sigma^0_L(H, \mu) = 0$, the first term of $\Delta^0_H(H, \mu)$ is positive and the second negative. Holding the

---

22 The value function in each case will be the same as future presidents do not observe the reports in the current period and so will hold the same equilibrium beliefs when the director deviates.
continuation value fixed and changing the probability of external verification in the current period it follows that there is a unique $\mu^*_5$ that ensures that $\Delta^5_H(H, \mu) = 0$. $\Delta^5_H(H, \mu)$ will therefore be positive when $\mu \leq \mu^*_5$. This implies that if $\mu_L, \mu_H \geq \mu^*_5$ there will be no misreporting to presidents of type $\theta$, if $\mu_L < \mu^*_5 \leq \mu_H$ there will be positive misreporting of $L$ signals to presidents of type $\theta$ in periods where $\mu = \mu_L$ and no misreporting when $\mu = \mu_H$. Finally, if $\mu_L, \mu_H < \mu^*_5$ there will be a positive probability of misreporting of $L$ signals in all periods where the president is of type $\theta$. Note that this does not prove a comparative static result with respect to $\mu_L$ or $\mu_H$, because changes in $\mu_L$ and $\mu_H$ will alter both the director’s continuation value and the President’s prior belief over quality. We return to this issue in the proof of Proposition 2.

**Case b) President is a Dove**

When $\theta < .5$ analogous reasoning leads us to find that the director will always report $L$ signals truthfully, but will distort $H$ signals with some unique probability $\sigma^*_5(\theta, \mu) \in [0, 1)$. To see this we write the net gain to reporting $L$ rather than $H$ after observing each signal:

$$\Delta^5_H(L, \mu) = \left[(1 - \mu) \left[ G(\lambda^5(H, \emptyset, \mu)) - G(\lambda^5(L, \emptyset, \mu)) \right] + \mu \left[ (\pi G(\lambda^5(L, \mu)) + (1 - \pi) G(0)) - ((1 - \pi) G(\lambda^5(H, \mu)) + \pi G(0)) \right] \right] \bar V$$

$$\Delta^5_H(L, \mu) = \left[ (1 - \mu) \left[ G(\lambda^5(L, \emptyset, \mu)) - G(\lambda^5(H, \emptyset, \mu)) \right] + \mu \left[ (1 - \pi) G(\lambda^5(L, L, \mu)) + \pi G(0) - (\pi G(\lambda^5(H, H, \mu)) + (1 - \pi) G(0)) \right] \right] \bar V$$

Similar analysis to that in the previous case shows that, in a SSMPE in which some types of presidents condition their action on the reports, $\sigma^*_5(L, \mu) = 1$ and normal directors will always truthfully report $L$ signals when the president is a dove. Bias in reports to dovish presidents therefore consist of at most misreporting of $H$ signals. For each combination of $\theta < .5$ and $\mu \in \{\mu_L, \mu_H\}$ there exists a unique equilibrium level of misreporting of $H$ signals which we denote by $\sigma^*_5(\theta, \mu) \in [0, 1)$. The equilibrium level of misreporting to a president of type $\theta$ is equal to zero if $\Delta^5_H(L, \mu) \leq 0$ at $\sigma^*_5(L, \mu) = 0$. Otherwise, $\sigma^*_5(\theta, \mu)$ is determined by the equilibrium condition $\Delta^5_H(L, \mu) = 0$. Finally, an identical argument can be used to prove the existence of an $\mu^*_3$ that determines whether or not there is a positive probability of misreporting.

An important point to note is that, for any $\theta \geq .5$, there exists a $\theta' < .5$ that is equidistant from one half. When all presidents have common priors over quality it follows that the problems facing the director in each case are equal and opposite. Therefore, in equilibrium $\sigma^*_5(\theta, \mu) = \sigma^*_5(\theta', \mu)$ and equilibrium reporting in each case will be biased an equal amount in opposite directions. This in turn implies a symmetry in the beliefs of presidents of type $\theta$ and $\theta'$:

$$\lambda^5(H, H, \mu) = \lambda^5(L, L, \mu)$$

$$\lambda^5(H, \emptyset, \mu) = \lambda^5(L, \emptyset, \mu)$$

This will be useful in demonstrating that our assumption that all presidents hold common prior beliefs over quality in equilibrium is valid.

**Step 2: Show That Some Presidents Respond to Reports**

We have uniquely pinned down reporting given that presidents have the symmetric, stationary prior belief over quality, $\lambda_0$, and that the director’s reappointment decision is increasing in his reputation. As explained in the text, this will be true if there is some probability that future President’s will base their action on the director’s reports (i.e. that there exists $\theta$ such that $H(\theta, \mu) > .5 > H(\emptyset, \mu)$).

A sufficient condition is that, for some $\theta > .5$ and $\mu$, $\theta(H, \emptyset, \mu) < .5$. This will be the case if:
\[
\hat{\theta}(L, \theta) = P(S = H \mid r = L) = \frac{P(r = L \mid S = H)P(S = H)}{P(r = L)} = \frac{(1 - \lambda_0)(1 - \pi)\theta^\sigma_L(L, \mu)\theta}{\lambda_0(1 - \theta) + (1 - \lambda_0) [\theta(1 - \pi)\sigma^\theta_L(L, \mu) + (1 - \theta)\pi\sigma^\theta_L(L, \mu)]} < .5
\]

Rearranging gives the condition:

\[(1 - \lambda_0)(\theta - \pi)\sigma^\theta_L(L, \mu) < \lambda_0(1 - \theta)\]

Because the right hand side is strictly positive and \(\pi > .5\), this will always be true for some range of values of \(\theta\) sufficiently close to .5. This implies that, given the reporting strategies derived above, the conditions in Lemma 1 are met and the director’s reappointment probability is increasing in his reputation.

**Step 3: Show that a Symmetric, Stationary Prior Belief Exists**

We have solved for the unique outcome when all presidents share the arbitrary stationary, symmetric prior belief over quality, \(\lambda_0 \in (0, 1)\). In this section we show that such a prior belief exists.

We initially assume that all previous Presidents held the same prior belief over quality, \(\lambda_0\). We then calculate an incoming President’s prior belief over quality and show that 1) it is independent of his type, \(\theta\) and 2) there exists at least one \(\lambda_0\) which is a fixed point so that the incoming President’s prior belief will be the same as that of his predecessors and so beliefs will be stationary.

If all presidents in previous periods used stationary strategies and had the same prior belief, \(\lambda_0\), then an incoming president’s belief about the director’s survival probability in previous periods will also be stationary. However, it could differ according to the president’s type. To capture this, we denote a president of type \(\theta\)'s belief that a given director survives a given period as \(\hat{\theta}\). We initially assume that all previous Presidents held the same prior belief over quality, \(\lambda_0\), and 2) there exists at least one \(\lambda_0\) which is a fixed point so that the incoming President’s prior belief will be the same as that of his predecessors and so beliefs will be stationary.

If all presidents in previous periods used stationary strategies and had the same prior belief, \(\lambda_0\), then an incoming president’s belief about the director’s survival probability in previous periods will also be stationary. However, it could differ according to the president’s type. To capture this, we denote a president of type \(\theta\)'s belief that a given director survives a given period as \(\hat{\theta}\). Similarly we denote his belief as to the survival probability of a director conditional on him being a good type by \(\bar{\hat{\theta}}\).

Because time is doubly infinite, at any time \(t \in \{0, 1, 2, \ldots\}\) we can calculate the probability that the incumbent director is good from the point of view of a president of type \(\theta\) as:

\[
\lambda_0(\theta) = \frac{(1 - \bar{\hat{\theta}})\gamma + (1 - \bar{\hat{\theta}})\bar{\hat{\theta}}\gamma + (1 - \bar{\hat{\theta}})(\bar{\hat{\theta}}\gamma)^2 + (1 - \bar{\hat{\theta}})(\bar{\hat{\theta}}\gamma)^3 + \ldots}{(1 - \bar{\hat{\theta}})^\gamma}
\]

The probability the current director is good is equal to the probability a new director was appointed last period who was good (which occurs with probability \((1 - \bar{\hat{\theta}})\gamma\)) plus the probability that a good director was appointed two periods ago and survived until the present day (which occurs with probability \((1 - \bar{\hat{\theta}})\gamma\bar{\hat{\theta}}\)) and so on for an infinite number of periods.

We first show that this belief does not depend on the incoming President’s type. This will be the case if \(\bar{\hat{\theta}}\) and \(\bar{\hat{\theta}}\) are both independent of \(\theta\). To economise on notation in the derivation that follows we write \(\lambda^\theta(r, F, \mu)\) as \(\lambda^\theta(r, F)\) and \(\sigma_\theta(r, \mu)\) as \(\sigma_\theta(r)\) and denote the probability that \(\mu\) takes a given value by \(p(\mu)\).

A president of type \(\theta\) believes that a good director survives with probability:

\[
\bar{\hat{\theta}}_\theta = \sum_{\mu \in \{\mu_L, \mu_H\}} p(\mu) \int \left( (1 - \mu) \left[ \theta G \left( \lambda^\theta(\mu, H, 0) \right) + (1 - \theta) G \left( \lambda^\theta(\mu, L, 0) \right) \right] + \mu \left[ \theta G \left( \lambda^\theta(\mu, H, H) \right) + (1 - \theta) G \left( \lambda^\theta(\mu, L, L) \right) \right] \right) J(\bar{\hat{\theta}}) d\bar{\hat{\theta}}
\]

\[
= \sum_{\mu \in \{\mu_L, \mu_H\}} p(\mu) \int \left( (1 - \mu) \left[ \theta G \left( \lambda^\theta(\mu, L, 0) \right) + \theta \left( G \left( \lambda^\theta(\mu, H, 0) \right) - G \left( \lambda^\theta(\mu, L, 0) \right) \right) \right] + \mu \left[ G \left( \lambda^\theta(\mu, L, 0) \right) + \theta \left( G \left( \lambda^\theta(\mu, H, H) \right) - G \left( \lambda^\theta(\mu, L, L) \right) \right) \right] \right) J(\bar{\hat{\theta}}) d\bar{\hat{\theta}}
\]
Examining the first term multiplied by \( \theta \) inside the summation we see that, because president’s types are chosen independently it can be rewritten as:

\[
\theta \int_{0}^{1} G \left( \lambda^{\hat{\mu}}(H, \emptyset) \right) - G \left( \lambda^{\hat{\mu}}(L, \emptyset) \right) f(\tilde{\theta}) d\tilde{\theta}
\]  

(13)

In addition, we can use our observation from step 1 of the proof that, when all presidents have a common prior then, for any type \( \theta > .5 \) there is a \( \theta' < .5 \) who acts in an equal and opposite manner. This, combined with the assumption that \( f \) is symmetric around \( .5 \) allows us to conclude that, for each value of \( \mu \):

\[
\int_{0}^{\frac{5}{2}} G(\lambda^{\hat{\mu}}(L, \emptyset) f(\tilde{\theta}) d\tilde{\theta} = \int_{0}^{\frac{5}{2}} G(\lambda^{\hat{\mu}}(H, \emptyset) f(\tilde{\theta}) d\tilde{\theta}
\]

\[
\int_{0}^{\frac{5}{2}} G(\lambda^{\hat{\mu}}(L, \emptyset) f(\tilde{\theta}) d\tilde{\theta} = \int_{0}^{\frac{5}{2}} G(\lambda^{\hat{\mu}}(H, \emptyset) f(\tilde{\theta}) d\tilde{\theta}
\]

Substituting this into (13) shows that the term is equal to zero. An identical argument can be shown for the second term multiplied by \( \theta \) and so, given that all other types have a common prior belief over quality, \( \bar{\pi}_{\theta}^0 = \bar{\pi}_{\theta}^0 \forall \theta \).

Showing the same result for the survival probability of a normal type is slightly more involved. We first calculate the probabilities that a president of type \( \theta \) assigns to each reporting history occurring, given that his predecessor was of type \( \theta' \) and that the probability of external verification was given by \( \mu \):

\[
P^{\mu}_{L_0} = (1 - \mu) \left( \left[ \theta \pi + (1 - \theta)(1 - \pi) \right] \sigma^{\hat{\mu}}_L(L) + [(1 - \theta) \pi + \theta(1 - \pi)] \sigma^{\hat{\mu}}_L(L) \right)
\]

\[
P^{\mu}_{H_0} = (1 - \mu) \left( \left[ \theta \pi + (1 - \theta)(1 - \pi) \right] \sigma^{\hat{\mu}}_H(H) + [(1 - \theta) \pi + \theta(1 - \pi)] \sigma^{\hat{\mu}}_H(H) \right)
\]

\[
P^{\mu}_{L_L} = \mu (1 - \theta) \left( \pi \sigma^{\hat{\mu}}_L(L) + (1 - \pi) \sigma^{\hat{\mu}}_H(L) \right)
\]

\[
P^{\mu}_{H_L} = \mu (1 - \theta) \left( \pi \sigma^{\hat{\mu}}_L(H) + (1 - \pi) \sigma^{\hat{\mu}}_H(H) \right)
\]

\[
P^{\mu}_{H_H} = \mu \theta \left( \pi \sigma^{\hat{\mu}}_H(H) + (1 - \pi) \sigma^{\hat{\mu}}_L(L) \right)
\]

\[
P^{\mu}_{L_H} = \mu \theta \left( \pi \sigma^{\hat{\mu}}_H(L) + (1 - \pi) \sigma^{\hat{\mu}}_L(L) \right)
\]

Then a president of type \( \theta \) believes that the unconditional probability a normal type survives is:

\[
\bar{p}_n = \Sigma_{\mu \in (L, H, L)} p(\mu) \left[ (1 - \mu) \int \left[ \theta \pi + (1 - \theta)(1 - \pi) \right] \sigma^{\hat{\mu}}_H(H) G(\lambda^{\hat{\mu}}(H, \emptyset)) \right]
\]

\[
+ [(1 - \theta) \pi + \theta(1 - \pi)] \sigma^{\hat{\mu}}_L(L) G(\lambda^{\hat{\mu}}(L, \emptyset)) \right]
\]

\[
+ \left[ \left( \theta \pi + (1 - \theta)(1 - \pi) \right] \sigma^{\hat{\mu}}_L(L) G(\lambda^{\hat{\mu}}(L, \emptyset)) \right]
\]

\[
+ \left[ \left( \theta \pi + (1 - \theta)(1 - \pi) \right] \sigma^{\hat{\mu}}_H(H) G(\lambda^{\hat{\mu}}(H, H)) \right]
\]

\[
+ \left[ \left( \theta \pi + (1 - \theta)(1 - \pi) \right] \sigma^{\hat{\mu}}_L(L) G(0) + \left( \pi \sigma^{\hat{\mu}}_H(H) + (1 - \pi) \sigma^{\hat{\mu}}_H(H) \right) G(0) \right]
\]

\[
f(\tilde{\theta}) d\tilde{\theta}
\]

31
Rearranging gives:

\[ \tilde{p}_n = \sum_{\mu \in \{L, H\}} \tilde{p}(\mu) \frac{(1 - \mu)}{(1 - \tilde{p}(\lambda_0))} \left[ (1 - \mu) \int \left[ \left[ \pi \sigma_{L}^{\mu}(H) + (1 - \pi)\sigma_{H}^{\mu}(H) \right] G \left( \lambda^{\mu}(H, \theta) \right) \right. \right. \\
+ \left. \left. \left[ \pi \sigma_{L}^{\mu}(L) + (1 - \pi)\sigma_{H}^{\mu}(L) \right] G \left( \lambda^{\mu}(L, \theta) \right) \right] \right] f(\tilde{\theta}) d\tilde{\theta} \]

A similar symmetry argument, this time also using the fact that, for \( \theta \geq 0 \) there exists a \( \theta' \leq 0 \) such that \( \sigma_{H}^{\mu}(H) = \sigma_{L}^{\mu}(L), \sigma_{H}^{\mu}(L) = \sigma_{L}^{\mu}(H) \) and vice versa, shows that all the terms multiplied by \( \theta \) are equal to zero.

Given that all presidents in previous periods held the stationary, prior belief \( \lambda_0 \), all types of president in the current period will have prior belief given by:

\[ \lambda(\lambda_0) = \frac{(1 - \tilde{p}(\lambda_0))}{(1 - \tilde{p}(\lambda_0))} \gamma \]  

(14)

Where the President’s belief about the director’s survival probability are themselves a function of the past presidents’ prior, \( \lambda_0 \). Any \( \lambda_0' \in (0, 1) \) which is a fixed point of (14) will support an SSMPE of the form described in 1. To show that at least one such \( \lambda_0' \) exists we note that 1) \( \lambda(\cdot) \) is continuous in \( \lambda_0 \) (as reappointment probabilities are continous in the President’s prior belief over quality), 2) When \( \lambda_0 = 0 \) all types will be considered bad no matter what they do so that \( \tilde{p}_n = \bar{p}_n \) and \( \lambda(0) = \gamma \) 3) When \( \lambda_0 = 1 \) all types will be considered good no matter what they do so that \( \tilde{p}_n = \hat{p}_n \) and so \( \lambda(1) = 1 \).

Because the function \( \lambda(\lambda) \) is continuous, is initially above the 45 degree line and is eventually below the 45 degree line. It follows that there will exist at least one \( \lambda^* \) where the function crosses the 45 degree line and \( \lambda(\lambda^*) = \lambda^* \). This implies that at least one symmetric, steady state prior will exist that supports the above behaviour as an equilibrium. Hence a SSMPE of the form in Proposition 1 exists.

Because the analysis in step 1 of the proof pinned down the unique outcome in any informative SSMPE with arbitrary prior \( \lambda_0 \) it follows that the equilibrium is unique up to the choice of \( \lambda_0 \). Therefore all SSMPEs in which reports ever influence presidents’ actions (so that the result in lemma 1 applies), will take the form described in 1 and equilibrium bias will consist of at most misreporting of signals that contradict the President’s prior beliefs.

**Step 4: Comparative Statics**

To see the comparative statics on \( \theta \) we begin with the hawkish case. In this case the equilibrium condition when \( \sigma_{H}(\theta, \mu) > 0 \) is that \( \Delta_{L}^{\theta}(H, \mu) = 0 \). Because the director’s flow payoff is always strictly positive it follows that \( V > 0 \) and so the equilibrium condition in equation (10) collapses to:

\[ 0 = (1 - \mu) \left[ G(\lambda^{\theta}(H, \theta, \mu)) - G(\lambda^{\theta}(L, \theta, \mu)) \right] + \mu \left[ (1 - \pi)G(\lambda^{\theta}(H, H, \mu)) + \pi G(0) \right] - \left( \pi G(\lambda^{\theta}(L, L, \mu)) + (1 - \pi)G(0) \right) \]
Holding the probability of garbling \(L\) signals fixed, this expression is strictly increasing in \(\theta\). In addition, it is strictly decreasing in the probability that the director misreports \(L\) signals, \(\sigma_h(\theta, \mu)\). It therefore follows that, after an increase in \(\theta\), \(\sigma_h(\theta, \mu)\) must increase to preserve the equilibrium condition. Hence \(\sigma^*_h(\theta, \mu)\) is increasing in \(\theta\).

Similarly, in the case where the president is a dove, the equilibrium condition collapses to:

\[
0 = (1 - \mu) \left[ G(\lambda^\theta(H, \emptyset, \mu)) - G(\lambda^\theta(H, \emptyset, \mu)) \right]
+ \mu \left[ \left( (1 - \pi)G(\lambda^\theta(L, L, \mu)) + \pi G(0) \right) - \left( (1 - \pi)G(\lambda^\theta(H, H, \mu)) + (1 - \pi)G(0) \right) \right]
\]

For fixed \(\sigma_d(\theta, \mu)\) this is strictly decreasing in \(\theta\). In addition it is strictly decreasing in the probability of garbling \(H\) signals, \(\sigma_d(\theta, \mu)\). It follows that, after an increase in \(\theta\) \(\sigma_d(\theta, \mu)\) must fall to preserve the equilibrium condition. Hence, \(\sigma^*_d(\theta, \mu)\) is decreasing in \(\theta\) for all \(\mu\).

**Proof of Proposition 2**

To show that the probability of misreporting by normal directors is greater when \(\mu = \mu_1\) than when \(\mu = \mu_h\) we focus on the case of a hawkish president. Suppose that bias was actually greater when \(\mu = \mu_h\) then there are two cases. Either \(\sigma^*_h(\theta, \mu_1) = 0\) and \(\sigma^*_h(\theta, \mu_h) > 0\) or \(\sigma^*_h(\theta, \mu_1) > 0\) and \(\sigma^*_h(\theta, \mu_h) > 0\). Beginning with the first case, it implies that, at \(\sigma_h(\theta, \mu_1) = 0:\)

\[
0 \geq (1 - \mu_1) \left[ G(\lambda^\theta(H, \emptyset, \mu_1)) - G(\lambda^\theta(L, \emptyset, \mu_1)) \right]
+ \mu_1 \left[ \left( (1 - \pi)G(\lambda^\theta(H, H, \mu_1)) + (1 - \pi)G(0) \right) - \left( (1 - \pi)G(\lambda^\theta(L, L, \mu_1)) + \pi G(0) \right) \right]
\]

But, that at \(\sigma_h(\theta, \mu_h) = 0:\)

\[
0 < (1 - \mu_h) \left[ G(\lambda^\theta(H, \emptyset, \mu_h)) - G(\lambda^\theta(L, \emptyset, \mu_h)) \right]
+ \mu_h \left[ \left( (1 - \pi)G(\lambda^\theta(H, H, \mu_h)) + (1 - \pi)G(0) \right) - \left( (1 - \pi)G(\lambda^\theta(L, L, \mu_h)) + \pi G(0) \right) \right]
\]

So that there is misreporting of \(L\) signals to a hawkish president of type \(\theta\) when \(\mu = \mu_h\) and truthful reporting when \(\mu = \mu_1\). Note that, because the director’s reputation depends on \(\mu\) only through its effect on the President’s belief about his reporting strategy, when we evaluate both expressions at the same probability of misreporting, the director’s reputation after each history is the same for both levels of \(\mu\). Hence the two expressions differ only in how much weight they put on the first term relative to the second.

From lemma 2 we know that, when \(\sigma_h(\theta, \mu) = 0\) and there is no garbling of signals, the terms multiplied by \((1 - \mu)\) are positive and the terms multiplied by \(\mu\) are negative. Because expression (16) puts more weight on the negative term than (15) it must be that, if (15) is negative, so is (16). This implies that \(\sigma^*_h(\theta, \mu_h) = 0\), a contradiction.

Similarly in the second case where \(\sigma^*_h(\theta, \mu_h) > \sigma^*_h(\theta, \mu_1) > 0\), it must be that there is \(\sigma^*_h(\theta, \mu_1)\) such that, at this level of misreporting, there is no incentive to misreport any further when \(\mu = \mu_1:\)

\[
0 = (1 - \mu_1) \left[ G(\lambda^\theta(H, \emptyset, \mu_1)) - G(\lambda^\theta(L, \emptyset, \mu_1)) \right]
+ \mu_1 \left[ \left( (1 - \pi)G(\lambda^\theta(H, H, \mu_1)) + (1 - \pi)G(0) \right) - \left( (1 - \pi)G(\lambda^\theta(L, L, \mu_1)) + \pi G(0) \right) \right]
\]

but, when \(\mu = \mu_h\), there is such an incentive:

\[
0 < (1 - \mu_h) \left[ G(\lambda^\theta(H, \emptyset, \mu_h)) - G(\lambda^\theta(L, \emptyset, \mu_h)) \right]
+ \mu_h \left[ \left( (1 - \pi)G(\lambda^\theta(H, H, \mu_h)) + (1 - \pi)G(0) \right) - \left( (1 - \pi)G(\lambda^\theta(L, L, \mu_h)) + \pi G(0) \right) \right]
\]

Once again, the two expressions are evaluated at the same level of bias (this time \(\sigma^*_h(\theta, \mu_1)\)) and so they differ only in how much weight is put on the first term relative to the second. From Lemma 2 we know that the second term must be negative because it is negative at zero bias and strictly decreasing in the probability.
of misreporting. In addition, the equilibrium condition in (18) implies that the first term must be positive at $\sigma_1^*(\theta, \mu_t)$.

Because $\mu_h > \mu_l$ it must therefore be that the right hand side of (18) is strictly less than the right hand side of (17). This implies that $\sigma_1^*(\theta, \mu_h) < \sigma_1^*(\theta, \mu_l)$, a contradiction. For the dovish case a similar examination of the equilibrium condition on $\Delta_L^0(H, \mu)$ shows that $\sigma_2^*(\theta, \mu_h) \geq \sigma_2^*(\theta, \mu_l)$ and that this inequality is strict whenever $\sigma_2^*(\theta, \mu_l) > 0$.

Although we can show that there will be more bias in periods with $\mu = \mu_l$ than in periods with $\mu = \mu_h$, we cannot sign the comparative static of increasing either $\mu_l$ or $\mu_h$. This is because, when we increase the probability of the state being revealed, we also change the probability that good directors survive relative to normal directors. This in turn will affect the President’s prior belief on quality. Changes in $\lambda_0$ move the president’s posterior beliefs after all histories (except those where a normal director fails to match the state) in the same direction. As a result it is not possible to sign the effect on the difference in reappointment probability and it is this that determines equilibrium bias. As a result it is not possible to sign the effect on the equilibrium level of bias.

**Deriving Predictions 2 and 3**

To derive prediction 1 we calculate the change in the expected report error conditional on $\mu$ when we move from $\mu_l$ to $\mu_h$:

$$E(\text{error} \mid \mu_h) - E(\text{error} \mid \mu_l) = \kappa \left[ \int_{\frac{1}{2}}^{1} \sigma_1^*(\theta, \mu_h) - \sigma_1^*(\theta, \mu_l) f(\theta) d\theta - \int_{0}^{\frac{1}{2}} \sigma_2^*(\theta, \mu_h) - \sigma_2^*(\theta, \mu_l) f(\theta) d\theta \right]$$

Using the same symmetry result in the proof of proposition 1 we can write $\int_{\frac{1}{2}}^{1} \sigma_1^*(\theta, \mu) f(\theta) d\theta = \int_{0}^{\frac{1}{2}} \sigma_2^*(\theta, \mu) f(\theta) d\theta$ for $\mu \in \{\mu_l, \mu_h\}$ and so the term cancels to zero.

To derive Prediction 3 we note that there will be a negative interaction effect between $\theta$ and $\mu$ if the marginal effect of moving from $\theta$ to $\theta' > \theta$ on the report error is always larger when the probability of external verification is given by $\mu_l$ rather than $\mu_h$:

$$E(\text{error} \mid \theta', \mu_h) - E(\text{error} \mid \theta, \mu_h) < E(\text{error} \mid \theta', \mu_l) - E(\text{error} \mid \theta, \mu_l)$$

In the case where the President is a hawk this requires that:

$$\sigma_1^*(\theta', \mu_h) - \sigma_1^*(\theta, \mu_h) \leq \sigma_1^*(\theta', \mu_l) - \sigma_1^*(\theta, \mu_l) \quad (19)$$

While in the case of a dove it requires that:

$$-\sigma_2^*(\theta', \mu_h) + \sigma_2^*(\theta, \mu_h) \leq -\sigma_2^*(\theta', \mu_l) + \sigma_2^*(\theta, \mu_l)$$

$$\sigma_2^*(\theta, \mu_h) - \sigma_2^*(\theta', \mu_h) \leq \sigma_2^*(\theta, \mu_l) - \sigma_2^*(\theta', \mu_l) \quad (20)$$

In other words: if the marginal effect of making the president more extreme (making a hawk more hawkish or a dove more dovish) on the probability of misreporting is always larger when $\mu = \mu_l$ than when $\mu = \mu_h$ we will observe a negative interactive effect between $\theta$ and $\mu$ on the report error. Whether or not this will be the case is unclear because we cannot sign the relative size of the effect of changes in $\theta$ on the incentive to misreport.

**Appendix B: Extensions to The Baseline Model**

**Allowing for Good Directors to Lie**

In this section we demonstrate that the equilibrium in Proposition 1 is robust to allowing the good type to misreport his signal. This makes clear that the salient feature that makes a director “good” is the quality
of his signal rather than his honesty. This is an identical result to that in Gentzkow and Shapiro (2006).

Focussing on the case of a hawkish President we note that the normal director’s equilibrium strategy, \( \sigma_{\text{h}}^{*}(\theta, \mu) \), is determined so as to leave him indifferent between reporting \( H \) and \( L \) after an \( L \) signal by equating his probability of reappointment in each case:

\[
(1 - \mu)G(\lambda^\theta(H, \emptyset, \mu)) + \mu \left((1 - \pi)G(\lambda^\theta(H, H, \mu)) + \pi G(0)\right) = (1 - \mu)G(\lambda^\theta(L, \emptyset, \mu)) + \mu \left(\pi G(\lambda^\theta(L, L, \mu)) + (1 - \pi)G(0)\right)
\]

(21)

We now show that, given that the President’s beliefs are as in Proposition 1, a good director has a strict incentive to report \( L \) signals truthfully. The fact that the good director receives a perfect signal of the state of the world means that if he reports \( H \) after an \( L \) signal, he will fail to match the state for sure if it is revealed. This gives his reappointment probability if he reports \( H \) as:

\[
(1 - \mu)G(\lambda^\theta(H, \emptyset, \mu)) + \mu G(0)
\]

This is strictly smaller than the left hand side of (21). Similarly, if he reports \( L \) he knows that he will match the state correctly for sure if it is revealed and so his reappointment probability is given by:

\[
(1 - \mu)G(\lambda^\theta(L, \emptyset, \mu)) + \mu G(\lambda^\theta(L, L, \mu))
\]

This is strictly greater than the right hand side of (21). Hence, given the strategies and beliefs in Proposition 1, good directors have a strict incentive to report \( L \) signals correctly to hawkish Presidents and so the equilibrium is robust to allowing good directors to lie. If we relabel the states we can equivalently show that good directors have a strict incentive to report \( H \) signals correctly to dovish Presidents.

Ruling Out Uninformative Equilibria

While Proposition 1 characterizes reporting in all SSMPE in which reports influence presidential actions, there is still the potential that uninformative SSMPE could exist. Because reputation only has an instrumental role in our model, once equilibria are uninformative and Presidents do not condition their actions on the reports, there is no incentive for them to reappoint good directors and so the reputational mechanism we are interested in breaks down.

To ensure that reports are uninformative despite the presence of good types who continue to report the state perfectly, an uninformative equilibrium requires that normal types deliberately misreport their signals just enough to counteract the information being conveyed by the good types. However, as long as good types are sufficiently prevalent in the population, no amount of misreporting can counteract them. In this case all SSMPE are informative and so must take the form in Proposition 1.

**Proposition 3.** There exists \( \gamma^* = \frac{2\pi - 1}{2\pi} \in (0, 1) \) such that, if \( \gamma \geq \gamma^* \), all SSMPEs are informative and take the form described in Proposition 1.

**Proof.** A babbling equilibrium requires that Presidents’ beliefs about the state are unchanged by reports:

\[
\hat{\theta}(H, \theta, \mu) = \hat{\theta}(L, \theta, \mu) = \theta \forall \mu
\]

(22)

Which implies that their actions are also independent of reports:

\[
a^\theta(H, \mu) = a^\theta(L, \mu) \forall \theta, \mu
\]

This means that \( I(\theta, \mu) = 0 \) for all \( \theta \) and \( \mu \), so the President’s expected legacy payoff is independent of quality: \( v(\lambda_1) = v(\gamma) \). This implies that both good and normal directors are reappointed with equal probability: \( \bar{p}^\theta = \bar{p}^\alpha = G(0) \) and so a President’s prior belief over the quality of an incumbent director is the same as for a new director (\( \lambda_0 = \gamma \)).
Using the fact that $\lambda_0 = \gamma$ we can rewrite condition (22) as:

$$\frac{\gamma}{1 - \gamma} = (2\pi - 1)(\sigma^L_H(H, \mu) - \sigma^H_H(H, \mu))$$

This states that the normal type must report his signal incorrectly just often enough to mask the informative signals being sent by the good type. Clearly the right hand side is maximized at $(2\pi - 1)$ when $\sigma^H_H(H, \mu) = 1$ and $\sigma^L_L(H, \mu) = 0$, so a sufficient condition to rule out babbling equilibria is that:

$$\frac{\gamma}{1 - \gamma} > 2\pi - 1$$

Or that $\gamma > \frac{2\pi - 1}{\pi} \in (0, 1)$. Under this parameter condition, reports to Presidents of all types will be informative and so will shift their beliefs about the state. As in the proof to proposition 1 we can then always find a range of Presidents with $\theta$ close enough to a half who condition their action on the director’s report. This then implies that the President’s continuation value from reappointing the director is increasing in his posterior belief over quality and so reputation matters in equilibrium. This in turn implies that equilibrium behaviour in any SSMPE must take the form described in Proposition 1.

The proportion of good types in the population required to rule out uninformative equilibria is increasing in the quality of the normal type’s signal, $\pi$. The intuition for this result is that, in order to counteract the good type’s reports the normal type must fail to match the true state. The higher is the quality of the normal type’s signal the more effective is its misreporting. Ironically, making normal directors’ signals more informative can actually make equilibrium reporting less informative by making coordination on an uninformative equilibrium possible.

### Allowing for the Director to Care About the President’s Action

In this section we extend the results in Proposition 1 to allow for the director to care about the President’s action. Formally, we assume that the director, as well as receiving the ego rent, $A$, each period he is in office, also receives a payoff of one if the President’s action matches the state in a given period and zero otherwise. However, we assume that the director cares only about actions that occur during his time in office. Doing otherwise would open up the possibility of normal directors actually wanting themselves to be removed so as to make way for a good quality replacement.

All else being equal, the director would like the President to choose the action that corresponds to his signal. As long as the President conditions his action on the director’s report, the effect of this is to discourage misreporting as sending the wrong report makes it less likely that the President will take the correct action. In this case, a low value of $\mu$ is no longer sufficient for misreporting to occur in equilibrium because, even if misreporting increases the director’s probability of reappointment, this benefit may be outweighed by the increased likelihood of having the President take the incorrect action. The effect of this trade-off is to introduce a second parameter condition alongside the one on $\mu$: the director’s ego rent must also be sufficiently large so that the reputational effect dominates.

When the President’s beliefs are such that he does not condition his action on the director’s report, allowing for policy preferences on the part of the agency director has no effect on his reporting decision. This follows because, when the President ignores reports, misreporting will have no effect on the probability that the correct action is taken. Because the Presidents who ignore reports in equilibrium are those with the strongest prior beliefs, allowing for directors to care about policy introduces an additional mechanism by which increasing the strength of the President’s prior beliefs can increase the level of bias: once a President becomes sufficiently ideological that he ceases to condition his actions on his advisor’s reports, one of the constraints that might prevent pandering dissapears. Intuitively, once it is clear to the director that he will not be listened to anyway, he may as well curry favour by pandering to the President’s prior beliefs.

In summary, the results in our baseline model are robust to allowing for the directors to care about the President’s actions. If ego rents are small so that the desire to match the state dominates, directors will still pander to Presidents whose beliefs are extreme enough that they ignore director’s reports when choosing
their actions. If ego rents are large, then reputational incentives will dominate and directors will pander to Presidents even though they realise that doing so makes it more likely that the President will take the incorrect action.

Proposition 4 summarises these results:

**Proposition 4.** The game has a SSMPE with the following properties:

1. If the President is a hawk then normal agency directors always report $H$ signals truthfully, but distort $L$ signals with the unique probability $\sigma_h^* (\theta, \mu) \in [0, 1)$. In the range where $\sigma_h^* (\theta, \mu)$ is strictly positive it is strictly increasing in $\theta$.

2. If the President is a dove then normal agency directors always report $L$ signals truthfully, but distorts $H$ signals with the unique probability $\sigma_d^* (\theta, \mu) \in [0, 1)$. In the range where $\sigma_d^* (\theta, \mu)$ is strictly positive it is strictly decreasing in $\theta$.

3. For each type of President there exists $\mu_0^\#$ such that if $\mu_h, \mu_l \geq \mu_0^\#$ there is never a reputational gain from misreporting to a president of type $\theta$. If $\mu_l, \mu_h < \mu_0^\#$ then there is a reputational gain to misreporting. Finally, if $\mu_h \geq \mu_0^\# > \mu_l$, there is a reputational gain from misreporting when $\mu = \mu_l$, but not when $\mu = \mu_h$.

4. If the President is either an extreme hawk ($\theta \geq \lambda + (1 - \lambda)\pi$) or an extreme dove ($\theta < (1 - \lambda)(1 - \pi)$) there will be positive misreporting whenever there is a reputational gain from misreporting.

5. Otherwise there exists $A_0^\# (\mu)$ such that there will be positive misreporting if both $A \geq A_0^\# (\mu)$ and $\mu < \mu_0^\#$.

6. The probability of misreporting is higher when the probability of external verification is low. Formally, $\sigma_h^* (\theta, \mu_l) \leq \sigma_h^* (\theta, \mu_l)$ and $\sigma_d^* (\theta, \mu_h) \leq \sigma_d^* (\theta, \mu_l)$. These inequalities hold strictly when respectively $\sigma_h^* (\theta, \mu_l) > 0$ and $\sigma_d^* (\theta, \mu_l) > 0$.

*Proof.* To avoid repetition we focus on the case of a hawkish President. The equivalent analysis for when the President is a Dove is almost identical.

If $I(\theta, \mu) = 0$ and the President does not condition his action on the director’s report then the fact that the director cares about policy will not have any effect on the director’s incentive to report truthfully. If instead $I(\theta, \mu) = 1$, then the director’s expected policy payoff is given by $\pi$ if he reports his signal truthfully and $(1 - \pi)$ otherwise. We can therefore write the net gain in terms of the director’s policy payoff to reporting $H$ after respectively receiving the signal $H$ and $L$ as:

$$
\Lambda^0_H (H, \theta, \mu) = (2\pi - 1)I(\theta, \mu) \geq 0
$$

$$
\Lambda^0_L (H, \theta, \mu) = (1 - 2\pi)I(\theta, \mu) \leq 0
$$

Similarly, we write the net gain in terms of reappointment probability from reporting $H$ to a hawkish president of type $\theta$ after receiving signal $s$ in the case of feedback and no feedback as:

$$
\Lambda^0_f (H, \theta, \mu) = \hat{\theta}_D (s) \left( G(\lambda^0 (H, \theta, \mu)) - G(0) \right) + (1 - \hat{\theta}_D (s)) \left( G(0) - G(\lambda^0 (L, \theta, \mu)) \right)
$$

$$
\Lambda^0_s (H, \theta, \mu) = G(\lambda^0 (H, \theta, \mu)) - G(\lambda^0 (L, \theta, \mu))
$$

With this new notation we can write the overall gain to reporting $H$ rather than $L$ to a hawkish president of type $\theta$ after receiving respectively an $H$ and $L$ signal as:

$$
\Delta^0_H (H, \mu) = \Lambda^0_H (H, \theta, \mu) + \left[ (1 - \mu)\Lambda^0_{nf} (H, \theta, \mu) + \mu\Lambda^0_H (H, \theta, \mu) \right] \bar{V}
$$

$$
\Delta^0_L (H, \mu) = \Lambda^0_L (H, \theta, \mu) + \left[ (1 - \mu)\Lambda^0_{nf} (H, \theta, \mu) + \mu\Lambda^0_L (H, \theta, \mu) \right] \bar{V}
$$

The fact that $\Lambda^0_H (H, \theta, \mu) \geq \Lambda^0_L (H, \theta, \mu)$ means that, as in the baseline case, $\Delta^0_H (H, \mu) \geq \Delta^0_L (H, \mu)$ so that the gain to reporting $H$ rather than $L$ to a hawkish president is larger after an $H$ signal. This means that an
identical analysis to that in the proof of Proposition 1 can be used to argue that, when faced with a hawkish President, normal directors will never misreport $H$ signals and will at most garble $L$ signals.

Examining equation (24) we see that the second term is identical to equation (10) and so is decreasing in $\sigma_L^0(H,\mu)$ as before. However, unlike before, this is insufficient to prove the existence of a unique probability of misreporting. This is because there could be two points at which $\Delta_L^0(H,\mu) = 0$: one in which the President does not condition his action on the director’s report and the equilibrium condition is given by:

$$\left[(1-\mu)\Lambda^{nL}_L(H,\mu) + \mu\Lambda^{nf}_L(H,\mu)\right]V = 0 \quad (25)$$

and another where he does condition his action on the director’s report and the equilibrium condition is that:

$$(1-2\pi) + \left[(1-\mu)\Lambda^{nL}_L(H,\mu) + \mu\Lambda^{nf}_L(H,\mu)\right]V = 0 \quad (26)$$

To prove that this is not the case, suppose that we have two reporting strategies for which $\Delta_L^0(H,\mu) = 0$: $\sigma$ for when the President does not conditions his action on the report and $\sigma'$ for when he does. Then, from the equilibrium conditions, $\sigma'$ solves $(1-2\pi) + \left[(1-\mu)\Lambda^{nL}_L(H,\mu) + \mu\Lambda^{nf}_L(H,\mu)\right]V = 0$ and $\sigma$ solves $\left[(1-\mu)\Lambda^{nL}_L(H,\mu) + \mu\Lambda^{nf}_L(H,\mu)\right]V = 0$. Because $1 - 2\pi < 0$ it must be that, evaluated at $\sigma'$, $\left[(1-\mu)\Lambda^{nL}_L(H,\mu) + \mu\Lambda^{nf}_L(H,\mu)\right]V > 0$. Because this term is decreasing in $\sigma_L(H,\mu)$ it must be that $\sigma > \sigma'$. However, this implies that, under the strategy $\sigma$, good types are relatively more likely to report $L$ than under strategy $\sigma'$. This in turn implies that a hawkish president’s posterior belief that the state is $H$ after observing signal $L$ is lower when he believes the equilibrium strategy is $\sigma$ than when he believes it is $\sigma'$. However, this implies that, if he conditions his action on the report in an equilibrium with strategy $\sigma'$, he must also do so when the equilibrium strategy is $\sigma$. But, this contradicts our assumption that the President did not condition his action on the report when the equilibrium strategy was $\sigma$.

Given that the equilibrium is informative and that Presidents hold the prior belief over quality $\lambda_0$ there is a unique probability of misreporting $L$ signals to a president of type $\theta$ when the probability of external verification is $\mu$: $\sigma_H^0(\theta,\mu) \in [0,1)$. To find the conditions under which there is a non-zero probability of misreporting we first find the conditions under which $(1-\mu)\Lambda^{nL}_L(H,\mu) + \mu\Lambda^{nf}_L(H,\mu)$ is strictly positive when evaluated at $\sigma_L(H,\mu) = 0$. From Lemma 2 the first term is positive and the second term is negative when reporting is truthful. There is therefore a unique $\mu_0^*$ that ensures there is a reputational benefit from lying if $\mu < \mu_0^*$. Furthermore, this threshold value of $\mu$ is identical to that in the baseline model.

However, even when the condition on $\mu$ is satisfied, $\sigma_H^0(H,\mu) = 0$ may still be the equilibrium outcome. This will be true if, evaluated at $\sigma_H^0(H,\mu) = 0$, the reputational gain from reporting $H$ after an $L$ signal is outweighed by the cost of sending the wrong signal to the President and having him potentially take the wrong action:

$$\left[(1-\mu)\Lambda^{nL}_L(H,\theta,\mu) + \mu\Lambda^{nf}_L(H,\theta,\mu)\right]V < -\Delta_L^0(H,\theta,\mu) \quad (27)$$

The right hand side of (27) will be equal to zero (and the condition on $\mu$ will be sufficient to ensure a positive probability of misreporting) if, given that he believes normal directors are reporting truthfully, the President ignores intelligence reports when making his action decision. In the case of a hawkish President this will be true if $\theta(L,\theta,\mu) > .5$ when evaluated at $\sigma_H^0(H,\mu) = 0$. Rearranging this inequality gives the condition $\theta \geq \lambda_0 + (1-\lambda_0)\pi$.23 If this condition holds then their will be a positive probability of misreporting as long as the condition on $\mu$ is satisfied.

If instead, $\theta \in ([.5,\lambda_0 + (1-\lambda_0)\pi]$ so that the President will condition his action on the director’s report, then, whether or not there is misreporting of $L$ signals depends on the relative importance of reputational concerns. This is captured by the size of the continuation value from reappointment, $\bar{V}$: when the condition on $\mu$ is satisfied the left hand side of (27) is positive and strictly increasing in $\bar{V}$. Hence, as long as $\bar{V}$ is sufficiently large there will be bias whenever the condition on $\mu$ holds. Because $\bar{V}$ is itself continuous and

23Similarly rearranging the condition that $\theta(L,\theta,\mu) < .5$ for a dovesh President gives the condition that $\theta < (1-\lambda)(1-\pi)$
strictly increasing in the ego rent, $A$, there will exist a unique threshold $A^*(\theta, \mu)$ such that if $A > A^*(\theta, \mu)$ and $\mu < \mu^*_h$ there will be a positive probability of misreporting $L$ signals to hawkish presidents.

To show that the level of misreporting to a hawkish president is increasing in $\theta$ we note that, holding the probability of misreporting $L$ signals fixed, $(1 - \mu)\Lambda^F_L(H, \theta, \mu) + \mu\Lambda^f_L(H, \theta, \mu)$ is strictly increasing in $\theta$. In addition, $\theta$ has no effect on $\Lambda^H$ until it crosses the threshold $\lambda_0 + (1 - \lambda_0)\pi$, at which point it jumps discretely from $1 - 2\pi < 0$ to 0. Hence $\Lambda^H_L(H, \mu)$ is strictly increasing in $\theta$ and so, as before, $\sigma_h^*(\theta, \mu)$ must be increasing in $\theta$.

To show that bias is greater when $\mu = \mu_l$ than when $\mu = \mu_h$. Suppose that bias was actually greater when $\mu = \mu_h$ then there are two cases. Either $\sigma_h^*(\theta, \mu) = 0$ and $\sigma_h^*(\theta, \mu) > 0$ or $\sigma_h^*(\theta, \mu) > 0$ and $\sigma_h^*(\theta, \mu) > 0$. Beginning with the first case it implies that, at $\sigma_h(\theta, \mu_l) = 0$:

\[
0 \geq \Lambda^H_L(H, \theta, \mu) + \left[ (1 - \mu_l)\Lambda^F_L(H, \theta, \mu) + \mu_l\Lambda^f_L(H, \theta, \mu) \right] \bar{V}
\]

So that there is no net gain to misreporting $L$ signals when $\mu = \mu_l$. But, that at $\sigma_h(\theta, \mu_l) = 0$:

\[
0 < \Lambda^H_L(H, \theta, \mu) + \left[ (1 - \mu_h)\Lambda^F_L(H, \theta, \mu) + \mu_h\Lambda^f_L(H, \theta, \mu) \right] \bar{V}
\]

So that there is misreporting of signals when $\mu = \mu_l$. As we are evaluating the incentive to lie when the President believes that all directors are reporting truthfully the three terms $\Lambda^H$, $\Lambda^F$ and $\Lambda^f$ will be the same in each case. The right hand side of the two expressions differ only in that expression (28) puts more weight on $\Lambda^f$ than $\Lambda^F$. As in the baseline case we know from Lemma 2 that $\Lambda^F$ is positive and $\Lambda^f$ is negative and so the right hand side of 29 must be less than the right hand side of 28, a contradiction as it implies that $\sigma_h^*(\theta, \mu_h) = 0$.

In the second case where $\sigma_h^*(\theta, \mu_h) > \sigma_h^*(\theta, \mu_l) > 0$ it must be that there is $\sigma_h^*(\theta, \mu_l)$ such that, at this level of misreporting:

\[
0 = \Lambda^H_L(H, \theta, \mu) + \left[ (1 - \mu_l)\Lambda^F_L(H, \theta, \mu) + \mu_l\Lambda^f_L(H, \theta, \mu) \right] \bar{V}
\]

But, at this same level of reporting:

\[
0 < \Lambda^H_L(H, \theta, \mu) + \left[ (1 - \mu_h)\Lambda^F_L(H, \theta, \mu) + \mu_h\Lambda^f_L(H, \theta, \mu) \right] \bar{V}
\]

Again the terms $\Lambda^H$, $\Lambda^F$ and $\Lambda^f$ will be the same in the two cases as they are evaluated at the same level of bias. This implies a contradiction as when $\mu_h > \mu_l$ the right hand side of (31) is strictly less than the right hand side of (30) and so $\sigma_h^*(\theta, \mu_h) < \sigma_h^*(\theta, \mu_l)$. 

\[
\square
\]

Appendix C: The Textual Data

In the main text we explain how our text-based measure is constructed conditional on having the word frequencies from the liberal, conservative and presidential texts: \( \{f^L_i\}_{i=1}^N, \{f^C_i\}_{i=1}^N \), and \( \{f^P_i\}_{i=1}^N \). In this section we describe in more detail how we obtained the documents needed to calculate these frequencies.

The Congressional Data

We first obtained DW nominate scores for each congressman from the 82nd through 100th US Congress from the voteview project. We then constructed the congressional reference documents as follows:

1. We used a Perl script to systematically download the entire Congressional record for the years 1952 to 1990. This corresponds to the 82nd through 100th US Congress.

2. We converted the PDF files into plain text format. Because the congressional record is displayed in columns which were not recognised by the conversion software we wrote another script that parsed each day’s record into a single block of text. In addition, this script deleted extraneous material such as roll call votes, the representatives’ extensions of remarks and Senate proceedings.
3. We then used another script to separate out the individual debates. To ensure that we only used material related to defence and national security matters we deleted all debates that did not contain any of the words: “war”, “warfare”, “Soviet”, “weaponry”, “missile”, “bomb”, “bomber”, “ICBM”, “ICBMs”, “MBRM”, “MRBMs”, “IRBM”, “IRBMs”, “SLBM”, “SLBMs”, “naval”, “army”, “missile”, “MIRV”, “MIRVs”, “megaton”, “megatons” or “warhead”.

4. We then identified the points at which different congressmen started speaking. This allowed us to separate out the words spoken on defence-related matters by each congressman during each day’s record. By looping the script over all days during a given congress and combining the words spoken we were able to construct text files for each congressman that contained all of the words said by them during all defence related debates during that congress.

5. Once these text files were constructed we then identified the most liberal and conservative deciles of congressmen and pooled all of their speeches so that we had two reference texts for each congress.

6. Finally we deleted the congressmen’s names, the names of their states, all 2 and 3 letter words, each of a list of around 200 “stop words” considered by linguists to be of zero informational value and a set of common legislative phrases (e.g. I yield to the gentleman..., ) that might otherwise bias the Wordscore estimates.24

The Presidential Data

To identify speeches made by each President before he received security clearance, the first step was to identify exactly when and to what extent each Presidential candidate was briefed by the intelligence services. This was possible thanks to the release by the CIA’s center for the study of intelligence of the monograph “Getting to know the President: CIA Briefings of Presidential Candidates” which relates in some detail the first interactions between the CIA and its predecessors with each of the Presidents in our sample (Helgerson (1993)).

The first column of Table 1 summarises the details of the first interactions between the President and the intelligence services, while the page references refer to the page in “Getting to know the president” where this information can be found. The second column then describes the selection rule used to determine which speeches contributed to our text-based measure.

In the case of Eisenhower, he himself acknowledged that his security briefings were “kept very general for political reasons” and no sensitive information was conveyed. Similarly Ronald Reagan’s one briefing prior to winning the 1980 general election was described as “a circus” in which no useful information was conveyed. Hence in both cases we feel justified in using defence related speeches from the beginning of the campaign until their subsequent election. In contrast, Jimmy Carter actually requested to be briefed before he had even won the Democratic nomination. Thankfully for our identification strategy the CIA declined his request, but commenced detailed briefings as soon as he was nominated. For this reason we restrict our analysis to focus only on speeches he made during the primary campaign.

In the cases of Nixon and Kennedy we base our pre-presidential hawkishness scores on speeches made during the 1960 election campaign. in the case of President Nixon this is not ideal as he did not take office until 1968. The reason we take this route is primarily one of expediency: the iconic nature of the 1960 campaign means that there is a vast body of campaign speeches in electronic format, whereas the same is not true for the 1968 election. For both Nixon and Kennedy we do not use any campaign speeches made after the 17th of September as this is the date the CIA started to give detailed briefings to the two candidates.

The biggest challenge is choosing which speeches to use for Presidents Johnson and Ford who assumed office directly by virtue of being Vice President and without the need for an election campaign. While the reports in our sample were Top Secret they would have had clearance to view them through their position on the National Security Council. Thankfully both Presidents served in the legislative branch prior to becoming

24It should be noted that deleting stop words is not strictly necessary as for these words, it should be the case that \( f^L_i \approx f^C_i \). Hence the Wordscores algorithm endogenously identifies these words as having little informational content. However, deleting these words reduced the computation time.
Table 1: Details of First Interactions and Sources Used

<table>
<thead>
<tr>
<th>President</th>
<th>First Interactions</th>
<th>Speeches Used</th>
<th>Congress</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eisenhower</td>
<td>No specific briefings prior to presidency (p. 29)</td>
<td>Defence-related speeches from 3/01/1951 until election</td>
<td>82nd</td>
</tr>
<tr>
<td>Kennedy</td>
<td>2-hour briefing on 23rd July 1960 frequent briefings from 17th September (p. 50)</td>
<td>Defence-related campaign speeches prior to 17th September.</td>
<td>86th</td>
</tr>
<tr>
<td>Johnson</td>
<td>Abrupt start to briefings following Kennedy assassination (p. 69)</td>
<td>Defence-related senate speeches from 03/01/1960 until 17/09/1960</td>
<td>86th</td>
</tr>
<tr>
<td>Nixon</td>
<td>Detailed briefings from outset of 1968 campaign (p. 79)</td>
<td>Defence-related 1960 campaign speeches prior to 17th September</td>
<td>86th</td>
</tr>
<tr>
<td>Ford</td>
<td>Detailed briefings while Vice-President once Watergate threatened Nixon’s presidency (p. 94)</td>
<td>Defence-related House speeches while representative during 92nd Congress</td>
<td>92nd</td>
</tr>
<tr>
<td>Carter</td>
<td>In depth briefings following nomination (p. 105)</td>
<td>Defence-related campaign speeches from announcement until nomination (including acceptance speech).</td>
<td>94th</td>
</tr>
<tr>
<td>Reagan</td>
<td>One short, chaotic briefing prior to election, no information conveyed (p. 129)</td>
<td>Defence-related campaign speeches from announcement until election (including victory speech).</td>
<td>96th</td>
</tr>
</tbody>
</table>

Table 2: Implied Ideological Rankings of Presidents

<table>
<thead>
<tr>
<th>Nominate</th>
<th>Text Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carter</td>
<td>Johnson</td>
</tr>
<tr>
<td>Kennedy</td>
<td>Carter</td>
</tr>
<tr>
<td>Johnson</td>
<td>Eisenhower</td>
</tr>
<tr>
<td>Eisenhower</td>
<td>Nixon</td>
</tr>
<tr>
<td>Ford</td>
<td>Kennedy</td>
</tr>
<tr>
<td>Nixon</td>
<td>Ford</td>
</tr>
<tr>
<td>Reagan</td>
<td>Reagan</td>
</tr>
</tbody>
</table>

Vice President. As a result we are able to use defence-related legislative speeches from President Ford’s time in the House of Representatives and Johnson’s in the senate. Naturally, when mapping President Ford to the House of Representatives, we remove his own speeches from the congressional text.

The speeches of Johnson and Ford were obtained from the Congressional Record. Those of the other presidents were collected from a variety of sources, most notably the American Presidency Project and the Presidential Libraries. Once the Presidential and Pre-Presidential speeches were identified and converted into plain text format we removed all 2 or 3 letter words along with the same stop words and legislative phrases as for the Congressional texts.
Figure 1: ICBM Launchers: Current Year Estimates and Actual Values
Figure 2: Average Error By Presidential Nominate Scores
Figure 3: Average Error By Presidential Text Scores
Figure 4: Doomsday Clock Readings
Table 3: Baseline Results and Results with Forward Prediction Interaction

<table>
<thead>
<tr>
<th></th>
<th>(1) Raw Error</th>
<th>(2) Raw Error</th>
<th>(3) Raw Error</th>
<th>(4) Raw Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>nominate</td>
<td>122.7780**</td>
<td>-</td>
<td>41.7981**</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(43.9328)</td>
<td></td>
<td>(11.9131)</td>
<td></td>
</tr>
<tr>
<td>nominate x forward</td>
<td>-</td>
<td>-</td>
<td>122.2863*</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(60.2285)</td>
<td></td>
</tr>
<tr>
<td>textscore</td>
<td>-</td>
<td>128.3956**</td>
<td>-</td>
<td>67.9675</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(41.4920)</td>
<td></td>
<td>(42.7537)</td>
</tr>
<tr>
<td>textscore x forward</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>97.5168*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(45.7739)</td>
</tr>
<tr>
<td>prediction length</td>
<td>-12.3360</td>
<td>-11.5837</td>
<td>-13.0829</td>
<td>-10.4313</td>
</tr>
<tr>
<td></td>
<td>(9.4133)</td>
<td>(9.4294)</td>
<td>(7.6841)</td>
<td>(8.0283)</td>
</tr>
<tr>
<td>icbm</td>
<td>-89.2173</td>
<td>-86.4642</td>
<td>-87.4130</td>
<td>-84.8373</td>
</tr>
<tr>
<td></td>
<td>(79.9567)</td>
<td>(79.3773)</td>
<td>(78.8349)</td>
<td>(78.6334)</td>
</tr>
<tr>
<td>bomber</td>
<td>1.3863</td>
<td>2.4523</td>
<td>1.5947</td>
<td>2.2492</td>
</tr>
<tr>
<td></td>
<td>(20.3192)</td>
<td>(19.8592)</td>
<td>(20.5393)</td>
<td>(20.1107)</td>
</tr>
<tr>
<td>keyhole</td>
<td>147.8808*</td>
<td>84.3208</td>
<td>167.1798**</td>
<td>102.8801</td>
</tr>
<tr>
<td></td>
<td>(64.1984)</td>
<td>(72.1566)</td>
<td>(59.7965)</td>
<td>(60.9977)</td>
</tr>
<tr>
<td>trend</td>
<td>-8.8676****</td>
<td>-7.5588</td>
<td>-8.9400***</td>
<td>-7.7697</td>
</tr>
<tr>
<td></td>
<td>(1.8497)</td>
<td>(4.5205)</td>
<td>(1.7793)</td>
<td>(4.6217)</td>
</tr>
<tr>
<td>constant</td>
<td>158.3686**</td>
<td>168.8970*</td>
<td>160.3360**</td>
<td>170.3417*</td>
</tr>
<tr>
<td></td>
<td>(49.3107)</td>
<td>(70.7754)</td>
<td>(47.9906)</td>
<td>(70.5820)</td>
</tr>
</tbody>
</table>

| N   | 317 | 317 | 317 | 317 |
| R²  | 0.2071 | 0.2109 | 0.2299 | 0.2251 |

i) Standard errors, clustered by president, in parentheses
ii) * p<0.10, ** p<0.05, *** p<0.01
iii) Specifications (1) and (2) give results for baseline specification.
iv) Specifications (3) and (4) interact ideology measures with a forward prediction dummy
### Table 4: Staff Turnover and Selection Effects

<table>
<thead>
<tr>
<th></th>
<th>(1) Raw Error</th>
<th>(2) Raw Error</th>
<th>(3) Raw Error</th>
<th>(4) Raw Error</th>
<th>(5) Raw Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>nominate</td>
<td>151.1746</td>
<td>-</td>
<td>-</td>
<td>83.2905**</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(58.9210)</td>
<td></td>
<td></td>
<td>(32.1732)</td>
<td></td>
</tr>
<tr>
<td>text score</td>
<td>-</td>
<td>156.4014**</td>
<td>-</td>
<td>-</td>
<td>89.9572**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(46.0389)</td>
<td></td>
<td></td>
<td>(27.4089)</td>
</tr>
<tr>
<td>mins to midnight</td>
<td>-</td>
<td>-</td>
<td>-18.1584**</td>
<td>-8.1579</td>
<td>-9.7031***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(6.4276)</td>
<td>(4.4896)</td>
<td>(2.1837)</td>
</tr>
<tr>
<td>icbm</td>
<td>-98.5658</td>
<td>-94.7287</td>
<td>-90.6895</td>
<td>-89.1198</td>
<td>-86.6780</td>
</tr>
<tr>
<td></td>
<td>(94.8776)</td>
<td>(94.1032)</td>
<td>(80.8267)</td>
<td>(80.1773)</td>
<td>(79.6334)</td>
</tr>
<tr>
<td>bomber</td>
<td>4.0521</td>
<td>5.3900</td>
<td>1.4630</td>
<td>1.3081</td>
<td>1.9667</td>
</tr>
<tr>
<td>keyhole</td>
<td>-</td>
<td>-</td>
<td>-95.0202</td>
<td>52.1905</td>
<td>-1.7127</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(67.7006)</td>
<td>(79.8556)</td>
<td>(70.2310)</td>
</tr>
<tr>
<td></td>
<td>(1.1688)</td>
<td>(5.0234)</td>
<td>(3.7535)</td>
<td>(3.1947)</td>
<td>(2.6922)</td>
</tr>
<tr>
<td>constant</td>
<td>169.5112*</td>
<td>181.1165*</td>
<td>188.3393**</td>
<td>178.7725**</td>
<td>194.5345**</td>
</tr>
<tr>
<td></td>
<td>(65.0388)</td>
<td>(83.7790)</td>
<td>(53.4925)</td>
<td>(52.2663)</td>
<td>(56.7386)</td>
</tr>
</tbody>
</table>

**N**

| 279 | 279 | 317 | 317 | 317 |

**R^2**

| 0.2212 | 0.2247 | 0.1973 | 0.2132 | 0.2239 |

i) Standard errors, clustered by president, in parentheses

ii) * p<0.10, ** p<0.05, *** p<0.01

iii) Specifications (1) and (2) examine the robustness of the baseline estimates to removing the post-Carter years

iv) Specifications (3), (4) and (5) include the reading on the doomsday clock as an additional control
Table 5: Evolution of Ideology Effect over a Presidency

<table>
<thead>
<tr>
<th></th>
<th>(1) Raw Error</th>
<th>(2) Raw Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>nominate x year 1</td>
<td>60.6072</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(31.2995)</td>
<td>-</td>
</tr>
<tr>
<td>nominate x year 2</td>
<td>194.9518**</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(77.4629)</td>
<td>-</td>
</tr>
<tr>
<td>nominate x year 3</td>
<td>109.9613*</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(54.6163)</td>
<td>-</td>
</tr>
<tr>
<td>nominate x year 4</td>
<td>204.1516**</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(70.4823)</td>
<td>-</td>
</tr>
<tr>
<td>textscore x year 1</td>
<td>-</td>
<td>52.6773</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>(28.8160)</td>
</tr>
<tr>
<td>textscore x year 2</td>
<td>-</td>
<td>254.5045***</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>(47.7158)</td>
</tr>
<tr>
<td>textscore x year 3</td>
<td>-</td>
<td>80.0555**</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>(28.9160)</td>
</tr>
<tr>
<td>textscore x year 4</td>
<td>-</td>
<td>156.7585**</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>(63.5276)</td>
</tr>
<tr>
<td>pred length</td>
<td>-13.8182</td>
<td>-12.1675</td>
</tr>
<tr>
<td></td>
<td>(10.1078)</td>
<td>(10.2812)</td>
</tr>
<tr>
<td>icbm</td>
<td>-120.0650</td>
<td>-113.1677</td>
</tr>
<tr>
<td></td>
<td>(87.7387)</td>
<td>(85.1466)</td>
</tr>
<tr>
<td>bomber</td>
<td>3.0418</td>
<td>5.7863</td>
</tr>
<tr>
<td></td>
<td>(23.5017)</td>
<td>(26.5557)</td>
</tr>
<tr>
<td>keyhole</td>
<td>209.8247**</td>
<td>106.9310</td>
</tr>
<tr>
<td></td>
<td>(81.6920)</td>
<td>(88.1375)</td>
</tr>
<tr>
<td>trend</td>
<td>-12.7908***</td>
<td>-8.0802</td>
</tr>
<tr>
<td></td>
<td>(2.4971)</td>
<td>(5.6615)</td>
</tr>
<tr>
<td>elec time</td>
<td>-27.8838***</td>
<td>-32.9314*</td>
</tr>
<tr>
<td></td>
<td>(3.3631)</td>
<td>(16.2665)</td>
</tr>
<tr>
<td>constant</td>
<td>261.8938***</td>
<td>229.2028*</td>
</tr>
<tr>
<td></td>
<td>(61.0221)</td>
<td>(108.3207)</td>
</tr>
</tbody>
</table>

\(N\)   248 248  
\(R^2\) 0.2890 0.3120

i) Standard errors, clustered by president, in parentheses
ii) * p<0.10, ** p<0.05, *** p<0.01

48
Table 6: Robustness Checks

<table>
<thead>
<tr>
<th></th>
<th>(1) Raw Error</th>
<th>(2) Raw Error</th>
<th>(3) Raw Error</th>
<th>(4) Raw Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>nominate</td>
<td>108.1681**</td>
<td>-</td>
<td>98.7182*</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(32.4761)</td>
<td>-</td>
<td>(49.6969)</td>
<td>-</td>
</tr>
<tr>
<td>text score</td>
<td>-</td>
<td>122.8558***</td>
<td>-</td>
<td>103.7091*</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>(28.7003)</td>
<td>-</td>
<td>(47.9997)</td>
</tr>
<tr>
<td>actual lag</td>
<td>0.1162</td>
<td>0.1365</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.0940)</td>
<td>(0.0823)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>pred length</td>
<td>-13.1846</td>
<td>-12.4750</td>
<td>-11.6188</td>
<td>-10.9127</td>
</tr>
<tr>
<td></td>
<td>(9.5227)</td>
<td>(9.5100)</td>
<td>(9.7875)</td>
<td>(9.7992)</td>
</tr>
<tr>
<td>icbm</td>
<td>-144.2934</td>
<td>-150.7403</td>
<td>-89.3435</td>
<td>-87.0009</td>
</tr>
<tr>
<td></td>
<td>(116.4122)</td>
<td>(111.7450)</td>
<td>(80.3390)</td>
<td>(79.4287)</td>
</tr>
<tr>
<td>bomber</td>
<td>12.2103</td>
<td>15.0426</td>
<td>2.3730</td>
<td>3.3066</td>
</tr>
<tr>
<td></td>
<td>(24.0255)</td>
<td>(23.0161)</td>
<td>(20.2067)</td>
<td>(19.9799)</td>
</tr>
<tr>
<td>keyhole</td>
<td>159.1007**</td>
<td>110.4300*</td>
<td>59.3707</td>
<td>2.0887</td>
</tr>
<tr>
<td></td>
<td>(57.7035)</td>
<td>(52.6629)</td>
<td>(111.0423)</td>
<td>(106.5184)</td>
</tr>
<tr>
<td></td>
<td>(4.2834)</td>
<td>(4.7270)</td>
<td>(6.7279)</td>
<td>(11.4630)</td>
</tr>
<tr>
<td>trend 2</td>
<td>-</td>
<td>-</td>
<td>0.3833</td>
<td>0.4190</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>(0.3214)</td>
<td>(0.4506)</td>
</tr>
<tr>
<td>_cons</td>
<td>189.8948**</td>
<td>210.7538**</td>
<td>182.9578***</td>
<td>194.9497**</td>
</tr>
<tr>
<td></td>
<td>(62.0595)</td>
<td>(75.6038)</td>
<td>(46.6309)</td>
<td>(66.9824)</td>
</tr>
<tr>
<td>N</td>
<td>317</td>
<td>317</td>
<td>317</td>
<td>317</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.2380</td>
<td>0.2551</td>
<td>0.2118</td>
<td>0.2172</td>
</tr>
</tbody>
</table>

i) Standard errors, clustered by president, in parentheses

ii) * p<0.10, ** p<0.05, *** p<0.01

iii) Specifications (1) and (2) control for lagged Soviet strength

iv) Specifications (3) and (4) allow for a quadratic trend specification
<table>
<thead>
<tr>
<th></th>
<th>(1) Percentage Error</th>
<th>(2) Percentage Error</th>
<th>(3) Percentage Error</th>
<th>(4) Percentage Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>nominate</td>
<td>1.5236**</td>
<td>-</td>
<td>1.1232***</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.4525)</td>
<td>(0.2486)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nominate x forward</td>
<td>-</td>
<td>-</td>
<td>0.6056</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>(0.3749)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>text score</td>
<td>-</td>
<td>0.6451</td>
<td>-</td>
<td>0.7970</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>(0.6732)</td>
<td></td>
<td>(0.5757)</td>
</tr>
<tr>
<td>text score x forward</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-0.2458</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td></td>
<td>(0.5458)</td>
</tr>
<tr>
<td>pred length</td>
<td>-0.0469</td>
<td>-0.0485</td>
<td>-0.0504</td>
<td>-0.0511</td>
</tr>
<tr>
<td></td>
<td>(0.0434)</td>
<td>(0.0401)</td>
<td>(0.0415)</td>
<td>(0.0441)</td>
</tr>
<tr>
<td>icbm</td>
<td>0.8772</td>
<td>0.7850</td>
<td>0.8875</td>
<td>0.7803</td>
</tr>
<tr>
<td></td>
<td>(1.2269)</td>
<td>(1.1724)</td>
<td>(1.2323)</td>
<td>(1.1684)</td>
</tr>
<tr>
<td>bomber</td>
<td>-0.3711**</td>
<td>-0.3328**</td>
<td>-0.3697**</td>
<td>-0.3314**</td>
</tr>
<tr>
<td></td>
<td>(0.1493)</td>
<td>(0.1321)</td>
<td>(0.1507)</td>
<td>(0.1308)</td>
</tr>
<tr>
<td>keyhole</td>
<td>2.7933***</td>
<td>1.6104*</td>
<td>2.8904***</td>
<td>1.5622*</td>
</tr>
<tr>
<td></td>
<td>(0.6528)</td>
<td>(0.7316)</td>
<td>(0.6174)</td>
<td>(0.7294)</td>
</tr>
<tr>
<td>trend</td>
<td>-0.2392***</td>
<td>-0.1804*</td>
<td>-0.2397***</td>
<td>-0.1797*</td>
</tr>
<tr>
<td></td>
<td>(0.0504)</td>
<td>(0.0903)</td>
<td>(0.0493)</td>
<td>(0.0896)</td>
</tr>
<tr>
<td>constant</td>
<td>3.2899***</td>
<td>2.8065*</td>
<td>3.2998***</td>
<td>2.7994*</td>
</tr>
<tr>
<td></td>
<td>(0.4144)</td>
<td>(1.2256)</td>
<td>(0.3976)</td>
<td>(1.2206)</td>
</tr>
<tr>
<td>N</td>
<td>306</td>
<td>306</td>
<td>306</td>
<td>306</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.1915</td>
<td>0.1411</td>
<td>0.1944</td>
<td>0.1415</td>
</tr>
</tbody>
</table>

i) Standard errors, clustered by president, in parentheses
ii) * p<0.10, ** p<0.05, *** p<0.01
iii) Specifications (1) and (2) give baseline results
iv) Specifications (3) and (4) include interaction with forward prediction dummy
Table 8: Pre-Carter Years and Superpower Relations with Percentage Error

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Percentage Error</td>
<td>Percentage Error</td>
<td>Percentage Error</td>
<td>Percentage Error</td>
<td>Percentage Error</td>
</tr>
<tr>
<td>nominate</td>
<td>1.7829**</td>
<td>-</td>
<td>-</td>
<td>1.0614</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.5526)</td>
<td>(0.5526)</td>
<td>(0.6261)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>textscore</td>
<td>-</td>
<td>0.6214</td>
<td>-</td>
<td>-</td>
<td>-0.3293</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.8562)</td>
<td></td>
<td></td>
<td>(0.3401)</td>
</tr>
<tr>
<td>mins to midnight</td>
<td>-</td>
<td>-</td>
<td>-0.2225***</td>
<td>-0.0979</td>
<td>-0.2544***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0559)</td>
<td>(0.0802)</td>
<td>(0.0500)</td>
</tr>
<tr>
<td>predlength</td>
<td>-0.0656</td>
<td>-0.0470</td>
<td>-0.0471</td>
<td>-0.0454</td>
<td>-0.0506</td>
</tr>
<tr>
<td></td>
<td>(0.0641)</td>
<td>(0.0511)</td>
<td>(0.0428)</td>
<td>(0.0438)</td>
<td>(0.0431)</td>
</tr>
<tr>
<td>icbm</td>
<td>1.0091</td>
<td>0.9169</td>
<td>0.8539</td>
<td>0.8872</td>
<td>0.8434</td>
</tr>
<tr>
<td></td>
<td>(1.4945)</td>
<td>(1.4270)</td>
<td>(1.2210)</td>
<td>(1.2333)</td>
<td>(1.2240)</td>
</tr>
<tr>
<td>bomber</td>
<td>-0.3761*</td>
<td>-0.3168*</td>
<td>-0.3680*</td>
<td>-0.3755**</td>
<td>-0.3718*</td>
</tr>
<tr>
<td></td>
<td>(0.1483)</td>
<td>(0.1167)</td>
<td>(0.1548)</td>
<td>(0.1522)</td>
<td>(0.1603)</td>
</tr>
<tr>
<td>keyhole</td>
<td>-</td>
<td>-</td>
<td>-0.2086</td>
<td>1.6652</td>
<td>-0.5526</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.3820)</td>
<td>(1.1501)</td>
<td>(0.4857)</td>
</tr>
<tr>
<td>trend</td>
<td>-0.2464***</td>
<td>-0.1829</td>
<td>-0.0895***</td>
<td>-0.1847**</td>
<td>-0.0675*</td>
</tr>
<tr>
<td></td>
<td>(0.0464)</td>
<td>(0.0936)</td>
<td>(0.0232)</td>
<td>(0.0606)</td>
<td>(0.0316)</td>
</tr>
<tr>
<td>trend 2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>constant</td>
<td>3.3813***</td>
<td>2.7747*</td>
<td>3.6380***</td>
<td>3.5567***</td>
<td>3.6292***</td>
</tr>
<tr>
<td></td>
<td>(0.3333)</td>
<td>(1.2649)</td>
<td>(0.6476)</td>
<td>(0.4118)</td>
<td>(0.5619)</td>
</tr>
<tr>
<td>N</td>
<td>268</td>
<td>268</td>
<td>306</td>
<td>306</td>
<td>306</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.1956</td>
<td>0.1388</td>
<td>0.1827</td>
<td>0.1960</td>
<td>0.1845</td>
</tr>
</tbody>
</table>

i) Standard errors, clustered by president, in parentheses

ii) * p<0.10, ** p<0.05, *** p<0.01

iii) Specifications (1) and (2) examine the robustness of the baseline estimates to removing the post-Carter years

iv) Specifications (3), (4) and (5) include the reading on the doomsday clock as an additional control
<table>
<thead>
<tr>
<th></th>
<th>(1) Percentage Error</th>
<th>(2) Percentage Error</th>
<th>(3) Percentage Error</th>
<th>(4) Percentage Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>nominate</td>
<td>0.1061</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.0778)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>text score</td>
<td>-</td>
<td>0.1776***</td>
<td>0.1291***</td>
<td>0.2927**</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>(0.0192)</td>
<td>(0.0309)</td>
<td>(0.0938)</td>
</tr>
<tr>
<td>text score x forward</td>
<td>-</td>
<td>-</td>
<td>0.0775*</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>(0.0339)</td>
<td>-</td>
</tr>
<tr>
<td>mins to midnight</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.0397</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>(0.0271)</td>
</tr>
<tr>
<td>pred length</td>
<td>-0.0268</td>
<td>-0.0249</td>
<td>-0.0242</td>
<td>-0.0261</td>
</tr>
<tr>
<td></td>
<td>(0.0160)</td>
<td>(0.0161)</td>
<td>(0.0155)</td>
<td>(0.0179)</td>
</tr>
<tr>
<td>icbm</td>
<td>-0.2455</td>
<td>-0.2354</td>
<td>-0.2338</td>
<td>-0.2321</td>
</tr>
<tr>
<td></td>
<td>(0.1720)</td>
<td>(0.1734)</td>
<td>(0.1733)</td>
<td>(0.1741)</td>
</tr>
<tr>
<td>bomber</td>
<td>-0.2491</td>
<td>-0.2478</td>
<td>-0.2483</td>
<td>-0.2448</td>
</tr>
<tr>
<td></td>
<td>(0.1862)</td>
<td>(0.1853)</td>
<td>(0.1856)</td>
<td>(0.1804)</td>
</tr>
<tr>
<td>keyhole</td>
<td>0.3461*</td>
<td>0.2748***</td>
<td>0.2876***</td>
<td>0.4661**</td>
</tr>
<tr>
<td></td>
<td>(0.1531)</td>
<td>(0.0353)</td>
<td>(0.0311)</td>
<td>(0.1424)</td>
</tr>
<tr>
<td>trend</td>
<td>-0.0386**</td>
<td>-0.0363***</td>
<td>-0.0363***</td>
<td>-0.0404***</td>
</tr>
<tr>
<td></td>
<td>(0.0123)</td>
<td>(0.0028)</td>
<td>(0.0026)</td>
<td>(0.0063)</td>
</tr>
<tr>
<td>constant</td>
<td>0.8639**</td>
<td>0.8513***</td>
<td>0.8510***</td>
<td>0.5210**</td>
</tr>
<tr>
<td></td>
<td>(0.2998)</td>
<td>(0.1751)</td>
<td>(0.1727)</td>
<td>(0.1398)</td>
</tr>
<tr>
<td>N</td>
<td>254</td>
<td>254</td>
<td>254</td>
<td>254</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.1767</td>
<td>0.2046</td>
<td>0.2064</td>
<td>0.2169</td>
</tr>
</tbody>
</table>

i) Standard errors, clustered by president, in parentheses
ii) * p<0.10, ** p<0.05, *** p<0.01
iii) Specifications (1) and (2) present results from the baseline specification
iv) Specification (3) interacts text score with a forward prediction dummy
iv) Specification (4) includes the doomsday clock reading as an additional control