Accountability Traps*

Scott Ashworth†  Ethan Bueno de Mesquita‡  Amanda Friedenberg§

December 18, 2013

Abstract

We introduce a concept of accountability traps. Two conditions characterize an accountability trap: (i) a society is caught in a self-reinforcing pattern of behavior with low accountability and (ii) within the same set of institutions, there is another self-reinforcing pattern of behavior with greater accountability and higher voter welfare. We show that a canonical model of elections is consistent with accountability traps. In the model, the source of accountability traps is “bad expectations” about governance. Specifically, if a society finds itself in an accountability trap, it is playing an equilibrium in which the voter expects politicians to shirk and politicians expect the voter to apply low standards for reelection. The results have implications for the role of institutional reform in development.

*For helpful comments and feedback, we are indebted to John Duggan, Rafael Hortala-Vallve, Dimitri Landa, Adam Meirowitz, David Myatt, Andrew Rehfeld, Rainer Schwabe, Dan Silverman, and seminar audiences at University of Arizona, Arizona State, Berkeley, Chicago, Maryland, Rochester, Virginia, the European Political Science Association Meetings, the Annual Meetings of the Midwest Political Science Association, the Society for Advancement of Economic Theory, and the World Congress of the Game Theory Society. This research was supported by NSF grants SES-0819152 (Bueno de Mesquita) and SES-0819200 (Friedenberg). In addition, Friedenberg thanks New York University and University College London for their hospitality.

†Harris School of Public Policy Studies, University of Chicago, email: sashwort@uchicago.edu
‡Harris School of Public Policy Studies, University of Chicago, email: bdm@uchicago.edu.
§W.P. Carey School of Business, Arizona State University, email: amanda.friedenberg@asu.edu
Electoral accountability is an important feature of democratic societies. The hope is that giving citizens formal accountability—that is, the formal right to reward or punish policy makers—will lead to policy choices that are better for citizen welfare.

This linkage is premised on the claim that formal accountability does, in fact, affect decisions. Two distinct empirical literatures have examined this issue. The first asks if voters actually use the channel of formal accountability—that is, if, in practice, they vote in a way that provides incentives (Canes-Wrone, Brady and Cogan, 2002; Wolfers, 2007; Berry and Howell, 2007; Ansolabehere and Jones, 2010). The second asks if formal accountability does, in fact, increase effective accountability—that is, if incumbent choices are in fact shaped by the incentives voters provide (Carey et al., 2006; Besley and Case, 1995, 2003; Alt, Bueno de Mesquita and Rose, 2011; de Janvry, Finan and Sadoulet, 2012; Finan and Ferraz, 2011; Gagliarducci and Nannicini, Forthcoming).

Of course, these two issues are deeply interwoven. An incumbent’s actions are shaped by how she expects voters to respond, and a voter’s electoral response to outcomes is conditioned on what actions he believes the incumbent has taken. This raises the possibility that different societies—with identical institutions—might differ in terms of how well they reap the welfare benefits of formal accountability: They may differ in the incumbent’s expectations of the voters’ behavior and the voters’ expectations of the incumbent’s behavior. As a consequence, they may differ in the actions chosen by the incumbent and so may differ in the voters’ welfare. If societies with identical institutions can differ in their level of effective accountability, it raises the possibility that there is scope for improving democratic performance and voter welfare without changing institutions.

To this end, we introduce the concept of an accountability trap. A model of electoral accountability is simply a game that describes a fixed set of institutions. The model is consistent with an accountability trap if two conditions hold. First, there are (at least) two equilibria of the model that differ in terms of their level of effective accountability. Second, voter welfare is higher in the equilibrium with a higher level of effective accountability. A society is caught in an accountability trap if it is in fact playing an equilibrium with both a lower level of effective accountability and a lower level of voter welfare than some other equilibrium.

We study a canonical model of the principal-agent relationship between a voter and a politician (Lohmann, 1998; Persson and Tabellini, 2000; Alesina and Tabellini, 2007, 2008). Within the context of this model, we provide necessary and sufficient conditions for the existence of an accountability trap. We split the analysis into two steps. First, we provide a necessary and sufficient condition for the existence of multiple equilibria that differ in their
level of effective accountability. Second, we provide necessary and sufficient conditions for voter welfare to be higher when there is a higher level of effective accountability. The conditions identified in these two steps are consistent and, so, are jointly necessary and sufficient for the existence of accountability traps.

In our model, the source of accountability traps is “bad expectations” about governance. Specifically, if a society finds itself in an accountability trap, it is playing an equilibrium where the voter expects the politician to shirk and the politician expects the voter to apply low standards for reelection.

The idea that “bad expectations” can lead to bad outcomes while “good expectations” can lead to good outcomes is a lesson familiar from the literature on poverty traps. Accountability traps can be seen as a political analogue to this important idea. But there are key differences between the two concepts. The equilibrium multiplicity of poverty traps is typically viewed as arising from an incentive to coordinate actions. By contrast, our model of accountability has multiple equilibria even when we abstract away from features typically associated with such multiplicity, e.g., incentives to coordinate actions, technological complementarities, signaling of private information, indifference, or an infinite time horizon.

Instead, in our model, accountability traps will be generated by a feature that—in our view—is intrinsic to the accountability relationship between a voter and a politician: the fact that the politician’s early behavior influences the voter’s ability to infer information about the politician’s characteristics. This is an idea reminiscent of Dewatripont, Jewitt and Tirole (1999a,b). They consider a principal-agent relationship where the agent’s early behavior influences the principal’s ability to infer the agent’s ability. They too have a multiplicity result based on expectations about the agent’s behavior. But, both the approach and the results here will differ in important ways from those in Dewatripont, Jewitt and Tirole (1999a,b). Here, any production function will be consistent with multiplicity whereas, Dewatripont, Jewitt and Tirole’s (1999a) multiplicity result requires technological complementarities. We discuss this further at the end of Section 3.

The link between the incumbent’s behavior and the voter’s ability to learn about the incumbent’s characteristics leads to a subtle connection between accountability and voter welfare. In particular, it is often presumed that increasing effective accountability is necessarily desirable from the perspective of voter welfare. But, we will see in Example 4.1 that this need not be the case. Elections allow voters to create effective accountability and to create electoral selection, i.e., to retain politicians whom the voters believe serve their future interests. A higher level of effective accountability can interfere with electoral selection. This creates the possibility that the overall quality of governance—as measured
in voter welfare—is higher with a lower level of effective accountability. If so, increasing effective accountability may not be a desirable goal.

The idea of accountability traps fits with a number of empirical observations: Countries with similar democratic institutions display considerable variation in the quality of governance outcomes. Further, in many societies with relatively bad governance, voters do not harshly sanction poorly performing politicians. (See, e.g., Bardhan, 1997; Chang, Golden and Hill, 2010; Golden, 2010.) This is consistent with the idea that different societies (with the same set of institutions) can differ in their expectations about governance. Cynical voters can have low expectations for government performance and, consequently, politicians are cynical and do not bother to work hard on the voters’ behalf. The result is poor governance outcomes that are not harshly sanctioned. Indeed, such outcomes are often associated with such cynicism. (See, e.g., the discussion in Bardhan, 1997.)

The existence of such accountability traps has important implications for improving effective accountability: It is often presumed that if a society seeks to increase effective accountability it should reform its institutions, i.e., change formal accountability. Indeed, the main focus of the literature is on the comparative statics of effective accountability with respect to changes in institutions, e.g., the rewards of office, the informational environment, electoral rules, or the length of term limits.¹ But, in the presence of an accountability trap, the first-order difficulty faced by the society may not be the current set of institutions. It may instead be societal expectations that leave it trapped in a situation with a low level of effective accountability. Hence, institutional reform may not be effective—even a reform that moves to an institution that, on average, has been associated with a higher level of accountability in structurally similar societies.

How then can effective accountability be improved? The answer depends on the underlying source of the accountability trap. In this paper, accountability traps arise because the politician’s early behavior influences the voter’s ability to infer information about the politician’s characteristics. (Again, we view such an inference problem as an intrinsic feature of the accountability relationship.) To escape such an accountability trap, a society must shift expectations.

There are at least two ways in which societal expectations can be shifted. First, small

institutional reforms may serve to shift expectations and break an accountability trap. For instance, small increases in the material benefits associated with holding office may render the “bad expectations” (i.e., associated with a low level of effective accountability) inconsistent with equilibrium expectations. (See Section 3.) Second, communication and leadership can play an important role in shifting expectations. (See Dewan and Myatt, 2007, 2008, 2012b and Bidner and Francois, Forthcoming.) And, indeed, improvements in effective accountability do appear consistent with such shifts. As Golden (2010) notes, improvements in accountability often happen suddenly, “as part of a wave of public revulsion.” Such sudden shifts in accountability are consistent with shifts in voter expectations—as opposed to large-scale institutional reform, which is often a long-term process.

Of course, there are other possible sources of accountability traps, e.g., incentives to coordinate voter actions or an infinite time horizon. (Section 5 discusses other models that can potentially generate accountability traps, due to these or other features.) Accountability traps that arise from these sources can be addressed in a relatively straightforward manner. For instance, if the view is that accountability traps necessarily arise because of miscoordination among voters, a clear policy prescription would be to find ways of increasing coordination. Likewise, if the view is that accountability traps necessarily arise because of a long-term (i.e., infinite) relationship, a natural policy prescription would be to impose term limits. An important contribution of our approach is to show that accountability traps can persist, even when these other sources are eliminated. Such persistence can arise because of an intrinsic feature of the accountability relationship.

The paper proceeds as follows. Section 1 presents a canonical model of elections. Section 2 gives a preview of the approach; it highlights both conceptual choices and formal issues that we will encounter. Sections 3-4 provide necessary and sufficient conditions for the existence of accountability traps. Section 5 discusses both other political traps and other sources of accountability traps; in so doing, it highlights why the model in Section 1 fits the question at hand.

1 The Model

There are three players: an Incumbent (I), Challenger (C) and Voter (V). We refer to each Politician (P) as “she” and the Voter as “he.” Each Politician P is either a high type, viz. \(\overline{\theta}\), or a low type, viz. \(\underline{\theta}\), where \(\overline{\theta} > \underline{\theta}\). Write \(\Pr(\overline{\theta}) \in (0, 1)\) for the probability that a Politician is of high type. This probability is commonly understood by the players.

In each period, the Politician in office takes an action \(a\) from a set of actions \(A\), where
A is a closed subset of the real line. Higher actions should be thought of as higher levels of effort. We will let $a$ represent the smallest element of $A$.

The level of public goods produced for the voter is a function of the action chosen by the Politician, the type of the Politician in office, and an idiosyncratic shock. If the Politician (in office) chooses action $a$ and is of type $\theta$, the level of public goods produced would be $f(a, \theta)$, absent the idiosyncratic shock. The function $f$ is strictly increasing in actions ($a$) and type ($\theta$). We refer to the function $f$ as the **production function**. The total level of public goods produced is the sum of the level produced via the production function and a random shock: If in period $t$ the Politician in office chooses action $a$, is of type $\theta$, and the random shock is $\epsilon_t$, the level of public goods produced in that period is

$$f(a, \theta) + \epsilon_t.$$ 

Each $\epsilon_t$ is the realization of a random variable. These random variables (i.e., the shocks in periods one and two) are independent of each other, and are independent of the Politicians’ abilities. In particular, each of these random variables are distributed according to an absolutely continuous CDF $\Phi$ with a PDF $\phi$ satisfying the following requirements: First, the distribution of the random variables is **atomless** (i.e., the probability of any realization of $\epsilon_t$ is zero). Second, the PDF $\phi$ is **symmetric about zero** (i.e., for each $x \in \mathbb{R}$, $\phi(x) = \phi(-x)$) and **single-peaked about zero** (i.e., $\phi$ is strictly increasing on $(-\infty, 0)$ and strictly decreasing on $(0, \infty)$).\(^2\) Third, the PDF $\phi$ satisfies the **(strict) monotone likelihood ratio property** (MLRP) relative to $A \times \Theta$: If $(a, \theta) > (a', \theta')$, then

$$\frac{\phi(g - f(a, \theta))}{\phi(g - f(a', \theta'))}$$

is strictly increasing in $g$.\(^3\) These assumptions are all satisfied by the Normal distribution.

Now we turn to the timing of the game:

**Time 1** Nature determines the realizations of each Politician’s type and of the random shock (in all periods). These realizations are not observed by any of the players.

**Time 2** Governance period 1 occurs:

2.1 The Incumbent chooses an action $a_1$. Her choice is not observed by the other players.

---

\(^2\)This implies that the random variable has support $\mathbb{R}$, since $\phi(x) > 0$ for all $x \in \mathbb{R}$.

\(^3\)We take the convention that $(a, \theta) > (a', \theta')$ if either $a \geq a'$ and $\theta > \theta'$ or $a > a'$ and $\theta \geq \theta'$. 

2.2 All players observe the level of public goods produced.

**Time 3** The Voter makes a choice to reelect the Incumbent or replace her with the Challenger. This choice is observed by all players.

**Time 4** Governance period 2 occurs:

4.1 The Politician in office (i.e., the winner of the election) chooses an action $a_2$. Her choice is not observed by the other players.

4.2 All players observe the level of public goods produced.

**Time 5** The game ends.

The Voter values public goods in each period. Writing $g_t$ for the level of public goods provided in governance period $t$, the Voter’s payoffs are $g_1 + g_2$.

Each Politician’s payoffs depend on both a benefit from holding office and the action chosen while in office. The benefit from holding office is given by $B > 0$. The cost of taking an action $a$ is given by a cost function $c(\cdot)$, where $c(\cdot)$ is strictly increasing and $B > c(a) \geq 0$.

A Politician’s payoff in governance period $t$ is $0$ if she is not in office and $B - c(a_t)$ if she is in office, where $a_t$ is the action she chose in period $t$. A Politician’s payoffs are given by the sum of her payoffs in each governance period.

2 Preview of the Approach

There will be two steps in showing that the model is consistent with an accountability trap. The first is to show that there are multiple equilibria of the model that differ in their level of effective accountability. (See Section 3.) The second is to show that the equilibrium with the higher level of effective accountability is also the equilibrium with a higher level of Voter welfare. (See Section 4.)

Begin with the existence of multiple equilibria that differ in their level of effective accountability. In the second period, no Politician has an incentive to take a costly action, as doing so offers no electoral benefit. So, if there are multiple equilibria that differ in their level of effective accountability, they must involve different first-period actions.

At first glance, an argument familiar from Fearon (1999) might suggest that finding two such equilibria may not be feasible: Since the Incumbent and the Challenger choose the same second-period action, there is only one thing that distinguishes the candidates at the time of the election—namely the Voter’s expectation of each Politician’s ability.
The Voter strictly prefers to elect the Politician with the strictly higher level of ability, i.e., to “select good types.” Quite generally, this incentive to select good types will lead the Voter to have a strict preference for one candidate over the other. If the first-period outcome is sufficiently favorable (respectively, poor), the Voter will think that the Incumbent (respectively, Challenger) is strictly more likely to be a high type and, so, will strictly prefer to elect the Incumbent (respectively, Challenger). And, if the Incumbent anticipates this strict preference, the Incumbent will, in turn, quite generally, have a unique best response.

So, given a prior belief about the Incumbent’s type and Incumbent’s action, the Voter quite generally has a strict preference. But the key is that equilibria with different levels of effective accountability will be associated with different Voter expectations about the Incumbent’s actions. In turn, different Voter expectations about what action the Incumbent will take affects what the Voter demands of the Incumbent and, therefore, how hard the Incumbent is willing to work. Thus, despite the strict preference, in principle, there may be multiple equilibria that differ in their level of effective accountability.

The formal analysis establishes when such multiplicity can or cannot occur. To address the question, we fix the production technology and the beliefs (about type). We ask: Do there exist a pair \((B, c)\), i.e., a benefit of reelection and a cost function, so that there are multiple pure strategy perfect Bayesian equilibria that differ in their first-period action? We provide a necessary and sufficient condition on the production technology and beliefs, which generates multiple equilibria that differ in their level of effective accountability. For any given production function and distribution of noise, this condition can be met provided there is a sufficiently high probability that any given Politician is a high type.

Why is this the question of interest? We view both the Incumbent’s benefit from reelection and the cost of effort as fundamentally subjective. The application might suggest certain material benefits of reelection (e.g., salary, prestige, and so on) and material costs of higher actions (e.g., foregone rents, time not devoted to other policy areas, and so on). But the application itself cannot pin down the Politician’s utility from these material outcomes. On the other hand, for a given application, the analyst may have intuitions or empirical knowledge about the nature of the production function or the beliefs. For instance, in a particular application, the analyst may think that it is quite likely that there are many high types in the pool of potential politicians and so it is quite likely that any given Politician is a high type. Or, in a particular application, the analyst may think effort and type are complements/substitutes. Thus, it is of interest to understand conditions on production and beliefs that imply that there is a benefit of reelection and cost function so that, in the
associated game, there are multiple equilibria that differ in terms of first-period effort.\footnote{We use the phrase ‘multiple equilibria’ as multiple pure strategy Perfect Bayesian equilibria. Of course, the existence of multiple pure strategy equilibria implies the existence of multiple mixed strategy equilibria.}

The next step is to show that the equilibrium with the higher level of effective accountability is also the equilibrium with a higher level of Voter welfare. The equilibrium with a higher level of effective accountability (i.e., the equilibrium that involves electoral incentives shaping Incumbent actions to a greater extent) corresponds to the equilibrium with a higher first-period action. Thus, we are interested in determining when an equilibrium in which the Incumbent has incentives to exert high effort is better for Voter welfare than an equilibrium in which the Incumbent has incentives to exert low effort.

At first glance, it may seem obvious that the equilibrium with a higher level of effective accountability also involves a higher level of Voter welfare: When the Incumbent chooses a more productive first-period action, it increases the first-period production of public goods and so increases first-period Voter welfare. But, nonetheless, the equilibrium associated with a more productive action may not be the equilibrium associated with a higher level of Voter welfare. It may be more difficult for the Voter to “select good types” in the equilibrium with a higher level of effective accountability. If so, second-period Voter welfare will be lower in the equilibrium with higher effective accountability. If this effect is sufficiently large, overall Voter welfare can be lower with a higher level of effective accountability—that is, higher levels of effective accountability may actually be detrimental for outcomes and so Voter welfare.

We will show that, if effort and ability are (weakly) complementary for the production of public goods, higher first-period effort will \textit{not} lead to lower second-period Voter welfare. Note, this complementarity is not needed to get a result on the existence of multiple equilibria—the necessary and sufficient conditions we identify above are consistent with effort and ability being strict substitutes. Instead, the complementarity is sufficient to guarantee that, if there are multiple equilibria, the one with a higher level of effective accountability also has a higher level of Voter welfare. And, while it is sufficient for this conclusion, it is not necessary to reach the conclusion. Even if a higher level of effective accountability leads to a lower second-period Voter welfare, overall Voter welfare will still increase if the first-period welfare gains are sufficiently large to compensate for the second-period welfare losses. We show that this is necessarily the case if the prior probability of a politician being a high or low type is sufficiently large.

Thus, we have two sets of necessary and sufficient conditions: The first generates multiple equilibria that differ in their level of effective accountability. The second ensures that,
if an equilibrium has a higher level of effective accountability, it also has a higher level of
Voter welfare. These conditions are consistent. So, taken together, they establish necessary
and sufficient conditions under which the model is consistent with accountability traps.

It is worth noting that the fact that higher effective accountability need not go along
with higher Voter welfare points to a subtlety in the normative interpretation of effective
accountability that is often missed by the literature. Suppose there are multiple equilibria
that differ in their level of effective accountability, but there is no accountability trap. Then
improving effective accountability is actually bad for Voter welfare, counter to standard
intuitions about accountability, governance, and development. Understanding this subtlety
is of independent import. We discuss it in greater detail at the end of Section 4.

3 Analysis: Multiple Equilibria

We begin by analyzing equilibrium behavior in the game relative to a fixed benefit of
reelection and cost function \((B, c)\). Then we turn to the question of choosing the pair \((B, c)\)
such that there are multiple equilibria that differ in terms of first period Incumbent effort.

Second Governance Period There are no electoral benefits from choosing a costly
action \((a_2 > a)\) in the second governance period, as there is no future election. As such,
the politician in office will choose the lowest possible action \((a_2 = a)\), independent of the
history.

The Voter’s Electoral Decision Because both candidates will choose the lowest possible
action in the second governance period, the Voter’s electoral decision depends only on his
expectation about the politicians’ types. The Voter will prefer to reelect the Incumbent if
and only if his posterior beliefs, conditional on observing \(g_1\), imply that the Incumbent’s
expected type is at least as high as the Challenger’s. Otherwise, the Voter strictly prefers
the Challenger.

Here we see the logic of selecting good types: The Voter updates his belief about the
Incumbent’s type. His reelection decision can be described by a cutpoint in the space of
posterior beliefs. The cutpoint is the probability that the Challenger is the high type, viz.
Pr(\(\theta\)). If the Voter’s posterior belief lies strictly above the cutpoint, then he reelects the
Incumbent. If the Voter’s posterior belief lies strictly below the cutpoint, then he elects the
Challenger.

The Voter’s posterior depends both on the level of public goods he observes in the first
period, $g_1$, and his belief about the Incumbent’s behavior in the first governance period. Specifically, the Voter’s posterior belief that the Incumbent is of type $\overline{\theta}$, given a level of public goods $g_1$ and an expected first-period action $\bar{a}$, written $\Pr(\overline{\theta}|g_1, \bar{a})$, is

$$
\Pr(\overline{\theta}|g_1, \bar{a}) = \frac{\Pr(\overline{\theta})\phi(g_1 - f(\bar{a}, \overline{\theta}))}{\Pr(\overline{\theta})\phi(g_1 - f(\bar{a}, \overline{\theta})) + (1 - \Pr(\overline{\theta}))\phi(g_1 - f(\bar{a}, \overline{\theta}))}.
$$

The Voter reelects the incumbent if and only if $\Pr(\overline{\theta}|g_1, \bar{a}) \geq \Pr(\overline{\theta})$, which holds if and only if

$$
\frac{\phi(g_1 - f(\bar{a}, \overline{\theta}))}{\phi(g_1 - f(\bar{a}, \overline{\theta}))} \geq 1.
$$

(1)

By the MLRP, the left-hand side of Equation 1 is strictly increasing in $g_1$ and strictly decreasing in $\bar{a}$. This yields two important observations. First, if the Voter believes the Incumbent’s first-period action is $\bar{a}$, the Voter reelects the Incumbent if and only if the outcome $g_1$ is greater than a threshold $\hat{g}(\bar{a})$, where $\hat{g}(\bar{a})$ is implicitly defined by:

$$
\frac{\phi(\hat{g}(\bar{a}) - f(\bar{a}, \overline{\theta}))}{\phi(\hat{g}(\bar{a}) - f(\bar{a}, \overline{\theta}))} = 1.
$$

(2)

Second, the threshold for reelection is strictly increasing in the Voter’s belief about the Incumbent’s action, i.e., $\hat{g}(\bar{a})$ is strictly increasing in $\bar{a}$. (See Lemma A.5.) That is, if the Voter believes the Incumbent worked harder in the first period, the Voter holds the Incumbent to a more stringent standard.

The logic of this second fact is as follows. For a given level of public goods, the Voter’s posterior belief will differ based on what action he initially believed the Incumbent took. In particular, the harder the Voter believes the Incumbent worked, the less he believes any given outcome is a sign of the Incumbent being a high type, since a larger portion of any success is credited to effort rather than to competence. As such, for any given level of public goods, $g_1$, the Voter’s posterior belief that the Incumbent is a high type will be greater if he believes the Incumbent chose some lower action $a''$ instead of some higher action $a' > a''$.

Figure 3.1 illustrates this point. Regardless of his beliefs about the Incumbent’s action, the Voter always uses the same cutpoint in the space of posterior beliefs. (The cutoff is $\Pr(\overline{\theta})$ on the $y$-axis.) However, in order to do so, he uses different cutpoints with respect to the level of public goods, depending on his belief about the Incumbent’s action. In particular, if the Voter believes the Incumbent took an action $\bar{a}$, he will reelect her if and only if the level public goods rises above $\hat{g}(\bar{a})$. The fact that $\hat{g}(a') > \hat{g}(a'')$ for $a' > a''$ reflects the fact that the Voter—selecting good types—is willing to reelect for a lower level of public goods
provision if he believes the Incumbent did not exert effort (i.e., chose a lower action).

**First-Period Action** Now we turn to the Incumbent’s first-period choice. By choosing a higher action, the Incumbent increases the probability that she will produce a high enough level of public goods to exceed the Voter’s threshold. Thus taking a higher action increases the likelihood that she will be reelected. But the extent to which it does so depends on the Voter’s expectation of the Incumbent’s first-period action. We now review why.

We saw that the Voter sets a standard for the level of public goods provision: He reelects the Incumbent if and only if the level of public goods meets some threshold. The threshold depends on the action the Voter believes the Incumbent chose. By choosing a higher action, the Incumbent increases the probability of reelection. The amount by which the probability of reelection increases depends on the Voter’s beliefs about the Incumbent’s action.

Suppose the Incumbent chooses the action $a$ and the Voter expects the Incumbent to choose action $a_*$. The Incumbent is reelected if and only if the level of public goods exceeds the Voter’s threshold. That is, the Incumbent is reelected if $f(a, \theta_I) + \epsilon_1 > \hat{g}(a_*)$ and replaced if $f(a, \theta_I) + \epsilon_1 < \hat{g}(a_*)$. Write $Pr(a|a_*)$ for the probability that Incumbent is reelected if she chooses $a$ when the Voter expects $a_*$. This is the probability that

$$\epsilon_1 \geq \hat{g}(a_*) - f(a, \theta_I).$$
(Here we use the fact that the random variable generating $\epsilon_1$ is atomless.) Thus, the probability is given by:

$$
Pr(a|a_*) = \Pr(\theta) \left[ 1 - \Phi (\hat{g}(a_*) - f(a, \theta)) \right] + \Pr(\theta) \left[ 1 - \Phi (\hat{g}(a_*) - f(a, \theta)) \right].
$$

Now, consider two actions $a, a' \in A$. If the Voter expects the Incumbent to take the action $a_*$, then the **incremental increase in probability of reelection** from choosing $a'$ instead of $a$ is

$$
IR(a', a|a_*) = Pr(a'|a_*) - Pr(a|a_*).
$$

In the second governance period, the Incumbent’s value of holding office is $B - c(a)$. (This uses the fact that the Incumbent chooses the low action, irrespective of the history.) So, if the Voter expects the Incumbent to take action $a_*$, then the **incremental benefit** that the Incumbent obtains by choosing the action $a'$ rather than $a$, given that the Voter believes the Incumbent took action $a_*$, is $IR(a', a|a_*)(B - c(a))$. The **incremental cost** to the Incumbent of choosing $a'$ rather than $a$ is $c(a') - c(a)$. The Incumbent will prefer $a'$ over $a$, given the Voter’s belief, when the incremental benefit of choosing $a'$ over $a$ is at least as large as the incremental cost of choosing $a'$ over $a$.

There is an equilibrium where the Incumbent chooses the action $a_1 = a_*$ in the first period if and only if, for each action $a \in A$, $IR(a_*, a|a_*)(B - c(a)) \geq c(a) - c(a)$. That is, there is an equilibrium where the Incumbent chooses the action $a_1 = a_*$ in the first period if and only if the incremental benefit of choosing $a_*$ over any other action (when the Voter expects her to choose $a_*$) is higher than the incremental cost of choosing $a_*$ over any other action.

**Choosing the Benefit of Reelection and Cost Function** Recall the question we are asking: Do there exist a benefit of reelection and a cost function so that, in the associated game, there are multiple equilibria that differ in their level of effective accountability (i.e., first-period action)? With this in mind, say that a pair $(B, c)$ justifies an action $a'$ if, when $B$ is the benefit of reelection and $c$ is the cost function, there is a perfect Bayesian equilibrium in which the politician takes action $a'$ in the first period. In light of the above discussion, it suffices to look for a pair $(B, c)$ that justifies two actions: $a_*$ and $a_{**}$.

Notice, $(B, c)$ justifies $a_*$ if and only if

$$
IR(a_*, a|a_*)(B - c(a)) \geq c(a_*) - c(a) \quad \text{for all } a \in A.
$$
and \((B, c)\) justifies \(a_{**}\) if and only if

\[
\text{IR}(a_{**}, a'|a_{**})(B - c(a)) \geq c(a_{**}) - c(a') \quad \text{for all } a' \in A.
\] (4)

In Equation 3 take \(a = a_{**}\) and in Equation 4 take \(a' = a_\ast\). Then, \((B, c)\) justifies both \(a_\ast\) and \(a_{**}\) only if

\[
\text{IR}(a_{**}, a_\ast|a_{**}) \geq \frac{c(a_{**}) - c(a_\ast)}{B - c(a)} \geq -\text{IR}(a_\ast, a_{**}|a_\ast).
\]

Notice that \(-\text{IR}(a_\ast, a_{**}|a_\ast) = \text{IR}(a_{**}, a_\ast|a_\ast)\). Thus \((B, c)\) justifies both \(a_\ast\) and \(a_{**}\) only if

\[
\text{IR}(a_{**}, a_\ast|a_{**}) \geq \frac{c(a_{**}) - c(a_\ast)}{B - c(a)} \geq \text{IR}(a_{**}, a_\ast|a_\ast).
\] (5)

This leads to a necessary condition \(a_\ast\) and \(a_{**}\) to be justifiable. Equation 5 says that, if \(a_\ast\) and \(a_{**}\) are justifiable, then

\[
\text{IR}(a_{**}, a_\ast|a_{**}) \geq \text{IR}(a_{**}, a_\ast|a_\ast).
\] (6)

To interpret this necessary condition, take \(a_{**} > a_\ast\). Equation 6 says that the incremental increase in probability of reelection of moving from the lower action \((a_\ast)\) to the higher action \((a_{**})\) must be higher when the Voter expects the Incumbent to take the higher action than when the Voter expects the Incumbent to take the lower action.

Of course, we are interested in providing a sufficient condition for \(a_\ast\) and \(a_{**}\) to be justifiable. If Equation 6 is satisfied, then we can find a benefit of reelection and a cost function satisfying Equation 5. If \(a_\ast\) and \(a_{**}\) were the only possible actions, this would suffice to show that Equations 3 and 4 are satisfied, establishing our result. But, Equation 6 is silent about actions distinct from \(a_\ast\) and \(a_{**}\). Hence, if there are more than two possible actions (i.e., if \(A\) contains more than two elements), we may have a benefit of reelection and cost function that satisfies Equation 5 but which fails to justify either \(a_\ast\) or \(a_{**}\). In particular, for a given benefit of reelection and cost function, Equation 3 may fail for some action \(a\), even if it holds for \(a_{**}\). Similarly, Equation 4 may fail for some action \(a'\), even if it holds for \(a_\ast\).

Nonetheless, we will see that Equation 6 is also sufficient to ensure that \(a_\ast\) and \(a_{**}\) are justifiable. That is, if Equation 6 holds, then we can always construct some benefit of reelection and some cost function such that Equations 3 and 4 are satisfied.
Proposition 3.1 Consider actions, $a_{**}, a_{\ast} \in A$, so that $a_{**} > a_{\ast}$. There exists a benefit of reelection and a cost function that justifies both $a_{\ast}$ and $a_{**}$ if and only if

$$\text{IR} (a_{**} , a_{\ast} | a_{**}) \geq \text{IR} (a_{**} , a_{\ast} | a_{\ast}).$$

Proposition 3.1 identifies a necessary and sufficient condition for two distinct actions $a_{\ast}$ and $a_{**}$ to be justifiable. Notice, the condition is local—i.e., defined relative to only the actions $a_{\ast}$ and $a_{**}$ that we are attempting to justify. So, for instance, the condition may be satisfied for any two actions—but not for all actions—and, even so, it is sufficient to generate the multiple equilibria result.

It is surprising that the local condition suffices. To see why it does, consider the case where $a_{\ast} = a_{**}$. The idea will be to fix a benefit of reelection, viz. $B$, and constants $n_{\ast}$ and $n_{**}$. The constants will turn out to be the costs associated with the high and the lowest actions—i.e., when we later choose a cost function $c$ that justifies both $a_{\ast}$ and $a_{**}$, it will satisfy $c(a_{\ast}) = n_{\ast}$ and $c(a_{**}) = n_{**}$. As such, we fix $B > n_{\ast} > 0$ and (in light of the necessity condition given by Equation 5), we fix $n_{**} > n_{\ast}$ so that

$$\text{IR} (a_{**} , a_{\ast} | a_{**}) \geq \frac{n_{**} - n_{\ast}}{B - n_{\ast}} \geq \text{IR} (a_{**} , a_{\ast} | a_{\ast}).$$

Suppose, in fact, that $c$ is a cost function with $c(a_{\ast}) = n_{\ast}$ and $c(a_{**}) = n_{**}$. Notice that, if $(B, c)$ justifies $a_{\ast}$, it must be that $c(\cdot)$ lies above the function $N(\cdot, a_{\ast}) : A \rightarrow \mathbb{R}$ where

$$N(a, a_{\ast}) = \text{IR} (a, a_{\ast} | a_{\ast})(B - n_{\ast}) + n_{\ast}.$$

If this condition were not satisfied, the Incumbent would have an incentive to deviate from $a_{\ast}$ to an alternate action. Analogously, if $(B, c)$ justifies $a_{**}$, it must be that $c(\cdot)$ lies above the function $N(\cdot, a_{**}) : A \rightarrow \mathbb{R}$ with

$$N(a, a_{**}) = \text{IR} (a, a_{**} | a_{**})(B - n_{**}) + n_{**}.$$

Figure 3.2 illustrates the functions $N(\cdot, a_{\ast})$ and $N(\cdot, a_{**})$. Each of these functions are strictly increasing. Moreover, by Equation 7, $N(a_{\ast}, a_{\ast}) = n_{\ast} \geq N(a_{\ast}, a_{**})$ and $N(a_{**}, a_{**}) = n_{**} \geq N(a_{**}, a_{\ast})$. Thus, we can take $c$ to be the upper envelope of the functions and get that $(B, c)$ justifies both $a_{\ast}$ and $a_{**}$.

To sum up, we have seen that Equation 6 is necessary and sufficient to justify both the lowest action $a_{\ast}$ and some higher action $a_{**}$. But what if the lower action is not the lowest
possible action—i.e., if \( a_* \neq a^? \). Now the construction involves choosing four parameters: the benefit of reelection \( B \), the cost of \( a \), the cost of \( a_* \), and the cost of \( a^{**} \). We must choose these parameters so that (i) the Incumbent has no incentive to deviate from one “equilibrium action” to the second and (ii) the Incumbent has no incentive to deviate from some “equilibrium action” to the lowest action. The proof demonstrates that the second requirement is implied by the first. Thus, Equation 6 is sufficient.

Finally, when the condition from Proposition 3.1 holds with strict inequality, there is a significant set of benefits of reelection and cost functions that justify both \( a_* \) and \( a^{**} \).

**Proposition 3.2** Suppose \( A \) is finite. Then the following are equivalent.

(i) There exists a non-empty open set of benefits of reelection and cost functions, so that each element of the set justifies both \( a_* \) and \( a^{**} \).

(ii) \( \text{IR}(a^{**}, a_* | a^{**}) < \text{IR}(a^{**}, a_* | a_*) \).

Proposition 3.1 (respectively 3.2) identifies conditions under which the model has multiple perfect Bayesian equilibria that differ in their level of effective accountability. We have seen that, when the model has multiple equilibria, they differ in the Incumbent’s first-period action, but not in the politicians’ second-period actions. In each such equilibrium, the Voter sets standards for the purpose of selecting good types. But, because the Voter has different expectations about Incumbent behavior in the two equilibria, he applies different reelection
standards in the two equilibria. In particular, the Voter reelects for lower levels of public goods when he expects a lower first-period action by the Incumbent.

**Primitives of the Model**  It is important to note that the condition identified in Proposition 3.1 (or Proposition 3.2) can be expressed in terms of primitives of the model—specifically, by referencing only the production technology and beliefs about type and shocks. Lemma A.5 and Proposition A.1 in the Appendix illustrate how to do so.\(^5\)

But, while the condition can be expressed in terms of primitives of the model, it may not be immediate to see which substantive assumptions about the technology and information do or do not meet the requirement. We show that the condition is consistent with any production function and any distribution of noise, provided politicians are sufficiently likely to be a high type.\(^6\)

**Proposition 3.3**  Fix actions \(a_{**} \neq a_*\). There exists a \(\hat{p}[a_*, a_{**}] \in (0, 1)\) so that \(a_*\) and \(a_{**}\) are justifiable if and only if \(\Pr(\theta) \geq \hat{p}[a_*, a_{**}]\).

This says that \(a_*\) and \(a_{**}\) are justifiable if the pool of candidates is sufficiently likely to be high quality. The phrase ‘sufficiently likely’ is determined by a benchmark \(\hat{p}[a_*, a_{**}]\). The Appendix specifies this benchmark based on (i) the distribution of noise and (ii) the value of the production function at \(a_*\) and \(a_{**}\). Thus, the benchmark is action-dependent. In particular, for a given distribution of noise and production function, \(\Pr(\theta)\) may meet the benchmark for actions \(a_*, a_{**}\) but may not meet the benchmark for other actions. If the set of actions is finite, every action is justifiable provided \(\Pr(\theta)\) is sufficiently high.

It is important to note that \(\hat{p}[a_*, a_{**}]\) need not be ‘close’ to 1, as we now show.

**Proposition 3.4**

(i) If \(f(a_{**}, \theta) - f(a_{**}, \theta) > f(a_*, \theta) - f(a_*, \theta)\), then \(\hat{p}[a_*, a_{**}] < \frac{1}{2}\).

(ii) If \(f(a_{**}, \theta) - f(a_*, \theta) = f(a_*, \theta) - f(a_*, \theta)\), then \(\hat{p}[a_*, a_{**}] = \frac{1}{2}\).

(iii) If \(f(a_{**}, \theta) - f(a_*, \theta) < f(a_*, \theta) - f(a_*, \theta)\), then \(\hat{p}[a_*, a_{**}] > \frac{1}{2}\).

Part (i) says that if, at \(a_*\) and \(a_{**}\), effort and type are strictly complementary, then \(\hat{p}[a_*, a_{**}]\) is strictly less than \(\frac{1}{2}\). Part (ii) says that if, at \(a_*\) and \(a_{**}\), effort and type are

---

\(^5\)This is one reason we restrict attention to two types: An analogue to Proposition 3.1 obtains with multiple types. By restricting attention to two types, it is easy to express IR (\(|\theta|\)) in terms of primitives of the model. See Footnote 7 for the second reason we restrict attention to two types.

\(^6\)Of course, the distribution of noise must satisfy the assumptions stated earlier.
neither strict complements nor strict substitutes, then \( \hat{p}[a_*, a_{**}] \) is equal to \( \frac{1}{2} \). Part (iii) says that if, at \( a_* \) and \( a_{**} \), effort and type are strict substitutes, then \( \hat{p}[a_*, a_{**}] \) is strictly greater than \( \frac{1}{2} \). Note, here, complementarity and substitutability are local conditions, i.e., defined only with respect to the first-period equilibrium actions \( a_* \) and \( a_{**} \).

To understand Proposition 3.3, fix some environment so that \( a_{**} \) and \( a_* \) are both justifiable. By Proposition 3.1, \( \text{IR}(a_{**}, a_*|a_{**}) - \text{IR}(a_{**}, a_*|a_*) \geq 0 \). The key is that \( \text{IR}(a_{**}, a_*|a_{**}) - \text{IR}(a_{**}, a_*|a_*) \) is strictly increasing in \( \text{Pr}(\theta) \). (See Remark A.1.) Thus, after increasing \( \text{Pr}(\theta) \), both \( a_{**} \) and \( a_* \) remain justifiable.

The fact that \( \text{IR}(a_{**}, a_*|a_{**}) - \text{IR}(a_{**}, a_*|a_*) = \text{IR}(a_{**}, a_*|a_{**}) + \text{IR}(a_*, a_{**}|a_*) \) is strictly increasing in \( \text{Pr}(\theta) \) is non-trivial: If \( \text{IR}(a, a'|a) \) were necessarily increasing in \( \text{Pr}(\theta) \), then it would be trivial that this difference is increasing. What makes Proposition 3.3 non-trivial is the fact that \( \text{IR}(a, a'|a) \) can decrease with an increase in \( \text{Pr}(\theta) \).

The fact that \( \text{IR}(a, a'|a) \) need not be increasing in \( \text{Pr}(\theta) \) leads to non-monotone comparative statics with respect to \( \text{Pr}(\theta) \): Fix parameters of the model so that there are multiple equilibria that differ in their level of effective accountability. If, for any choice of \( a, a' \), \( \text{IR}(a, a'|a) \) were increasing in \( \text{Pr}(\theta) \), then increasing \( \text{Pr}(\theta) \) would necessarily result in a new scenario that also has multiple equilibria which differ in their level of effective accountability. (This follows from Equations 3-4.) But—precisely because \( \text{IR}(a, a'|a) \) need not be increasing in \( \text{Pr}(\theta) \)—the new scenario may very well have a single equilibrium.

**Benefits of Holding Office** An institutional reform can alter the utility of holding office (i.e., \( B \)), by changing the politician’s salary or other material benefits from office. Interestingly, such a reform can break an accountability trap. To see this claim, consider an environment with an accountability trap: In this environment, there are two equilibria that differ in their level of effective accountability, i.e., with different first-period actions \( a_{**} \) and \( a_* \), and Voter welfare is higher in the equilibrium with the higher level of effective accountability. Suppose \( a_{**} \) is the highest action in \( A \). Refer to Equations 3-4. By increasing \( B \), \( a_{**} \) remains an equilibrium. (This uses the fact that, for all \( a \neq a_{**} \), \( \text{IR}(a_{**}, a|a_{**}) > 0 \). See Lemma A.1.) But there exists sufficiently large benefits to holding office so that \( a_* \) is no longer an equilibrium. (This uses the fact that \( \text{IR}(a_*, a_{**}|a_*) < 0 \).

This fact suggests that there may be situations in which small increases in the material benefits associated with holding office have large effects on the level of effective accountability. A small change in \( B \) can affect behavior—not only by causing a small increase in the incremental return to effort—but also by making the “bad expectations” associated with a low level of effective accountability inconsistent with equilibrium.
Connection to Dewatripont, Jewitt and Tirole (1999a,b) In the context of wage-based incentives, Dewatripont, Jewitt and Tirole also provide a multiplicity result. While there are some similarities, there are also important differences between their analysis and ours, which we are now equipped to highlight.

Dewatripont, Jewitt and Tirole’s multiplicity result imposes assumptions on the distribution of types and the production technology. In particular, it assumes the production technology exhibits strict complementarities between the agent’s effort and ability. Propositions 3.1 and 3.2 say that, in our context, multiplicity (with different levels of effective accountability) requires a particular form of complementarity—that the incremental return to high effort must be larger when the Voter expects high effort. But this type of complementarity is generated endogenously through Voter learning. In particular, Proposition 3.3 says that the strict complementarities between the Incumbent’s effort and ability is neither necessary nor sufficient for this type of complementarity: Effort and ability may be strict substitutes and still two actions will be justifiable provided Pr(θ) is sufficiently large. Conversely, effort and ability may be strict complements and, yet, if Pr(θ) is sufficiently low, no two actions will be justifiable.

There is an important difference between the approach that we take versus the approach in Dewatripont, Jewitt and Tirole (1999a). Dewatripont, Jewitt and Tirole begin by fixing a convex cost function as part of the description of the model. Conditional upon observing an outcome, the analysis chooses a benefit; the benefit is constrained to equal the conditional expectation of the agent’s type. In our analysis, we also choose a benefit, but the nature of the choice is quite different: We choose a benefit as part of the description of the model and we do not constrain the benefit to equal the conditional expectation of the agent’s type. Each of the choices fits for the question at hand: In Dewatripont, Jewitt and Tirole (1999a), the benefit is a wage and the choice of wage is a strategic variable. Thus, it makes sense that the wage is chosen after observing variables and is equal to the conditional expectation of the agent’s type. Here, however, the benefit reflects a subjective utility from holding office. Thus, the benefit is prior to the model and there is no reason the benefit ought to equal the conditional expectation of the politician’s type. Likewise, we do not impose a restriction that the chosen cost function must be convex; we allow for a more general class of preferences. For a given production technology and information structure that satisfies the requirements of Propositions 3.1 and 3.2, there may or may not be a benefit of reelection and a convex cost function so that, in the associated game, there are multiple equilibria that differ in their level of effective accountability.
4 Analysis: Welfare

The previous section gave a necessary and sufficient condition for the existence of multiple equilibria that differ in the level of effective accountability, i.e., in their first-period action. To show that an accountability trap is consistent with the canonical agency model of elections, we need one further step: to show that the equilibrium with a higher level of effective accountability (i.e., with a higher first-period action) also has a higher level of \textit{ex ante} Voter welfare. This section provides necessary and sufficient conditions under which this holds. The conditions are consistent with the conditions in Proposition 3.1 and so, taken together, we have necessary and sufficient conditions for an accountability trap.

**Voter Welfare** Fix an equilibrium with first-period action $a_\ast$. We begin by calculating expected Voter welfare at the equilibrium, which we refer to as \textit{Voter welfare at $a_\ast$}. This will be the sum of the Voter’s expected first-period payoffs, viz. $\textit{VW}_1(a_\ast)$, and his \textit{ex ante} expected second-period payoffs, viz. $\textit{VW}_2(a_\ast)$, from the equilibrium.

In particular, fix an equilibrium where the Incumbent’s first-period action is $a_\ast$. The expected first-period Voter welfare is

$$\text{VW}_1(a_\ast) = \Pr(\theta) f(a_\ast, \theta) + (1 - \Pr(\theta)) f(a_\ast, \bar{\theta}).$$

The \textit{ex ante} expected second-period Voter welfare is

$$\text{VW}_2(a_\ast) = \Pr(\theta_{P2} = \theta_{P2} \mid a_\ast) f(a_\ast, \theta) + (1 - \Pr(\theta_{P2} = \theta_{P2} \mid a_\ast)) f(a_\ast, \bar{\theta}),$$

where $\Pr(\theta_{P2} = \theta_{P2} \mid a_\ast)$ is the equilibrium probability that the Politician in office in the second period is type \bar{\theta}.\footnote{The restriction to two types allows us to express $\Pr(\theta_{P2} = \theta \mid a_\ast)$ in terms of primitives of the model. See the Appendix.} Note that $\Pr(\theta_{P2} = \theta_{P2} \mid a_\ast)$ depends on the first-period action chosen by the Incumbent. It also depends on the Voter’s strategy—in particular, the Voter must be using an equilibrium cut-off rule, as specified in Section 3. (But it does not depend on the specification of the Voter’s strategy at the cut-off value.)

**Increasing Effective Accountability** Suppose there are two equilibria: one with first-period action $a_\ast$ and the second with first-period action $a_{\ast\ast} > a_\ast$. The latter equilibrium has a higher level of effective accountability. It also involves a higher expected first-period Voter welfare, i.e., $\text{VW}_1(a_{\ast\ast}) > \text{VW}_1(a_\ast)$; all else equal, increasing the action strictly increases the
level of public goods the Voter receives. But it need not result in a higher ex ante expected second-period Voter welfare. It may be more difficult to select good types in the equilibrium with the first-period action $a_{ss}$. That is, it may be the case that $\Pr(\theta_{P2} = \bar{\theta}|a_s) > \Pr(\theta_{P2} = \bar{\theta}|a_{ss})$ and, if so, it will be the case that $\mathbb{V}W_2(a_s) > \mathbb{V}W_2(a_{ss})$.

**Weak Local Complements**  If increasing the level of effective accountability makes it (weakly) easier to select good types, then increasing effective accountability will increase both first- and second-period Voter welfare. When effort and type are weakly complementary this is exactly what happens. That is, when effort and type are weakly complementary, it is necessarily the case that $\Pr(\theta_{P2} = \bar{\theta}|a_{ss}) \geq \Pr(\theta_{P2} = \bar{\theta}|a_s)$ for $a_{ss} > a_s$.

**Proposition 4.1** Suppose $(B, c)$ justifies both $a_s$ and $a_{ss}$. Then, the following are equivalent:

(i) $\mathbb{V}W_2(a_{ss}) \geq \mathbb{V}W_2(a_s)$.

(ii) $f(a_{ss}, \bar{\theta}) - f(a_{ss}, \theta) \geq f(a_s, \bar{\theta}) - f(a_s, \theta)$.

**Corollary 4.1** Suppose $(B, c)$ justifies both $a_s$ and $a_{ss}$ with $a_{ss} > a_s$. If $f(a_{ss}, \bar{\theta}) - f(a_{ss}, \theta) \geq f(a_s, \bar{\theta}) - f(a_s, \theta)$, then Voter welfare at $a_{ss}$ is strictly higher than Voter welfare at $a_s$.

Suppose there are two equilibria: one with first-period action $a_s$, the second with first-period action $a_{ss}$, and $a_{ss} > a_s$. The corollary says that if, at $a_s$ and $a_{ss}$, effort and type are complementary, then the equilibrium with a higher level of effective accountability must also have a higher level of Voter welfare.

**Strict Local Substitutes** Despite the fact that expected second-period Voter welfare may be higher under the equilibrium with first-period action is $a_s$, it may still be the case that expected (total) Voter welfare is higher under the equilibrium with first-period action $a_{ss}$. This will occur if the first-period welfare gains from the equilibrium with the higher first-period action compensates for the second-period welfare losses from that equilibrium, i.e., if the gains from improved selection are small. The next proposition establishes a necessary and sufficient condition under which the gains from improved selection are small.

**Proposition 4.2** Suppose $(B, c)$ justifies both $a_s$ and $a_{ss}$ with $a_{ss} > a_s$. Then there exists $\underline{p}[a_s, a_{ss}], \bar{p}[a_s, a_{ss}] \in (0, 1)$ so that the following holds: Voter welfare at $a_{ss}$ is strictly higher than Voter welfare at $a_s$ if and only if $\Pr(\bar{\theta}) \in \left(0, \underline{p}[a_s, a_{ss}]\right) \cup \left(\bar{p}[a_s, a_{ss}], 1\right)$. 20
Proposition 4.2 says that, if the prior is close to 0 or 1, the Voter has relatively little to gain from improved selection. In that case, most politicians are of the same type. So, the likelihood that the Politician in office in the second period will be a high type is not very responsive to the first-period outcome. Hence, the first-period welfare effect of high effort dominates any negative second-period welfare effect of high effort.

The bounds \( p[a_0, a_{**}] \) and \( p[a_0, a_\ast] \) can be computed based on primitives of the model. In particular, they depend on (i) the distribution of noise, and (ii) the value of the production function at \( a_0 \), \( a_\ast \), and \( a_{**} \). Thus, the bounds are action-dependent.

**When Higher Effective Accountability is Bad** Refer to the bounds \( p[a_0, a_{**}] \) and \( p[a_0, a_\ast] \) in Proposition 4.2. There is no presumption that \( p[a_0, a_{**}] \leq p[a_0, a_\ast] \). In fact, even if effort and type are strict substitutes at \( a_\ast, a_{**} \), it may well be the case that the first-period welfare gains from choosing the high action over the low action dominate the second-period welfare losses from choosing the high action over the low action. In that case, \( (0, p[a_0, a_{**}]) \cup (p[a_0, a_{**}], 1) = (0, 1) \). That said, there are situations where the second-period welfare losses from choosing the high action over the low action do dominate the first-period welfare gains from choosing the high action over the low action.\(^8\) This is demonstrated with the following example.

**Example 4.1** Let \( A = \{a_0, a_\ast, a_{**}\} \), where \( a_{**} > a_\ast > a_0 \). The production function \( f \) is described by the following table.

<table>
<thead>
<tr>
<th>( a )</th>
<th>0</th>
<th>36</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a_{**} )</td>
<td>38</td>
<td>39</td>
</tr>
<tr>
<td>( a_\ast )</td>
<td>35</td>
<td>37</td>
</tr>
</tbody>
</table>

Figure 4.1: Production Function

Note effort and ability are strict substitutes. Take \( \Phi \) to be the CDF of the standard Normal distribution.

It is readily verified that \( p[a_\ast, a_{**}] < \hat{p}[a_\ast, a_{**}] < .522 < .65 < \overline{p}[a_\ast, a_{**}] \). So, for any \( \Pr(\theta) \in (.53, .65) \), (i) both \( a_\ast \) and \( a_{**} \) are justifiable and (ii) Voter welfare is strictly higher

\(^8\) There is a sense in which this result is the flip of a result from a recent paper by Dewan and Hortala-Vallve (2013). They show that informative campaigns can sometimes reduce voter welfare—increasing electoral selection while decreasing effective accountability.
at $a_*$ than at $a_{**}$. In particular, if $P_r(\bar{y}) = .6$,

$$\VW_1(a_*) + \VW_2(a_*) = .4(35) + .6(37) + 36[.6 + .6(.4)(2\Phi(1) - 1)] > 63.69$$

and

$$\VW_1(a_{**}) + \VW_2(a_{**}) = .4(38) + .6(39) + 36[.6 + .6(.4)(2\Phi(.5) - 1)] < 63.51.$$ 

In sum, when effort and type are strict substitutes, it is easier for the Voter to select good types at the equilibrium with a lower first-period action. This fact leads to an intriguing possibility: Sometimes, increasing effective accountability is actually bad for the Voter—i.e., overall Voter welfare is higher when the first period action is $a_*$ versus $a_{**}$.

## 5 Related Literature

We have introduced a concept of an accountability trap. In an accountability trap, the society finds itself in an equilibrium with a low level of effective accountability relative to some other equilibrium within the same institutional environment. Moreover, the equilibrium with higher effective accountability has a higher level of voter welfare. This section discusses related concepts and other potential sources of accountability traps.

### Other Notions of Accountability Traps

Our use of the phrase “accountability traps” is similar to, but distinct from, the use of the phrase in Landa (2010). Landa uses the term to mean that the society is caught in an equilibrium that has a lower level of voter welfare than some other pair of reelection rule and politician behavior, independent of whether the other reelection rule is optimal for the voter given the politician’s behavior. By contrast, our concept only applies when that other pair forms an equilibrium.

### Other Sources of Accountability Traps

A number of papers in the literature may, ex post, be seen as generating accountability traps. This set of papers has the feature that there are multiple equilibria within a given institutional game. These papers will generate an accountability trap if one of the equilibria has both a higher level of effective accountability and a higher level of voter welfare. We now discuss the extent to which the multiplicities can generate accountability traps.

One natural source of multiplicity is players who have an incentive to coordinate. In the electoral setting, there are two natural reasons why such coordination problems might
arise. First, voters can potentially coordinate or mis-coordinate on providing incentives to politicians. Second, politicians can potentially coordinate or mis-coordinate on working to provide good governance outcomes. Fearon (2011) provides a model that generates accountability traps based on voter coordination problems. Dewan and Myatt (2012) provide a model that generates accountability traps based on coordination problems amongst parliamentary ministers.

In a sense, such coordination problems are not intrinsic to the accountability relationship—they are a feature of voters’ or politicians’ relationships with one another, rather than of voters’ relationship with the politician. To abstract away from such coordination issues, consider the case of a single representative voter and single office holder. Even with these assumptions, there are several models that can be seen as reflecting the possibility of accountability traps.

One important example is a literature on accountability relationships where politicians (i.e., the incumbent and challenger) are identical (Barro, 1973; Ferejohn, 1986; Austen-Smith and Banks, 1989). Consider a version of our model, where politicians are identical. At the time of the election, the voter is indifferent between reelecting the incumbent or replacing her with a challenger. As a result, all voting rules are sequentially rational, so the voter can commit to any rule. This generates the possibility of accountability traps. In particular, there is an equilibrium in which the voter holds the incumbent to a high standard and the incumbent works hard. But there is also an equilibrium—which can be seen as an accountability trap—in which the voter always replaces the incumbent and the incumbent shirks.

In this set of papers, the source of multiple equilibria (and so accountability traps) is the fact that the voter is indifferent between reelecting the incumbent or replacing her with a challenger. Multiplicity can arise from such voter indifference, even when politicians are not ex ante identical. To illustrate this point, consider a two-period version of the models in Banks and Sundaram (1993) and Schwabe (2009). There, candidate heterogeneity reflects different costs of effort on the part of the politician and so does not directly enter voter preferences. Thus, at the election, the voter is indifferent between reelecting or replacing the incumbent, because no politician will take a costly action in the final period.

Accountability traps based on voter indifference are arguably not robust. The familiar argument dates back to Fearon (1999): If politicians differ (even by a tiny amount) in a way that is important for future outcomes, then the voter is no longer indifferent between the incumbent and challenger. Moreover, the natural assumption is that politicians do differ in a way that is important for future voter outcomes. Thus, voter indifference is not a robust
feature of the model and so not a robust source of accountability traps.

Several important subsequent papers deliver equilibrium multiplicity when there is candidate heterogeneity. One approach gives incumbents an inherent advantage. A second approach gives politicians’ incentive to signal. A third approach is based on a long-term (infinite) relationship between the politician and the voter.

Examples of the first approach are Myerson (2006) and Bidner and Francois (Forthcoming). In these papers, the voter has an exogenous benefit for retaining incumbents. With such a preference, there are equilibria similar to equilibria in models with identical politicians. In particular, there is an equilibrium where good types behave well and are reelected, while bad types behave poorly and are replaced. Unlike the case of identical politicians, these equilibria are robust to the presence of a small amount of politician heterogeneity—so long as incumbent heterogeneity is not important relative to the exogenous incumbency advantage. Thus, these models do generate accountability traps. But, the accountability traps depend on the presence of an exogenous incumbency advantage and, so, arguably do not emerge from the accountability relationship itself.

Examples of the second approach are Rogoff (1990), Besley (2006), Frisell (2009) and Gersbach (2009, 2010). In these papers, politicians have heterogenous characteristics: In Rogoff (1990), Besley (2006), and Frisell (2009) the characteristics indirectly enter the voter’s payoffs; in Gersbach (2009, 2010), the characteristics directly enter the voter’s payoffs. Politicians are privately informed about their characteristics and, thus, potentially have an incentive to signal their characteristic to the voter. Of course, such signaling games often have multiple equilibria.

The models in Rogoff and Besley have both pooling and separating equilibria. But the multiplicity in these models should be treated with caution: The pooling equilibrium in each model is inconsistent with Cho and Kreps’s (1987) intuitive criterion. (In fact, for this reason, the authors each dismiss the pooling equilibrium.) Even if we take for granted that Rogoff’s and Besley’s models generate multiplicity, it is not clear that they also generate accountability traps. In the pooling equilibrium in Besley (2006) all types of incumbents take the voter’s most preferred action, while in the separating equilibrium only good types take the voter’s most preferred action. As such, the pooling equilibrium is associated with a higher level of effective accountability. But it is not obvious that the pooling equilibrium also leads to a higher level of voter welfare: While it leads to higher first-period voter welfare, it leads to lower second-period voter welfare. This is because the separating equilibrium allows the voter to select good types. In Rogoff (1990), it is both not clear which equilibrium has a higher level of effective accountability and which equilibrium
is associated with a higher level of voter welfare.

The model in Frisell (2009) has two equilibria. The equilibria are ranked in terms of effective accountability. But, the specification of the model does not permit a voter welfare analysis.

By contrast, the models in Gersbach (2009, 2010) do provide a signaling-based story of accountability traps. In those models, there are both pooling and semi-separating equilibria. The pooling and (some of) the semi-separating equilibria survive the intuitive criterion. Moreover, voter welfare is higher in the semi-separating equilibria. So, the pooling equilibria can be seen as generating accountability traps.

Examples of the third approach are Meirowitz (2007) and Svolik (Forthcoming). They provide a story consistent with accountability traps based on a long-term (infinite) relationship between the politician and the voter. In each of the papers, there is candidate heterogeneity. In Meirowitz (2007), candidate heterogeneity reflects ideological positions which are known to the voter.\(^9\) In Svolik (Forthcoming), candidate heterogeneity reflects whether or not electoral incentives can be effective to get the politician to work. In the finite horizon versions of each of the models, there is a unique equilibrium and, in the unique equilibrium, the voter offers no electoral incentives. The key is that in the last period electoral incentives cannot be effective. But, when there is a long-term (infinite) relationship between the politician and the voter, there is a second equilibrium where the voter does provide electoral incentives in each period, which increases both the level of effective accountability and voter welfare.

We provide an alternate and complementary source of the equilibrium multiplicity. Much like the other papers, our analysis begins with a representative voter, so that the multiplicity (and the existence of accountability traps) will not be generated by the relationship between the voters themselves. Likewise, our analysis has a representative incumbent. We also assume politicians have heterogeneous characteristics—specifically, heterogeneity of skills. This heterogeneity is one natural assumption and is sufficient to ensure that the multiplicity (and so the existence of accountability traps) will not be generated by a fragile voter indifference. We assume that politicians do not know the realization of these characteristics. This is a natural assumption when these characteristics represent skill at tasks where the politician has had little prior experience (e.g., managing fiscal policy or foreign affairs).\(^{10}\) In this case, the multiplicity cannot be generated by politicians with different

\(^9\) The voter does not know which policies are feasible in any given period.

\(^{10}\) By contrast, when politician characteristics represent ideological positions, as in Besley (2006) and Meirowitz (2007), it is natural to assume politicians do know the realization of their characteristics, as Besley and Meirowitz do. Likewise for cost of effort in Banks and Sundaram (1993) and Schwabe (2009).
characteristics potentially signaling “who they are.” Finally, we look at a two-period model. So the multiplicity (and the existence of accountability traps) will not be generated by the nature of long-term interactions. Instead, here, multiplicity arises from a feature that—in our view—is intrinsic to the accountability relationship between a voter and a politician: the fact that the voter has an incentive to select good types in an environment where the politician’s early actions influence the voter’s inference problem.

As we emphasized in the introduction, from the perspective of policy prescription, it is important to have a complete understanding of why accountability traps arise. As we have seen, accountability traps can certainly be driven by ‘common sources’ of equilibrium multiplicity. But, we have shown that they can also arise from an intrinsic feature of the accountability relationship. These different sources require different policy responses.

Alternate Political Traps  Accountability is only one reason that societies with the same set of institutions might differ in their quality of government. They may also differ in terms of the quality of citizens that choose to become candidates (Messner and Polborn, 2004; Caselli and Morelli, 2004; Mattozzi and Merlo, 2008). For instance, in the context of a citizen candidate model, Caselli and Morelli (2004) show that there can be multiple equilibria of a game between citizens who are choosing whether to become candidates. (This multiplicity arises because the presence of high quality politicians has positive spillover effects for other politicians.) So societies can be characterized either by an equilibrium with high quality politicians or by an equilibrium with low quality politicians. Thus, Caselli and Morelli (2004) can be thought of as identifying sufficient conditions for the existence of what we can call “quality traps.”

Note, much as we abstract away from the entry decision, Caselli and Morelli (2004) abstract away from the accountability relationship. A recent literature has combined such citizen-candidate entry decisions with agency problems associated with the accountability relationship. (See, e.g., Besley, 2004; Dal Bó, Dal Bó and Di Tella, 2006; Brollo et al., Forthcoming). But these papers do not address the question of equilibrium multiplicity and so cannot speak to whether accountability traps and quality traps can co-exist. In fact, Proposition 3.3 may suggest a conflict: It says that, if there is an accountability trap, the pool of high quality candidates is ‘large’ (relative to the technology and distribution of noise). This appears to rule out a ‘large’ pool of low quality candidates. Of course, this does not necessarily establish a formal conflict. We leave the co-existence of these traps as an open question.
Appendix

Proofs for Proposition 3.1

We will write \( s^I \) for a strategy of the Incumbent, \( s^V \) for a strategy of the Voter, and \( s^C \) for a strategy of the Challenger. It will be useful to briefly describe these strategies.

A strategy of the Incumbent can be decomposed into a first-period plan, viz. \( s_1^I \), and a second-period plan \( s_2^I \). The first-period plan \( s_1^I \) is simply a first-period action \( a_1 \). Thus, we will talk about a first-period plan being justifiable. A second-period plan \( s_2^I \) maps each first-period level of public goods plus a decision to reelect the Incumbent to a second period action \( a_2 \). Since the Incumbent’s second-period information set can only be reached by a decision to reelect, we suppress reference to the decision and simply write \( s_2^I(g_1) \).

A strategy of the Voter maps each level of public goods provided (in the first period) to a reelection decision. Thus, \( s^V(g_1) = 1 \) \((s^V(g_1) = 0)\) will represent the fact that the Voter reelects (replaces) the Incumbent if \( g_1 \) is the level of public goods provided in the first period. A strategy for the Challenger \( s^C \) maps each level of public goods provided (in the first period) plus a decision to replace the Incumbent to a second period action \( a_2 \). Since the Challenger’s second-period information set can only be reached by a decision to replace, we suppress reference to the decision and simply write \( s^C(g_1) \).

Lemma A.1 Fix a pure-strategy perfect Bayesian equilibrium, viz. \((s^I_*, s^V_*, s^C_*)\), with \( s^I_1 = a_* \). If \( a > a' \) then

(i) \( \Pr(a | a_*) > \Pr(a' | a_*) \),

(ii) \( \text{IR}(a, a' | a_*) > 0 \), and

(iii) \( \text{IR}(a_*, a' | a_*) > \text{IR}(a_*, a | a_*) \).

Proof. The Incumbent is reelected if \( f(a_1, \theta) + \epsilon_1 > \hat{g}(a_*) \). Thus, if the Incumbent chooses first-period action \( a \), the probability of reelection is

\[
\Pr(a | a_*) = \Pr(\bar{\theta}) \left( 1 - \Phi \left( \hat{g}(a_*) - f(a, \bar{\theta}) \right) \right) + (1 - \Pr(\bar{\theta})) \left( 1 - \Phi \left( \hat{g}(a_*) - f(a, \bar{\theta}) \right) \right).
\]

Since \( f \) is increasing in \( a \), \( \Pr(a | a_*) \) is increasing in \( a \), establishing Part (i). Part (ii) follows from Part (i), since \( \text{IR}(a, a' | a_*) = \Pr(a | a_*) - \Pr(a' | a_*) \). Part (iii) follows from Part (i), since \( \text{IR}(a_*, a' | a_*) = \Pr(a_* | a_*) - \Pr(a_* | a_*) - \Pr(a | a_*) - \Pr(a | a_*) = \text{IR}(a_*, a | a_*) \).

Lemma A.2 Fix a first-period strategy \( s^I_1 \) with \( s^I_1 = a_* \). The pair \((B, c)\) justifies \( \{s^I_1\} \) if and only if \( c(a) \geq \text{IR}(a, a_* | a_*) (B - c(g_1)) + c(a_*) \) for each \( a \in A \).
Proof. First suppose that \((B, c)\) justifies \(s^I_1\). Then, for each action \(a \in A\),

\[
\Pr(a_s|a_s)(B - c(a)) + (B - c(a_s)) \geq \Pr(a|a_s)(B - c(a)) + (B - c(a)),
\]

and so \(c(a) \geq \Pr(a|a_s) - \Pr(a_s|a_s)(B - c(a)) + c(a_s)\) for each \(a \in A\).

Conversely, suppose that \(c(a) \geq \text{IR} (a, a_s|a_s)(B - c(a)) + c(a_s)\) for each \(a \in A\). Construct \((s^I_1, s^I_2, s^V, s^C)\) so that

- for each \(g_1, s^I_2(g_1) = s^C(g_1) = a_s\), and
- \(s^V(g_1) = 1\) if and only if \(g_1 \geq \hat{g}(a_s)\).

It is readily verified that \((s^I_1, s^I_2, s^V, s^C)\) is a perfect Bayesian equilibrium of the game. 

Proof of Proposition 3.1. Fix strategies \(s^I_1\) and \(r^I_1\) with \(s^I_1 = a_{**}, r^I_1 = a_s\) and \(a_{**} > a_s\).

Begin with the “only if.” Suppose \((B, c)\) justifies both \(a_{**}\) and \(a_s\). Applying Lemma A.2 to the strategy \(s^I_1\) and action \(a_s\) gives

I. \(c(a_s) \geq \text{IR} (a_{**, a_s}|a_{**})(B - c(a)) + c(a_{**})\), and

Applying Lemma A.2 to the strategy \(r^I_1\) and action \(a_{**}\) gives

II. \(c(a_{**}) \geq \text{IR} (a_{**, a_s}|a_s)(B - c(a)) + c(a_s)\).

Put I and II together. This gives

\[
c(a_s) + \text{IR} (a_{**, a_s}|a_{**})(B - c(a)) \geq c(a_{**}) \geq \text{IR} (a_{**, a_s}|a_s)(B - c(a)) + c(a_s),
\]

where we use the fact that \(\text{IR} (a_{**, a_s}|a_{**}) = -\text{IR} (a_{**, a_s}|a_{**})\). Since \(B > c(a)\) by assumption, we have \(\text{IR} (a_{**, a_s}|a_s) \geq \text{IR} (a_{**, a_s}|a_s)\), as required.

Now we turn to the “if” part. We suppose that \(\text{IR} (a_{**, a_s}|a_{**}) \geq \text{IR} (a_{**, a_s}|a_s)\) and we will show that we can construct a part \((B, c)\) that justifies both \(a_{**}\) and \(a_s\).

To do so, it will be useful to fix certain constants: First choose \(B\) and \(n\) so that \(B > n > 0\). If \(a_s = a\), fix \(n_s = n\). If \(a_s \neq a\), fix \(n_s\) so that:

i. \(n_s > n\);

ii. \(n + \text{IR} (a_s,a_s|a_s)(B - n) > n_s\); and

iii. \(n + [\text{IR} (a_{**, a_s}|a_{**}) - \text{IR} (a_{**, a_s}|a_s)](B - n) > n_s\).
The fact that requirements $i_{**}-ii_{**}$ can be satisfied simultaneously follows from Lemma A.1(ii). The fact that requirements $i_{**}-iii_{**}$ can be satisfied simultaneously follows from $\text{IR}(a_{**}, a|a_{**}) > \text{IR}(a_{**}, a_{**}|a_{**})$ (Lemma A.1(iii)) and $\text{IR}(a_{**}, a|a_{**}) \geq \text{IR}(a_{**}, a_{**}|a_{**})$ (by assumption), so that $\text{IR}(a_{**}, a|a_{**}) > \text{IR}(a_{**}, a_{**}|a_{**})$.

Now fix $n_{**}$

$$i_{**} \quad n_{**} \geq n + \text{IR}(a_{**}, a|a)(B - n);$$

$$ii_{**} \quad n + \text{IR}(a_{**}, a|a)(B - n) \geq n_{**};$$

and

$$iii_{**} \quad n + \text{IR}(a_{**}, a|a)(B - n) \geq n_{**} \text{ with strict inequality if } a \neq a.$$

The fact that requirements $i_{**}-ii_{**}$ can be satisfied simultaneously follows from the assumption that $\text{IR}(a_{**}, a|a) \geq \text{IR}(a_{**}, a_{**}|a_{**})$. Condition $iii_{**}$ follows from condition $ii_{**}$, if $a = a$. The fact that requirements $i_{**}-iii_{**}$ can be satisfied simultaneously when $a > a$ follows from condition $iii_{**}$ above. Note, it follows from Lemma A.1 and $i_{**}$ that $n_{**} > n$.

Construct a function $N : A \times \{a_{**}, a_{**}\} \to \mathbb{R}$ so that

$$N(a, a) = \text{IR}(a, a|a)(B - n) + n$$

and

$$N(a, a_{**}) = \text{IR}(a, a_{**}|a_{**})(B - n) + n_{**}.$$

It follows from Lemma A.1 that $N(\cdot, a)$ and $N(\cdot, a_{**})$ are strictly increasing in $a$. Moreover,

- $n \geq \max\{N(a, a), N(a, a_{**})\}$;
- $n = N(a, a) \geq N(a, a_{**})$;
- $n_{**} = N(a_{**}, a_{**}) \geq N(a_{**}, a_{**})$.

The first of these follows from requirement $ii_{**}$ on $n_{**}$ and requirement $iii_{**}$ on $n_{**}$. The second of these follows from requirement $ii_{**}$ on $n_{**}$. The third of these follows from requirement $i_{**}$ on $n_{**}$.

Now let $\hat{N} : A \to \mathbb{R}$ be the upper envelope of $N(\cdot, a)$ and $N(\cdot, a_{**})$, i.e., $\hat{N}(a) = \max\{N(a, a), N(a, a_{**})\}$ for each $a \in A$. It is strictly increasing. Moreover, it satisfies

- $\hat{N}(a) \leq n$,
- $\hat{N}(a) = n$, and

\footnote{Of course, $n = N(a, a)$ if $a = a$.}
\[ N(a) = n. \]

It follows that we can construct a strictly increasing function \( c : A \to \mathbb{R} \) that lies everywhere above \( \hat{N} \), i.e., for each \( a \in A \), \( c(a) \geq \hat{N}(a) \), with

- \( c(a) = n \),
- \( c(a_*) = n_* \), and
- \( c(a_{**}) = n_{**} \).

Applying Lemma A.2 we get that the \((B, c)\) constructed justifies both \( a_{**} \) and \( a_* \).  

**Proofs for Proposition 3.2**

We now turn to the proof of Proposition 3.2. For mathematical simplicity, throughout this section (i.e., on the proof of Proposition 3.2), we assume \( A \) is finite. We will now introduce preliminary definitions and results.

Recall, the benefit of reelection is a strictly positive number and a cost function \( c : A \to \mathbb{R} \) is a strictly increasing function with \( c(a) \geq 0 \). Thus, the set of all benefits of reelection and cost functions, \( C \), can be viewed as a strict subset of \( \mathbb{R}^{\mid A \mid + 1} \). We endow \( C \) with the relative topology. Let \( C^+ \) be the set of all \((B, c) \in C \) with \( B > c(a) \). We view \( C^+ \) as a topological subset of \( C \) and endow \( C^+ \) with the relative topology.

**Lemma A.3**

(i) The set \( C^+ \) is open in \( C \).

(ii) If \( U \subseteq C^+ \) is open in \( C^+ \), then it is also open in \( C \).

**Proof.** Note, \( C \setminus C^+ \) is closed in \( C \): Fix a sequence \((B^n, c^n) \in C \setminus C^+ \) that converges to \((B, c)\). For each \( n \), \( B^n \leq c^n(a) \). It follows that \( B \leq c(a) \). Thus, \( C \setminus C^+ \) is closed and \( C^+ \) is open.

Now fix some \( U \subseteq C^+ \) that is open in \( C^+ \). There exists some \( V \subseteq C \) open in \( C \) so that \( U = V \cap C^+ \). Since \( C^+ \) is open in \( C \), this says that \( U \) is also open in \( C \).

Fix some subset of first-period Incumbent plans \( S_1^I \). Say \( S_1^I \) is **strictly justifiable** if there is a non-empty \( U \subseteq C^+ \) open in \( C \) so that each \((B, c) \in U \) justifies each element of \( S_1^I \).

**Lemma A.4** Fix first-period strategies \( s_1^I \) and \( r_1^I \) with \( s_1^I = a_* \), \( r_1^I = a_{**} \) and \( a_{**} > a_* \). Suppose, there exists some \((\hat{B}, \hat{c})\) with

- \( \frac{\hat{c}(a) - \hat{c}(a_*)}{\hat{B} - \hat{c}(a)} > \text{IR} (a, a_*) | a_* \) for all \( a \in A \setminus \{a_*\} \) and
\[ \frac{\hat{c}(a) - \hat{c}(a_*)}{B - \hat{c}(a)} > \text{IR} (a, a_* | a_*) \text{ for all } a \in A \setminus \{a_*\}. \]

Then, \( \{s_1^*, r_1^*\} \) is strictly justifiable.

**Proof.** For each \( r' \in \{a_*, a_*\} \) and \( a \in A \setminus \{a'\} \), construct a function \( g[a, a'] : C^+ \to \mathbb{R} \) with \( g[a, a'](B, c) = \frac{c(a) - c(a')}{B - c(a)}/B - c(a) \). This function is continuous. It follows that the sets \( U[a, a'] = (g[a, a'] - (\text{IR} (a, a'|a'), \infty)) \) are each open in \( C^+ \). As such, the sets \( U[a'] = \cap_{a \in A \setminus \{a'\}} U[a, a'] \) are open in \( C^+ \) and so \( U = U[a_1] \cap U[a_*] \) is open in \( C^+ \). By Lemma A.3(ii), \( U \) is also open in \( C \).

By assumption, \( U \neq \emptyset \). For each \( (B, c) \in U \) and each \( a \in A \)

(i) \( c(a) \geq \text{IR} (a, a_* | a_*) (B - c(a)) + c(a_*), \) and

(ii) \( c(a) \geq \text{IR} (a, a_* | a_*) (B - c(a)) + c(a_*). \)

Thus, by Lemma A.2, each \( (B, c) \in U \) justifies both \( s_1^* \) and \( r_1^* \).

**Proof of Proposition 3.2.** Begin by showing that part (i) implies part (ii): To do so, fix strategies \( s_1^* \) and \( r_1^* \) with \( s_1^* = a_*, r_1^* = a_* \), and \( a_* > a_* \). Suppose, contra hypothesis, that \( \{s_1^*, r_1^*\} \) is strictly justifiable but \( \text{IR} (a_* | a_*) \leq \text{IR} (a_*, a_*) \). It follows from Proposition 3.1 that \( \text{IR} (a_* | a_*) = \text{IR} (a_*, a_*) \). Applying Lemma A.2, if \( (B, c) \) justifies \( \{s_1^*, r_1^*\} \), then \( c(a_*) = \text{IR} (a_*, a_* | a_*) (B - c(a)) + c(a_*) \). A consequence is that, for each triple \( B, c(a), c(a_*) \), there is a unique number \( c(a_*) \) so that the pair \( (B, c) \) justifies \( \{s_1^*, r_1^*\} \). As such, there is no non-empty open set of benefits of reelections and cost functions that justify \( \{s_1^*, r_1^*\} \).

Now we show that part (ii) implies part (i): To do so, fix strategies \( s_1^* \) and \( r_1^* \) with \( s_1^* = a_*, r_1^* = a_* \), and \( a_* > a_* \). Also, assume that \( \text{IR} (a_* | a_*) > \text{IR} (a_*, a_*) \).

Repeat the proof of Proposition 3.1: Choose \( B, n \), and \( n_* \) as before. But, now choose \( n_* \) so that

(i) \( n_* > n_0 + \text{IR} (a_* | a_*) (B - n); \)

(ii) \( n_* + \text{IR} (a_* | a_*) (B - n) > n_* ; \) and

(iii) \( n_* + \text{IR} (a_* | a_*) (B - n) > n_* . \)

To see that requirements i–iii can be satisfied simultaneously, use the fact that \( \text{IR} (a_* | a_*) > \text{IR} (a_* | a_*) \). Condition i holds trivially if \( a_* = a_* \). To see that requirements i–iii can be satisfied simultaneously when \( a_* > a_* \), use condition iii in the definition of \( n_* \). As before, it follows from Lemma A.1 and i that \( n_* > n_* \).
Use this new definition of $n_*$ to construct the function $N : A \times \{a_*, a_{**}\} \to \mathbb{R}$. The construction is the same relative to this new choice of $n_*$. With this, once again, $\hat{N}(a) = \max\{N(a,a_*) , N(a,a_{**})\}$ for each $a \in A$. The function exhibits the same properties as it previously did. It follows that we can construct a strictly increasing cost function $c : A \to \mathbb{R}$ with that lies everywhere strictly above $\hat{N}$, i.e., for each $a \in A$, $c(a) \geq \hat{N}(a)$, with

- $c(a) = n$,
- $c(a_*) = n_*$,
- $c(a_{**}) = n_{**}$, and
- $c(a) > \hat{N}(a)$ for each $a \in A \setminus \{a, a_*, a_{**}\}$.

It now follows from Lemma A.4 that $\{s_I^*, r_I^*\}$ is strictly justifiable.

**Proof of Proposition 3.3**

This appendix begins by demonstrating that the necessary and sufficient condition given in Proposition 3.1 is, in fact, a condition that can be stated only in terms of primitives of the model—specifically, it can be stated only in terms of requirements on the production technology and beliefs. We then turn to prove Proposition 3.3.

The following lemma will be useful.

**Lemma A.5** \( \hat{g}(a) = f(a, \theta) + f(a, \bar{\theta}) \).

**Proof.** From Equation 2, we have:

\[ \phi(\hat{g}(a) - f(a, \bar{\theta})) = \phi(\hat{g}(a) - f(a, \theta)) \]

Since the PDF $\phi$ is symmetric and single-peaked, there are two possible ways this can hold: (i) \( \hat{g}(a) - f(a, \bar{\theta}) = \hat{g}(a) - f(a, \theta) \) or (ii) \( [\hat{g}(a) - f(a, \bar{\theta})] = -[\hat{g}(a) - f(a, \theta)] \). If case (i) held, this would imply \( f(a, \bar{\theta}) = f(a, \theta) \), which contradicts the fact that $f$ is strictly increasing in $\theta$. Hence, case (ii) must hold. Rearranging, this implies \( \hat{g}(a) = \frac{f(a, \bar{\theta}) + f(a, \theta)}{2} \), as required. 

Define numbers

\[ X \equiv \left[ \Phi \left( \hat{g}(a_{**}) - f(a_*, \bar{\theta}) \right) - \Phi \left( \hat{g}(a_*) - f(a_*, \bar{\theta}) \right) \right] - \left[ \Phi \left( \hat{g}(a_{**}) - f(a_{**}, \bar{\theta}) \right) - \Phi \left( \hat{g}(a_{**}) - f(a_{**}, \bar{\theta}) \right) \right] \]

and

\[ Y \equiv \left[ \Phi \left( \hat{g}(a_{**}) - f(a_{**}, \theta) \right) - \Phi \left( \hat{g}(a_*) - f(a_*, \theta) \right) \right] - \left[ \Phi \left( \hat{g}(a_{**}) - f(a_{**}, \theta) \right) - \Phi \left( \hat{g}(a_{**}) - f(a_{**}, \theta) \right) \right] . \]
In light of Lemma A.5, \( X \) and \( Y \) are defined in terms of primitives of the model. Specifically, they depend only on the production function \( f \) and beliefs about shocks.

The next Proposition shows that the condition of Proposition 3.1 can be restated in terms of \( X, Y \), and beliefs about \( \theta \). To show this, we use the fact that

\[
\Pr(a'|a) = \Pr(\overline{\theta}) \left[ 1 - \Phi(\hat{g}(a) - f(a', \overline{\theta})) \right] + (1 - \Pr(\overline{\theta})) \left[ 1 - \Phi(\hat{g}(a) - f(a', \overline{\theta})) \right],
\]

since \( \Pr(a'|a) \) is the probability that \( f(a', \cdot) + \varepsilon \) meets the benchmark \( \hat{g}(a) \). Using this, we have the following:

**Lemma A.6** \( \text{IR}(a_{**}, a_*|a_{**}) \geq \text{IR}(a_{**}, a_*|a_*) \) if and only if \( \Pr(\overline{\theta})X \geq (1 - \Pr(\overline{\theta}))Y \).

An immediate consequence of Proposition 3.1 and Lemma A.6 is the following:

**Proposition A.1** Fix an action \( a_{**} > a_* \). The actions \( a_{**} \) and \( a_* \) are justifiable if and only if \( \Pr(\overline{\theta})X \geq (1 - \Pr(\overline{\theta}))Y \).

In what follows, we provide conditions on the production technology which imply that \( \Pr(\overline{\theta})X \geq (1 - \Pr(\overline{\theta}))Y \) when \( \Pr(\overline{\theta}) \) is sufficiently high. To do so, it will be useful to have two properties of the CDF \( \Phi \).

**Lemma A.7**

(i) \( \Phi(x) = 1 - \Phi(-x) \).

(ii) If \( y, z > 0 \) and \( x \in (-y - z, y) \), then \( \Phi(x + z) - \Phi(x) > \Phi(y + z) - \Phi(y) \).

**Proof.** Part (i) follows symmetry, since

\[
1 - \Phi(-x) = 1 - \int_{-\infty}^{-x} \phi(q) dq = 1 - \int_{x}^{\infty} \phi(q) dq = \int_{-\infty}^{x} \phi(q) dq = \Phi(x).
\]

For part (ii), fix \( z > 0 \) and note

\[
\Phi(x + z) - \Phi(x) = \int_{x}^{x+z} \phi(q) dq
\]

By single-peakedness, \( \phi \) is strictly increasing on \((-\infty, 0)\) and strictly decreasing on \((0, \infty)\). Thus, if \( x \in [0, y) \), it is immediate that \( \Phi(x + z) - \Phi(x) \geq \Phi(y + z) - \Phi(y) \). If \( x \in (-y - z, 0) \), then

\[
\Phi(x + z) - \Phi(x) \geq \Phi(-y - z + z) - \Phi(-y - z) = \Phi(y + z) - \Phi(y),
\]

33
where the equality follows from the Part (i) of this Lemma.

In what follows, we will fix a production function $f$ and actions $a_{**} > a_*$. It will be convenient to adopt the notation (for the output of production) described in Figure A.1. So, $\rho_* = f(a_* \theta) - f(a_*, \theta) > 0$, $\rho_{**} = f(a_{**} \theta) - f(a_{**}, \theta) > 0$, and $\psi_{**} = f(a_{**} \theta) - f(a_*, \theta) > 0$. The next Lemma is immediate from monotonicity of the production function.

<table>
<thead>
<tr>
<th>$a_{**}$</th>
<th>$\psi_{**}$</th>
<th>$\psi_{<strong>} + \rho_{</strong>}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_*$</td>
<td>$\psi_*$</td>
<td>$\psi_<em>$ + $\rho_</em>$</td>
</tr>
</tbody>
</table>

Figure A.1: The Function $f$

**Lemma A.8** $\psi_{**} + \rho_{**} - \rho_* > 0$.

**Proof.** Note that $f(a_{**}, \theta) - f(a_*, \theta) = \psi_{**} + \rho_{**} - \rho_*$, from which the claim follows.

**Lemma A.9** $X > 0$.

**Proof.** Applying Lemma A.7(i), $X$ can be written as

$$X = \left[ \Phi\left(\frac{1}{2} \rho_{**}\right) - \Phi\left(\rho_* - \psi_{**} - \frac{1}{2} \rho_{**}\right) \right] - \left[ \Phi\left(\rho_{**} + \psi_{**} - \frac{1}{2} \rho_{**}\right) - \Phi\left(\frac{1}{2} \rho_*\right) \right].$$

Write $d = \rho_{**} + \psi_{**} - \rho_*$ and note that, by Lemma A.8, $d > 0$. We can then rewrite $X$ as

$$X = \left[ \Phi\left(\rho_* - \psi_{**} - \frac{1}{2} \rho_{**} + d\right) - \Phi\left(\rho_* - \psi_{**} - \frac{1}{2} \rho_{**}\right) \right] - \left[ \Phi\left(\frac{1}{2} \rho_* + d\right) - \Phi\left(\frac{1}{2} \rho_*\right) \right].$$

Thus, by Lemma A.7(ii), $X > 0$ provided (i) $\rho_* - \psi_{**} - \frac{1}{2} \rho_{**} > -\frac{1}{2} \rho_* - d$ and (ii) $\frac{1}{2} \rho_* > \rho_* - \psi_{**} - \frac{1}{2} \rho_{**}$. Condition (i) is immediate and Condition (ii) is by Lemma A.8.

**Lemma A.10** $Y > 0$.

**Proof.** Note, $Y$ can be written as

$$Y = \left[ \Phi\left(\frac{1}{2} \rho_*\right) - \Phi\left(\frac{1}{2} \rho_* - \psi_{**}\right) \right] - \left[ \Phi\left(\frac{1}{2} \rho_{**} + \psi_{**}\right) - \Phi\left(\frac{1}{2} \rho_{**}\right) \right].$$

By Lemma A.8, $\frac{1}{2} \rho_{**} > \frac{1}{2} \rho_* - \psi_{**}$. Since $\psi_{**} > 0$ and $\frac{1}{2} \rho_* - \psi_{**} > -\frac{1}{2} \rho_{**} - \psi_{**}$, we can apply Lemma A.7(ii) to conclude that $Y > 0$. 

34
Lemma A.11 \( \Pr(\bar{\theta})X \geq (1 - \Pr(\bar{\theta}))Y \) is strictly increasing in \( \Pr(\bar{\theta}) \)


The following is immediate from Lemmata A.6 and A.11.

Remark A.1 \( \text{IR}(a_{**}, a_{*}|a_{**}) - \text{IR}(a_{**}, a_{*}|a_{**}) \) is strictly increasing in \( \Pr(\bar{\theta}) \).

Proof of Proposition 3.3. Take \( \hat{p}[a_{*}, a_{**}] = \frac{Y}{Y+X} \). By Lemmata A.9-A.10, \( \hat{p}[a_{*}, a_{**}] \in (0,1) \). Now the result follows from Proposition A.1. ■

Say that **effort and type are strict complements at** \( a_{*}, a_{**} \) if \( f(a_{**}, \bar{\theta}) - f(a_{**}, \theta) > f(a_{*}, \bar{\theta}) - f(a_{*}, \theta) \) or, equivalently, if \( \rho_{**} > \rho_{*} \). Say that **effort and type are strict substitutes at** \( a_{*}, a_{**} \) if \( f(a_{**}, \bar{\theta}) - f(a_{**}, \theta) > f(a_{*}, \bar{\theta}) - f(a_{*}, \theta) \) or, equivalently, if \( \rho_{*} > \rho_{**} \).

Proof of Proposition 3.4. Take \( \hat{p}[a_{*}, a_{**}] = \frac{Y}{Y+X} \). To show the Proposition it suffices to show (i) if effort and type are strict complements at \( a_{*}, a_{**} \), then \( Y - X < 0 \), (ii) if effort and type are neither strict complements nor strict substitutes at \( a_{*}, a_{**} \), then \( X = Y \), and (iii) if effort and type are strict substitutes at \( a_{*}, a_{**} \), then \( X - Y < 0 \). In what follows we set \( d = \frac{1}{2}(\rho_{**} - \rho_{*}) \). Then \( d > 0 \) (resp. \( -d > 0 \)) if effort and type are strict complements (resp. substitutes) at \( a_{*}, a_{**} \).

First, suppose that effort and type are strict complements at \( a_{*}, a_{**} \). Note,

\[
Y - X = \left[ \Phi \left( \frac{1}{2} \rho_{**} + \psi_{**} + d \right) - \Phi \left( \frac{1}{2} \rho_{**} + \psi_{**} \right) \right] - \left[ \Phi \left( \rho_{*} - \psi_{**} - \frac{1}{2} \rho_{**} + d \right) - \Phi \left( \rho_{*} - \psi_{**} - \frac{1}{2} \rho_{**} \right) \right].
\]

Since effort and type are strict complements at \( a_{*}, a_{**}, d > 0 \). Observe that \( \frac{1}{2} \rho_{**} + \psi_{**} > 0 \) and \( \rho_{*} - \psi_{**} - \frac{1}{2} \rho_{**} > -(\frac{1}{2} \rho_{**} + \psi_{**} + d) = -\rho_{**} - \psi_{**} + \frac{1}{2} \rho_{*} \). So, by Lemma A.7(ii), \( Y - X < 0 \).

Next, suppose that effort and type are neither strict substitutes nor strict complements at \( a_{*}, a_{**} \). Then, using the fact that \( \rho_{*} = \rho_{**} \), it is immediate that \( X = Y \).

Finally, suppose that effort and type are strict substitutes at \( a_{*}, a_{**} \). Note

\[
X - Y = \left[ \Phi \left( \rho_{**} + \psi_{**} - \frac{1}{2} \rho_{*} - d \right) - \Phi \left( \rho_{**} + \psi_{**} - \frac{1}{2} \rho_{*} \right) \right] - \left[ \Phi \left( \frac{1}{2} \rho_{*} - \psi_{**} - d \right) - \Phi \left( \frac{1}{2} \rho_{*} - \psi_{**} \right) \right].
\]

Since effort and type are strict substitutes at \( a_{*}, a_{**}, \) \( d < 0 \). By Lemma A.8, \( \rho_{**} + \psi_{**} - \frac{1}{2} \rho_{*} > 0 \). Moreover, \( \frac{1}{2} \rho_{*} - \psi_{**} > -(\rho_{**} + \psi_{**} - \frac{1}{2} \rho_{*} - d) = -\frac{1}{2} \rho_{**} - \psi_{**} \). So, by Lemma A.7(ii), \( X - Y < 0 \). ■
Proofs for Section 4

Fix an equilibrium where the action chosen in the first-period is $a^*$. Write $\Pr(\theta_2 = \theta | a^*)$ for the probability that the Politician in office in the second period is type $\theta$. Note, the event “the Politician in office in the second period is type $\theta$” can be realized in one of three ways: (i) the first-period Incumbent is type $\theta$ and is reelected, (ii) the first-period Incumbent and the Challenger are both type $\theta$ and the first-period Incumbent is not reelected, or (iii) the first-period Incumbent is type $\theta$, she is not reelected, and the Challenger is type $\theta$. Thus,

$$
\Pr(\theta_2 = \theta | a^*) = \Pr(\theta) \Pr(f(a^*, \theta) + \epsilon_1 \geq \hat{g}(a^*)) + 
\Pr(\theta)^2 \Pr(f(a^*, \theta) + \epsilon_1 \leq \hat{g}(a^*)) + (1 - \Pr(\theta)) \Pr(f(a^*, \theta) + \epsilon_1 \leq \hat{g}(a^*)).
$$

(This implicitly use the fact that the random variable generating $\epsilon_1$ is atomless, i.e., so that $\Pr(f(a^*, \theta) + \epsilon_1 = \hat{g}(a^*)) = 0$.) Using the fact that $\Pr(f(a^*, \theta) + \epsilon_1 \geq \hat{g}(a^*)) = 1 - \Phi(\hat{g}(a^*) - f(a^*, \theta))$,

$$
\Pr(\theta_2 = \theta | a^*) = \Pr(\theta) + (1 - \Pr(\theta)) \Pr(\theta)[2\Phi \left( \frac{f(a^*, \theta) - f(a^*, \theta)}{2} \right) - 1].
$$

The next lemma is an immediate consequence of this calculation.

**Lemma A.12** The following are equivalent:

(i) $\Pr(\theta_2 = \theta | a^*) \geq \Pr(\theta_2 = \theta | a^*)$.

(ii) $f(a^*, \theta) - f(a^*, \theta) \geq f(a^*, \theta) - f(a^*, \theta)$.

**Proof of Proposition 4.1.** This is immediate from Lemma A.12 and the fact that, for each action $a$, $VW_2[a] = \Pr(\theta_2 = \theta | a)[f(a, \theta) - f(a, \theta)] + f(a, \theta)$. ■

Fix some $a \in A$. It will be convenient to introduce a function $VW_1[a] : [0, 1] \to \mathbb{R}$ and $VW_2[a] : [0, 1] \to \mathbb{R}$ so that

$$
VW_1[a](p) = [f(a, \theta) - f(a, \theta)]p + f(a, \theta)
$$

and

$$
VW_2[a](p) = (f(a, \theta) - f(a, \theta)) \left( p + (1 - p)p \left( 2\Phi \left( \frac{f(a, \theta) - f(a, \theta)}{2} \right) - 1 \right) \right) + f(a, \theta).
$$

Then, define $VW[a] : [0, 1] \to \mathbb{R}$ so that $VW[a] = VW_1[a] + VW_2[a]$. Note, if there is an equilibrium where $a$ is the first-period action and $p \in (0, 1)$ is $p = \Pr(\theta)$, then $VW_1[a](p)$
is first-period expected Voter welfare, \( \text{VW}_2[a](p) \) is second-period \textit{ex ante} expected Voter welfare and \( \text{VW}[a](p) \) is expected Voter welfare.

**Proof of Proposition 4.2.** It will be convenient to adopt the notation for the production function described in Figure A.1. Suppose \((B,c)\) justifies \(a^{**}\) and \(a^*\), where \(a^{**} > a^*\). Consider the difference (in welfare) function \( \Delta[a^{**},a^*] = \text{VW}[a^{**}] - \text{VW}[a^*] \). Note, this is a continuous function with \( \Delta[a^{**},a^*](0) = \psi^{**} > 0 \) and (by Lemma A.8) \( \Delta[a^{**},a^*](1) = \psi^{**} + \rho^{**} - \rho^* > 0 \). Moreover, the derivative of the difference function with respect to \( p \) is

\[
(p^{**} - \rho^*) + 2[f(a,\bar{\theta}) - f(a,\theta)] \left[ \Phi\left(\frac{\rho^{**}}{2}\right) - \Phi\left(\frac{\rho^*}{2}\right) \right] (1 - 2p).
\]

Thus, if effort and type are strict substitutes (respectively, strict complements) at \(a^*,a^{**}\), the difference function is strictly decreasing (respectively, strictly increasing) on \((0,p_{\min}) \neq \emptyset\) (respectively, \((0,p_{\max}) \neq \emptyset\) and strictly increasing (respectively, strictly decreasing) on \((p_{\min},1) \neq \emptyset\) (respectively, \((p_{\max},1) \neq \emptyset\)). And, if effort and type are neither strict substitutes nor strict complements at \(a^*,a^{**}\), this is a constant function. It follows that there exists \( \hat{p}[a^*,a^{**}],\overbar{p}[a^*,a^{**}] \in (0,1) \) so that \( p \in (0,\hat{p}[a^*,a^{**}]) \cup (\overbar{p}[a^*,a^{**}],1) \) if and only if \( \Delta[a^{**},a^*](p) \geq 0 \). ■

**References**


