Accountability and Information in Elections*

Scott Ashworth†   Ethan Bueno de Mesquita‡   Amanda Friedenberg§

April 22, 2015

Abstract

Elections are thought to improve voter welfare through two channels: effective accountability (i.e., providing incentives for politicians to take costly actions) and electoral selection (i.e., retaining politicians with characteristics voters value). We study the relationship between effective accountability and electoral selection in a canonical career concerns model of elections. Increasing effective accountability directly impacts the informativeness of the voter’s signal about the politician’s type. We show that, if politicians’ actions and type are local substitutes (resp. complements) in the production of governance outcomes, an increase in effective accountability corresponds to a decrease (resp. increase) in Blackwell (1951) informativeness. Consequently, in the case of local substitutes, higher effective accountability can be associated with lower voter welfare. The fact that effective accountability impacts informativeness has counter-intuitive implications for voter behavior—increased effective accountability can be associated with the voter using a less stringent performance standard for reelection. We conclude by showing that institutional variation is not a prerequisite for generating variation in effective accountability. In particular, we provide necessary and sufficient conditions for there to be multiple equilibria that differ in terms of both effective accountability and electoral selection. This condition is consistent with any production environment.

*We have benefited from conversations with John Duggan, Rafael Hortala-Vallve, Ian Jewitt, Dimitri Landa, Adam Meironitz, David Myatt, Andrew Rehfeld, Rainer Schwabe, Dan Silverman and from seminar participants at University of Arizona, Arizona State, Berkeley, Chicago, Maryland, Rochester, Virginia, EPA, MPSA, SAET, and the World Congress of the Game Theory Society. We thank Alex Danieli for research assistance. The paper has been much improved due to comments by the Associate Editor and Referees. The research was supported by NSF grants SES-0819152 (Bueno de Mesquita) and SES-0819200 (Friedenberg). Friedenberg thanks New York University, University College London, Caltech, and Northwestern University for their hospitality. Earlier versions of this paper circulated under the titles: “Accountability Traps,” “Creating Incentives and Selecting Good Types Revisited,” and “Sanctioning and Learning in Elections.”

†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—i.e., the formal right to retain or replace policy makers—will lead to policy outcomes that are better for citizen welfare.

Formal accountability is thought to be linked to welfare improvements for two reasons. (See Fearon, 1999.) First, when voters have the formal right to retain or replace policy makers, the politicians’ actions are shaped by how they anticipate the voters will respond. That is, formal accountability creates what we will call effective accountability. Second, voters have the ability to retain politicians whose characteristics are best aligned with their own interests. That is, formal accountability creates electoral selection.

We show that there may be a trade-off between effective accountability and electoral selection. Higher levels of effective accountability may hinder the ability to learn about the characteristics of politicians. In turn, this may hinder electoral selection and be detrimental to voter welfare.

It is, of course, known that this trade-off can arise in an environment where an incumbent has an incentive to signal her type. Specifically, in signaling models, the quality of electoral selection depends on the correlation between an incumbent’s actions and her type: A separating equilibrium involves a higher level of electoral selection than a pooling equilibrium. But, there may be a pooling equilibrium in which all incumbents choose the voter’s most preferred action. (See Maskin and Tirole, 2004; Besley, 2006.)

We show that there may be a trade-off between effective accountability and electoral selection, even in a setting where politicians cannot signal their type. We address the trade-off in a signal-jamming model, in the spirit of Holmström’s (1999) career concerns model. (A long tradition in political economy uses such models to study the agency relationship between a voter and a politician. See, e.g., Lohmann, 1998; Persson and Tabellini, 2000; Alesina and Tabellini, 2007, 2008.) In these models, politicians do not know their types and so all types must choose the same action. The trade-off between effective accountability and electoral selection arises despite the fact that there cannot be correlation between action and type.

What then generates a relationship between effective accountability and electoral selection? Governance outcomes serve as a signal of the incumbent’s characteristics. The signals’ informativeness depends on the level of effective accountability. That is, different levels of effective accountability generate different statistical experiments from which the voter learns about the incumbent’s characteristics. This is true, despite the fact that all types must choose the same action.

We show that, when the politicians’ actions and characteristics are local substitutes (resp. complements) in the production of governance outcomes, an increase in effective accountability corresponds to a decrease (resp. increase) in Blackwell (1951) informativeness. Consequently, there is a trade-off between electoral selection and effective accountability if and only if the politician’s actions and characteristics are local substitutes (in the production of governance outcomes). So, in the case of complements, higher levels of effective accountability improve voter welfare. But, in the case of substitutes, the voter may be worse off with a higher level of effective accountability and a lower level of electoral selection.
The informational consequences of increasing effective accountability have counter-intuitive implications for voter behavior. An increase in effective accountability induces an improvement in the distribution of governance outcomes. As a result, when effective accountability is higher, the voter requires higher levels of public goods to infer neutral news about the incumbent’s characteristics. One might conjecture that when effective accountability is higher, the voter will hold the incumbent to a more stringent performance standard. But this need not be the case. Changes in effective accountability also alter the informativeness of outcomes, which creates a potentially countervailing effect. Whether this informational effect offsets the direct effect of increasing effective accountability depends on (i) whether actions and characteristics are local complements or substitutes and (ii) the pools of high-ability incumbents and challengers. (We later explain why.) We show that an increase in effective accountability may correspond to a less stringent performance standard.

In sum, different levels of effective accountability generate different statistical experiments from which the voter draws inferences about the incumbent’s characteristics. In turn, this generates a trade-off between effective accountability and electoral selection. The trade-off rests on the presumption that it is possible to vary the level of effective accountability—an endogenous variable—and, thus, the statistical experiment.

The standard approach is to think of variation in effective accountability as a result of variation in the institutions of formal accountability. That is, varying term length, compensation, campaign finance rules, staff resources, etc., may lead to variation in effective accountability. (See, e.g., Persson and Tabellini, 2000; Maskin and Tirole, 2004; Besley, 2006; Ashworth and Bueno de Mesquita, 2006; Gehlbach, 2007; Myerson, 2006 for theoretical studies and Besley and Case, 1995, 2003; Persson and Tabellini, 2000; Huber and Gordon, 2004; Besley, 2004; Alt, Bueno de Mesquita and Rose, 2011; de Janvry, Finan and Sadoulet, 2012; Gagliarducci and Nannicini, 2013; Finan and Ferraz, 2011 for empirical studies.) In the context of our model, varying institutions of formal accountability will serve to vary the preferences of politicians, i.e., the benefits of holding office and the costliness of actions.

But, importantly, we will also see that varying institutions is not a prerequisite for varying the level of effective accountability (and electoral selection). That is, within a fixed set of institutions, there may be variation in the behavior of politicians and voters. We provide necessary and sufficient conditions for there to be multiple equilibria that differ in terms of both their level of effective accountability and in electoral selection. We go on to show that any production environment is consistent with multiplicity: For any production environment, there are preferences of the incumbent which give rise to multiple equilibria that differ in their level of effective accountability.

In our model, equilibrium multiplicity is 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 raises the possibility that different societies—with identical institutions—might differ in both their level of effective accountability and electoral selection.

The fact that there can be multiple equilibria in a career concerns setting is known from Dewa-
tripont, Jewitt and Tirole (1999b) (henceforth, DJT). In their setting, the agent’s early behavior influences the principal’s ability to infer the agent’s ability. In the context of a multiplicative production environment (where the agent’s effort and type are strict complements), they show that this can give rise to multiple equilibria that differ in their level of first-period effort. They then give an example of an additive production function and strictly convex cost function; in the context of that example, they show uniqueness.\(^1\) While our multiplicity result does not insist that politicians have convex cost functions, it does not preclude it. In particular, we will give an example of an additive production and strictly convex cost function, in which multiplicity does arise. At first glance, these two examples may appear to contradict the message in DJT. But we will show that their different conclusions rest on a particular modeling assumption—specifically, the richness of the action sets. Section 5 will further discuss the relationship to DJT.

In Holmström’s (1999) seminal career concerns paper, the principal draws an inference about the agent’s type from a statistical experiment. This work has raised a broader question about how this inference is affected by additional observed variables. (Dewatripont, Jewitt and Tirole, 1999a is one example, amongst many.) These additional variables exogenously change the informativeness of the principal’s signal. To the best of our knowledge, the literature has not explored how variations in the endogenous variable, specifically effort, change the informativeness of the signal.

There is a growing literature that explores how changes in an endogenous variable change the informativeness of an experiment. (See Prendergast and Stole, 1996 and Deb and Stewart, 2015.) Each of these papers explores the question in an environment where the agent who chooses the endogenous variable can signal her own type. By contrast, in our paper, \textit{ex ante} the politicians and voter have symmetric uncertainty about the politician’s type.

The paper proceeds as follows. Section 1 presents a canonical model of electoral accountability. Section 2 describes the basic properties of equilibrium. Section 3 characterizes when there is and is not a trade-off between effective accountability and electoral selection and shows that higher levels of effective accountability do not necessarily correspond to higher voter welfare. Section 4 describes implications for voter behavior. Section 5 shows that variation in electoral behavior need not derive from institutional variation by providing necessary and sufficient conditions for the existence of multiple equilibria that differ in their level of effective accountability and electoral selection. Section 6 discusses the normative implications of our results.

\section{The Model}

There is an Incumbent (I), a Challenger (C), and a Voter (V). We refer to each Politician (P) as “she” and the Voter as “he.” In each of two periods, the Voter receives a level of public goods. This level is a function of the action chosen by the Politician in office, the type of the Politician in

\footnote{Dewatripont, Jewitt and Tirole (1999a, p. 187) say: “Uniqueness of equilibrium does however rely on the fact that effort and talent enter additively in the determination of output: as stressed in our companion paper (Dewatripont \textit{et al.} (1999[b])), complementarity between effort and talent can lead to multiple equilibria.” The literature has, at times, taken this to mean that complementarities are required for multiplicity.}
office, and an idiosyncratic shock.

The Politician in office chooses an action in \( A \subseteq \mathbb{R}_+ \). The set \( A \) is closed, with \( a \) its smallest element. Higher actions should be thought of as higher levels of effort. The set of types is \( \Theta = \{ \vartheta, \overline{\vartheta} \} \), where \( \vartheta \geq 0 \) is the low type and \( \overline{\vartheta} > \vartheta \) is the high type. Write \( \pi_P \in (0, 1) \) for the probability that Politician \( P \) is type \( \overline{\vartheta} \). These probabilities are commonly understood by the players.

The level of public goods produced in a period is a function of a production technology and a random shock. The production function \( f : A \times \Theta \to \mathbb{R}_+ \) is strictly increasing in action \( a \) and type \( \vartheta \). We will be interested in complementarity and substitutability between effort and type.

**Definition 1.1.** Fix actions \( a_{ss} > a_s \).

(i) Effort and type are *complements* at \( a_s, a_{ss} \) if \( f(a_{ss}, \overline{\vartheta}) - f(a_{ss}, \vartheta) \geq f(a_s, \overline{\vartheta}) - f(a_s, \vartheta) \).

(ii) Effort and type are *substitutes* at \( a_s, a_{ss} \) if \( f(a_s, \overline{\vartheta}) - f(a_s, \vartheta) \geq f(a_{ss}, \overline{\vartheta}) - f(a_{ss}, \vartheta) \).

These definitions of complementarity vs. substitutability are local; effort and type may be complements at two actions, but substitutes at other actions. If effort and type are complements (resp. substitutes) at \( a_s, a_{ss} \) but not substitutes (resp. complements) at \( a_s, a_{ss} \), we say that effort and type are *strict complements* (resp. *strict substitutes*) at \( a_s, a_{ss} \).

If, in period \( t \), the Politician in office chooses action \( a \), is of type \( \vartheta \), and the random shock is \( \epsilon_t \), the level of public goods produced in that period is \( f(a, \vartheta) + \epsilon_t \). Each \( \epsilon_t \) is the realization of a random variable. These random variables are independent of each other and are independent of the Politicians’ abilities. In particular, each of these random variables is distributed according to an absolutely continuous CDF, \( \Phi \), with a continuously differentiable PDF, \( \phi \). This distribution satisfies two additional requirements: First, for each \( x > x' \geq 0 \), the associated likelihood ratio defined by

\[
\phi(g - x) \phi(g - x')
\]

is onto with non-zero derivatives. Second, the distribution satisfies the (strict) monotone likelihood ratio property (MLRP) relative to all possible realizations of production: If \( x > x' \geq 0 \), then the associated likelihood ratio

\[
\frac{\phi(g - x)}{\phi(g - x')}
\]

is strictly increasing in \( g \). At times, we will be interested in specific examples of this framework.

**Specific Model Assumptions.** Call the PDF symmetric if, for each \( x \in \mathbb{R} \), \( \phi(x) = \phi(-x) \). If \( \phi \) is symmetric, the MLRP implies that \( \phi \) is single-peaked about zero, i.e., \( \phi \) is strictly increasing on \( (-\infty, 0) \) and strictly decreasing on \( (0, \infty) \). (See Torgersen, 1991, Theorem 9.4.9). Call the model symmetric if the density is symmetric and \( \pi_I = \pi_C \).

\(^2\)Observe that this is a requirement that is independent of the particular production function—put differently, it is a requirement that can be satisfied by any production function. Lemma B.1 states that the distribution satisfies strict MLRP and the likelihood ratios have non-zero derivatives and if and only log \( \phi \) is strictly concave. (Weak versions of this equivalence are standard. See, e.g., Torgersen, 1991, Theorem 9.4.9.)
Prior to the game being played, Nature determines the realizations of each Politician’s type and of the random shocks (in all periods). These realizations are not observed by any of the players. Figure 1.1 depicts the timeline: In the initial governance period, the Incumbent chooses an action $a_1$. The choice is not observed by the Voter. Instead, the Voter observes the level of public goods produced, $g_1$. This leads to the electoral stage, in which the Voter chooses to reelect the Incumbent or replace her with a Challenger. The winner of the election is the Politician in office in the second governance period. She chooses an action $a_2$. Again, the Voter observes the level of public goods produced, $g_2$.

![Timeline](image)

The Voter’s payoffs are the sum of public goods produced in the two periods. Each Politicians’ 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 and chooses action $a_t$. A Politician’s payoffs are given by the sum of her payoffs in each governance period.

We will be interested in understanding how changing the Politicians’ benefits of reelection and cost function affects the Voter’s ability to acquire information and the Voter’s welfare. With this in mind, it will be convenient to think of the model as parameterized by $B$ and $c$. Write $M(B, c)$ for the model in which the benefits of reelection and cost functions take on the values $B$ and $c$.

## 2 Properties of Equilibrium

In this section, we fix a model $M(B, c)$ and point to basic properties of equilibrium, which will be useful in our subsequent analysis. We focus on pure strategy Perfect Bayesian Equilibrium.

**The Voter’s Reelection Decision** In the second governance period, there are no electoral benefits from choosing a costly action. As such, the politician in office will choose the lowest possible action, i.e., $a_2 = \hat{a}$, independent of the history. So, the Voter’s electoral decision depends only on his expectation about the Politicians’ types. As a consequence, the Voter adopts a cutoff rule in the space of posterior beliefs. If the Voter expects the Incumbent’s first-period choice was $\hat{a}$ and he observed a level of public goods $g_1$, his posterior belief that the Incumbent is of high type is $\Pr_I(\theta | g_1, \hat{a})$. The Voter reelects the Incumbent if and only if his posterior

$$\Pr_I(\theta | g_1, \hat{a}) = \frac{\pi_I \phi(g_1 - f(\hat{a}, \theta))}{\pi_I \phi(g_1 - f(\hat{a}, \theta)) + (1 - \pi_I) \phi(g_1 - f(\hat{a}, \theta))}$$
is higher than the threshold $\pi_C$.

The fact that the Voter adopts a cutoff rule in the space of posterior probabilities implies that the Voter also adopts a cutoff rule in the space of public goods. To see this, define the likelihood ratio as a function of public goods and actions, i.e., $LR : \mathbb{R} \times A \to \mathbb{R}$ with

$$LR(g, a) = \frac{\phi(g - f(a, \theta))}{\phi(g - f(a, \theta))}.$$ 

Observe that $Pr_I(\theta|g_1, \tilde{a}) \geq \pi_C$ holds if and only if

$$LR(g_1, \tilde{a}) = \frac{\phi(g_1 - f(\tilde{a}, \theta))}{\phi(g_1 - f(\tilde{a}, \theta))} \geq \frac{1 - \pi_I}{\pi_I} \frac{\pi_C}{1 - \pi_C}. \quad (1)$$

Since $LR(\cdot, \tilde{a})$ is onto, there exists some $\hat{g}(\tilde{a})$ satisfying

$$LR(\hat{g}(\tilde{a}), \tilde{a}) = \frac{\phi(\hat{g}(\tilde{a}) - f(\tilde{a}, \theta))}{\phi(\hat{g}(\tilde{a}) - f(\tilde{a}, \theta))} = \frac{1 - \pi_I}{\pi_I} \frac{\pi_C}{1 - \pi_C}. \quad (2)$$

By the MLRP, the left-hand side of Equation (1) is strictly increasing in $g_1$. So, if the Voter believes the Incumbent’s first-period action is $\tilde{a}$ and the first-period outcome is $g_1$, the Voter reelects the Incumbent if $g_1 > \hat{g}(\tilde{a})$ and replaces her if $g_1 < \hat{g}(\tilde{a})$.

**Example 2.1.** Suppose the model is symmetric. Since $\pi_I = \pi_C$, Equation (2) says that $\hat{g}(a)$ solves $\phi(\hat{g}(a) - f(a, \theta)) = \phi(\hat{g}(a) - f(a, \theta))$. Since (a) the PDF $\phi$ is symmetric and single-peaked and (b) $\hat{g}(a) - f(a, \theta) > \hat{g}(a) - f(a, \theta)$, it follows that $(\hat{g}(a) - f(a, \theta)) = -(\hat{g}(a) - f(a, \theta))$. This implies

$$\hat{g}(a) = f(a, \theta) + f(a, \theta) = \frac{f(a, \theta) + f(a, \theta)}{2}.$$ 

**Example 2.2.** Let $\phi$ be the PDF of the standard normal distribution. Then,

$$\frac{\exp^{-\frac{1}{2}(\hat{g}(a) - f(a, \theta))^2}}{\exp^{-\frac{1}{2}(\hat{g}(a) - f(a, \theta))^2}} = \frac{\pi_C(1 - \pi_I)}{\pi_I(1 - \pi_C)}.$$ 

It follows that

$$\hat{g}(a) = \frac{\log \left( \frac{\pi_C(1 - \pi_I)}{\pi_I(1 - \pi_C)} \right) + f(a, \theta) + f(a, \theta)}{2}.$$ 

Observe that the Voter always uses the same cutpoint in the space of posterior beliefs, namely $\pi_C$. However, his cutpoint in the space of public goods depends on his belief about the Incumbent’s action. In Example 2.1, this cutpoint is increasing in the action. One natural conjecture is that this is always the case: If the Voter expects the Incumbent to choose a higher action, she holds the Incumbent to a higher benchmark for public good provision, since a larger portion of any success is credited to effort rather than competence. Section 4 will show that this conjecture is incorrect. (That analysis will be based on Example 2.2.)
The Incumbent’s First-Period Choice  Suppose the Incumbent chooses the action $a$ and the Voter expects the Incumbent to choose action $a_\ast$. The Incumbent is reelected (resp. replaced) if the level of public goods observed, $f(a,\theta_I) + \epsilon_1$, exceeds (resp. falls short of) the Voter’s threshold $\hat{g}(a_\ast)$. The probability that the Incumbent is reelected if she chooses $a$ when the Voter expects $a_\ast$, $\Pr(a|a_\ast)$, is then the probability that $\epsilon_1 \geq \hat{g}(a_\ast) - f(a,\theta_I)$,\(^3\) i.e.,

$$\Pr(a|a_\ast) = \pi_I \left[ 1 - \Phi \left( \hat{g}(a_\ast) - f(a,\theta_I) \right) \right] + (1 - \pi_I) \left[ 1 - \Phi \left( \hat{g}(a_\ast) - f(a,\theta_I) \right) \right]. \quad (3)$$

Now consider two actions $a,a' \in A$. If the Voter expects the Incumbent to take the action $a_\ast$, then the incremental increase in probability of reelection from choosing $a'$ instead of $a$ is $\text{IR} (a',a|a_\ast) = \Pr(a'|a_\ast) - \Pr(a|a_\ast)$.\(^4\) We can construct an equilibrium where the Incumbent chooses the action $a_\ast$ in the first period if and only if, for each action $a \in A$, $\text{IR} (a_\ast,a|a_\ast)(B - c(a)) \geq c(a_\ast) - c(a)$, i.e., if and only if the Incumbent’s incremental benefit from choosing $a_\ast$ over $a$ (when the Voter expects $a_\ast$) is higher than her incremental cost of choosing $a_\ast$ over $a$.

Set of Equilibria  Fix a model associated with the benefit of reelection and cost function $(B,c)$. In light of the analysis above, we will identify a pure-strategy Perfect Bayesian Equilibrium with the first-period action chosen by the Incumbent in that equilibrium. Write $E(B,c)$ for the set of first-period equilibrium actions, in the game parameterized by $(B,c)$. Say the pair $(B,c)$ justifies $a \in A$ if $a \in E(B,c)$. Say an action $a$ is justifiable if there exists $(B,c)$ that justifies $a$.

Recall, in any equilibrium, the second-period Politician chooses $a$. Thus, the set $E(B,c)$ represents the set of possible levels of effective accountability in the model $M(B,c)$. Varying the institutions of formal accountability varies the consequences of holding office or exerting effort and, in so doing, can lead to variation in $(B,c)$. In turn, varying $(B,c)$ can lead to variation in the level of effective accountability, in the sense that it can change the set of first-period equilibrium actions. But we will see that, even without varying the institutions of formal accountability (or $(B,c)$), we can have variation in the level of effective accountability. That is, for a given model $M(B,c)$, there may be multiple equilibria that differ in their first-period action.

To understand why this can occur, recall that the Voter 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. This last fact is what generates the possibility of multiple equilibria that differ in first-period action: Distinct expectations about the Incumbent’s first-period behavior are associated with distinct cut-off rules for the Voter. This, in turn generates distinct first-period incentives for the Incumbent, since the ex ante probability of reelection $\Pr(\cdot|\cdot) : A \times A \to [0,1]$ depends both on the Incumbent’s action and the Voter’s expectation of the Incumbent’s action. This intuition is familiar from the career concerns literature (DJT). In Section 5, we return to the

---

\(^3\)Here we use the fact that the random variable generating $\epsilon_1$ is atomless. This follows from the fact that $\Phi$ is continuous.

\(^4\)Observe that, as a matter of computation, we can and do define the function $\hat{g}(a)$ from Equation (2), irrespective of whether $a$ is a first-period equilibrium action. With this, we can and do define the function $\text{IR} (a',a|a_\ast)$ (by way of Equation (3)) for each triple of actions $(a',a,a_\ast)$, i.e., irrespective of whether $a_\ast$ is a first-period equilibrium action.
question of when such multiplicity does vs. does not obtain, and provide results that go beyond this familiar intuition.

3 Welfare and Information

One natural hypothesis is that, all else equal, a higher level of effective accountability is necessarily beneficial to the Voter. In this section, we show that this conjecture is incorrect in general, and provide sufficient conditions under which the conjecture is in fact correct.

Consider a model in which there is an equilibrium where the Incumbent plays the first-period action $a^\ast$. Expected Voter welfare is the sum of expected first-period Voter welfare and \textit{ex ante} expected second-period Voter welfare. Expected first-period Voter welfare is

$$VW_1(a^\ast) = \Pr(\theta)f(a^\ast, \theta) + (1 - \Pr(\theta))f(a^\ast, \theta).$$

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

$$VW_2(a^\ast) = \Pr(\theta_{P2} = \theta|a^\ast)f(a, \theta) + (1 - \Pr(\theta_{P2} = \theta|a^\ast))f(a, \theta),$$

where $\Pr(\theta_{P2} = \theta|a^\ast)$ is the equilibrium probability that the Politician in office in the second period is type $\theta$.\footnote{The term $\Pr(\theta_{P2} = \theta|a^\ast)$ can be expressed in terms of primitives of the model. See the Appendix.} Note that $\Pr(\theta_{P2} = \theta|a^\ast)$ depends on the first-period equilibrium action and the Voter’s equilibrium cutoff rule. With these definitions in hand, we can show that higher levels of effective accountability can be associated with lower Voter welfare.

Example 3.1. Let $A = \{a, a^\ast, a^{**}\}$, where $a^{**} > a^\ast > a$. The production function $f$ is described by the following table.

<table>
<thead>
<tr>
<th>$\theta$</th>
<th>$\theta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a^{**}$</td>
<td>38</td>
</tr>
<tr>
<td>$a^{\ast}$</td>
<td>35</td>
</tr>
<tr>
<td>$a$</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 1: Production Function

Take $\Phi$ to be the CDF of the standard Normal distribution and let $\pi_I = \pi_C = .6$.

Using Example 2.1, the incremental returns in reelection probabilities satisfy the following:

| $\text{IR}(a_{**}, a^\ast|a_{**}) \in (.498, .499)$ | $\text{IR}(a_{**}, a|a_{**}) \in (.534, .535)$ |
| $\text{IR}(a, a^\ast|a_{**}) \in (-.422, -.421)$ | $\text{IR}(a, a^\ast|a) \in (.268, .269)$ |
| $\text{IR}(a, a_{**}|a) \in (-.401, -.399)$ | $\text{IR}(a, a_{**}|a) \in (-.401, -.399)$ |

Table 2: Incremental Returns in Reelection Probabilities
We will consider two models parameterized by \((B_{**}, c_{**}) = (10, c_{**})\) and \((B_{*}, c_{*}) = (5, c_{*})\), where \(c_{**}(a_{**}) = c_{*}(a_{**}) = 5\), \(c_{**}(a_{*}) = c_{*}(a_{*}) = 1\), and \(c_{**}(g) = c_{*}(g) = 0\). It can be verified that \(\mathcal{E}(10, c_{**}) = \{a_{**}\}\) and \(\mathcal{E}(5, c_{*}) = \{a_{*}\}\). Thus, the model associated with \((B_{**}, c_{**})\) has a higher level of effective accountability than the model associated with \((B_{*}, c_{*})\). But, the Voter’s welfare is strictly higher in the model associated with \((B_{*}, c_{*})\) than in the model associated with \((B_{**}, c_{**})\). In particular,

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

and

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

Example 3.1 illustrates that higher levels of effective accountability may be associated with lower second-period Voter welfare, so much so that it may also be associated with lower levels of total Voter welfare. The key is that the Voter’s expected second-period welfare depends on the probability that the Politician in office in the second-period is of high type. This, in turn, depends on the Voter’s ability to learn about the Incumbent’s type, given the (equilibrium) level of first-period accountability. Higher levels of effective accountability may hinder the Voter’s ability to learn about the Incumbent’s type.

We begin by pointing out that higher levels of effective accountability can only hinder the Voter’s ability to learn about type if effort and type are substitutes, as they are in Example 3.1.

**Definition 3.1** (Lehmann, 1988). Say \(a_{**}\) is more informationally effective than \(a_{*}\) if \(\theta > \theta'\), 
\[
\Phi(g_{**} - f(a_{**}, \theta)) = \Phi(g_{*} - f(a_{*}, \theta)), \text{ and } \Phi(g_{**}' - f(a_{**}, \theta')) = \Phi(g_{*}' - f(a_{*}, \theta')),
\]
then \(g_{**} \geq g_{**}'\).

Because \(\Theta\) contains two types, \(a_{**}\) is more informative than \(a_{*}\) in the sense of Blackwell (1951) if and only if \(a_{**}\) is more informationally effective than \(a_{*}\). (See Blackwell and Girshick, 1979 or Jewitt, 2007.)

**Theorem 3.1.** Fix \(a_{**} > a_{*}\).

(i) If effort and type are complements at \((a_{*}, a_{**})\), then \(a_{**}\) is more informationally effective than \(a_{*}\).

(ii) If effort and type are substitutes at \((a_{*}, a_{**})\), then \(a_{**}\) is more informationally effective than \(a_{*}\).

**Proof.** For each \(\theta \in \Theta\), write \(\Phi_{**, \theta}\) (resp. \(\Phi_{*, \theta}\)) for the CDF defined by \(\Phi_{**, \theta}(g) = \Phi(g - f(a_{**}, \theta))\) (resp. \(\Phi_{*, \theta}(g) = \Phi(g - f(a_{*}, \theta))\)). Since \(\Phi_{**, \theta}\) is strictly increasing on \(\mathbb{R}\), for each \(x \in (0, 1)\), \(\Phi_{**, \theta}^{-1}(x)\) is well-defined. Observe that \(a_{**}\) is more effective than \(a_{*}\) if and only if, for each \(g \in \mathbb{R}\), 
\[
(\Phi_{**, \theta})^{-1}(\Phi_{*, \theta}(g)) = f(a_{**}, \theta) - f(a_{*}, \theta) + g.
\]

Start by noting that, for each \(x \in (0, 1)\),
\[
\Phi_{**, \theta}^{-1}(x) = f(a_{**}, \theta) + \Phi_{*, \theta}^{-1}(x).
\]
To see this, fix $y = \Phi^{-1}_{*,\theta}(x)$ and observe that $\Phi_{*,\theta}(y) = \Phi_{*,\theta}(\Phi^{-1}_{*,\theta}(x)) = x$. So, $x = \Phi(y - f(a_{**}, \theta))$. Since $\Phi$ is injective, $\Phi^{-1}(x) = y - f(a_{**}, \theta)$. Alternatively, $\Phi^{-1}_{*,\theta}(x) = y = \Phi^{-1}(x) + f(a_{**}, \theta)$. This establishes Equation (4).

Next observe that $(\Phi_{*,\theta})^{-1}(\Phi_{*,\theta}(g)) = f(a_{**}, \theta) + \Phi^{-1}(\Phi_{*,\theta}(g)) = f(a_{**}, \theta) + \Phi^{-1}(\Phi(g - f(a_{*}, \theta))) = f(a_{**}, \theta) - f(a_{*}, \theta) + g$,

where the first equality is by Equation (4), the second is by definition of $\Phi_{*,\theta}$, and the last equality uses the fact that $\Phi$ is injective. This establishes the desired result. ■

Theorem 3.1 and Proposition 4 in Jewitt, 2007 give the following second-period welfare implication:

**Corollary 3.1.** Fix $a_{**} > a_{*}$, where $a_{*}$ and $a_{**}$ are each justifiable.

(i) If effort and type are complements at $(a_{*}, a_{**})$, then $VW_2(a_{**}) \geq VW_2(a_{*})$.

(ii) If effort and type are substitutes at $(a_{*}, a_{**})$, then $VW_2(a_{*}) \geq VW_2(a_{**})$.

If action and type are complements at $(a_{*}, a_{**})$, then increasing the level of effective accountability from $a_{*}$ to $a_{**}$ increases both first- and second-period welfare. If action and type are substitutes at $(a_{*}, a_{**})$, then increasing the level of effective accountability from $a_{*}$ to $a_{**}$ increases first-period welfare but decreases second-period welfare. In Example 3.1, this led to a decrease in total welfare. However, in other environments, the increase in first-period welfare may dominate the decrease in second-period welfare. Proposition 3.1 provides a sufficient condition for this to occur.

**Proposition 3.1.** Suppose the model is symmetric. If $a_{**} > a_{*}$ and $a_{*}$ and $a_{**}$ are each justifiable, then there exists $\pi[\bar{a_{*}}, a_{**}] \in (0, 1)$ so that the following are equivalent:

(i) $\pi_{1} \in (0, \pi[\bar{a_{*}}, a_{**}]) \cup \pi[\bar{a_{*}}, a_{**}], 1)$.

(ii) $VW_1(a_{**}) + VW_2(a_{**}) \geq VW_1(a_{*}) + VW_2(a_{*})$.

Proposition 3.1 focuses on the symmetric model, where the pool of Incumbents and Challengers coincide. It 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.

### 4 The Reelection Decision and Information

There is a natural conjecture concerning the relationship between the Voter’s reelection standard and the level of effective accountability: More stringent standards correspond to higher levels of
effective accountability. Intuitively, when there is a higher level of effective accountability, the Voter
expects a higher level of effort from the Incumbent and so should use a higher cutoff rule with respect
to public goods. Indeed, this was true in Example 2.1. But this intuition is incomplete: higher
levels of effective accountability may be associated with lower cutoffs in public goods. Example 4.1
shows this in the case of complements and Example 4.2 shows this in the case of substitutes.

**Example 4.1.** Let $A = \{1, 2\}$, $\theta = 1$ and $\bar{\theta} = 2$. Let the production function be $f(a, \theta) = a \cdot \theta$, so
that effort and type are complements. Take $\Phi$ to be the CDF of the standard Normal distribution
and let $\pi_I = 0.1$ and $\pi_C = 0.8$.

We will show that there are models $\mathcal{M}(B_{**}, c_{**})$ and $\mathcal{M}(B_*, c_*)$ with $2 \in \mathcal{E}(B_{**}, c_{**})$, $1 \in \mathcal{E}(B_*, c_*)$, and $\hat{g}(1) > \hat{g}(2)$. In light of the analysis of Example 2.2, take

$$\hat{g}(a) = \frac{\log \left( \frac{8(9)}{\pi(2)} \right)}{a} + \frac{3a}{2}$$

for each $a \in A$. This implies that

$$\hat{g}(1) \in (5.08, 5.09) \quad \text{and} \quad \hat{g}(2) \in (4.79, 4.80).$$

With this, observe that the incremental returns in reelection probabilities satisfy

$$\text{IR}(2, 1|2) \in (.023, .024) \quad \text{and} \quad \text{IR}(2, 1|1) \in (.014, .015).$$

Take $(B_{**}, c_{**}) = (10, c_{**})$ and $(B_*, c_*) = (3, c_*)$, where $c_{**}(2) = c_*(2) = 1$, $c_{**}(1) = c_*(1) = 0$. It
then be verified that $\mathcal{E}(10, c_{**}) = \{2\}$ and $\mathcal{E}(3, c_*) = \{1\}.$

**Example 4.2.** The environment is as in Example 4.1 with two exceptions: The production function
is now $f(a, \theta) = \sqrt{a + \theta}$, so that effort and type are substitutes. The ex ante probabilities that the
Incumbent and Challenger are the high types are $\pi_I = 0.6$ and $\pi_C = 0.4$.

We will show that there are models $\mathcal{M}(B_{**}, c_{**})$ and $\mathcal{M}(B_*, c_*)$ with $2 \in \mathcal{E}(B_{**}, c_{**})$, $1 \in \mathcal{E}(B_*, c_*)$, and $\hat{g}(1) > \hat{g}(2)$. In light of the analysis of Example 2.2, take

$$\hat{g}(a) = \frac{4a}{2(2^{a/2})} + \frac{\sqrt{a + 2} + \sqrt{a + 1}}{2}$$

for each $a \in A$. Now

$$\hat{g}(1) \in (-0.98, -0.97) \quad \text{and} \quad \hat{g}(2) \in (-1.17, -1.16).$$

With this, observe that the incremental returns in reelection probabilities satisfy

$$\text{IR}(2, 1|2) \in (.001, .002) \quad \text{and} \quad \text{IR}(2, 1|1) \in (.003, .004).$$

Take $(B_{**}, c_{**}) = (100, c_{**})$ and $(B_*, c_*) = (1, c_*)$, where $c_{**}(2) = c_*(2) = 1$, $c_{**}(1) = c_*(1) = 0$. It
can be verified that $E(100, c_{**}) = \{2\}$ and $E(1, c_*) = \{1\}$.

Why might the Voter adopt a lower benchmark for reelection when there is a higher level of effective accountability? Observe that the Voter may receive bad news (resp. good news) about the Incumbent, relative to his prior, and nonetheless retain (resp. replace) her. This can occur if the ex ante probability that the Incumbent is a high type is higher (resp. lower) than the ex ante probability that the Challenger is a high type. With this in mind, observe that: (i) more informative experiments exacerbate good and bad news about the Incumbent, while (ii) less informative experiments temper good and bad news about the Incumbent. If the benchmark associated with the lower action provides good news about the Incumbent (relative to the prior $\pi_I$), then a more informative experiment will require a lower benchmark for reelection. Likewise, if the benchmark associated with the higher action provides bad news about the Incumbent (relative to the prior $\pi_I$), then a less informative experiment will require a lower benchmark for reelection. This accounts for Examples 4.1–4.2.

This argument suggests that, if the Voter’s benchmark provides bad news about the Incumbent and there are complements, then increasing effective accountability should increase the Voter’s benchmark. And, similarly, if the Voter’s benchmark provides good news about the Incumbent and there are substitutes, then increasing effective accountability should also increase the Voter’s benchmark. Indeed, this intuition is reflected in the following result:

**Theorem 4.1.** Suppose $\phi$ is symmetric. Fix $a_{**} > a_*$.

(i) If $\pi_I \geq \pi_C$ and effort and type are complements at $a_{**}, a_*$, then $\hat{g}(a_{**}) > \hat{g}(a_*)$.

(ii) If $\pi_C \geq \pi_I$ and effort and type are substitutes at $a_{**}, a_*$, then $\hat{g}(a_{**}) > \hat{g}(a_*)$.

To better understand this idea, recall that increasing the action $a$ has two effects on the benchmark $\hat{g}(a)$: The first effect is due to the increase in productivity, holding the level of informativeness fixed. Let $\nu(a)$ be the level of public goods that, in equilibrium, provides neutral news about the Incumbent—that is, that leaves the Voter’s beliefs about the Incumbent unchanged. Then $\nu(a)$ solves $\Pr_1(\theta | \nu(a), a) = \pi_I$. When $\phi$ is symmetric, we can repeat the analysis in Example 2.1 to get

$$\nu(a) = \frac{f(a, \bar{\theta}) + f(a, \theta)}{2}.$$  

Observe that $\nu(a)$ is increasing in $a$.

The second effect is due to the change in informativeness. Take

$$\iota(a) = \frac{f(a, \bar{\theta}) - f(a, \theta)}{2}$$

as a measure of informational effectiveness. By Theorem 3.1, if $a_{**} > a_*$, then the following equivalence holds: Effort and type are local complements (resp. local substitutes) at $a_{**}, a_{**}$ if and only if $\iota(a_{**}) \geq \iota(a_*)$ (resp. $\iota(a_*) \geq \iota(a_{**})$).
For each \( a \in A \), we can express the likelihood ratio in terms of the variables \((g, \nu(a), \iota(a))\),

\[
LR(g, a) = \frac{\phi(g - \nu(a) - \iota(a))}{\phi(g - \nu(a) + \iota(a))}.
\]

Fix \( a_{**} > a_{*} \) where both \( a_{*} \) and \( a_{**} \) are justifiable. Write \( LR_*(\cdot) = LR(\cdot, a_{*}) \) and \( LR_{**}(\cdot) = LR(\cdot, a_{**}) \). Figure 4.1 draws the function \( LR_*(\cdot) \). Observe that, by the MLRP, it is increasing in \( g \). The values \( \nu(a_{*}) \) and \( \hat{g}(a_{*}) \) solve \( LR_*(\nu(a_{*})) = 1 \) and \( LR_*(\hat{g}(a_{*})) = \beta \), where we write \( \beta \) for the benchmark \( \frac{1-\pi_I}{\pi_I - \pi_C} \). The value \( \hat{g}(a_{**}) \) also solves \( LR_{**}(\hat{g}(a_{**})) = \beta \). Thus, to determine the relationship between \( \hat{g}(a_{*}) \) and \( \hat{g}(a_{**}) \), we want to understand the relationship between the functions \( LR_*(\cdot) \) and \( LR_{**}(\cdot) \).

\[
\tilde{L}(g) = \frac{\phi(g - \nu(a_{**}) - \iota(a_{*}))}{\phi(g - \nu(a_{**}) + \iota(a_{*}))}.
\]

That is, this function is obtained from \( LR_* \) by increasing the \( \nu(a_{*}) \) variable to \( \nu(a_{**}) \), leaving the \( \iota(a_{*}) \) variable fixed. This change can be viewed as a thought experiment where we change the level of productivity expected by the Voter, holding fixed the informativeness of the signal. Figure 4.1 depicts this function below the function \( LR_* \) and intersecting the benchmark \( \beta \) at a higher level of public goods \( \tilde{g} \). Indeed, this must always be the case.

**Lemma 4.1.**

(i) For each \( g \), \( LR_*(g) > \tilde{L}(g) \).

(ii) If \( \tilde{L}(\hat{g}) = \beta \), then \( \tilde{g} > \hat{g}(a_{*}) \).

Now suppose we begin with the function \( \tilde{L} \) and allow the level of informativeness to change, going from \( \iota(a_{*}) \) to \( \iota(a_{**}) \). Figures 4.2a-4.2b depict the case of complements, which increases
informativeness. In this case, outcomes that were good news about the Incumbent are now better news and outcomes that were bad news about the Incumbent are now worse news. Thus, if \( g > \nu(a_{**}) \) (resp. \( g < \nu(a_{**}) \)) then \( LR_{**}(g) > \tilde{L}(g) \) (resp. \( LR_{**}(g) < \tilde{L}(g) \)). Figures 4.2c-4.2d depict the case of substitutes, which decreases informativeness. This tempers good news about the Incumbent and exacerbates bad news about the Incumbent. Thus, if \( g > \nu(a_{**}) \) (resp. \( g < \nu(a_{**}) \)) then \( LR_{**}(g) < \tilde{L}(g) \) (resp. \( LR_{**}(g) > \tilde{L}(g) \)).

**Lemma 4.2.** Suppose \( \phi \) is symmetric.

(i) Suppose effort and type are complements at \( a_*,a_{**} \). Then \( g > \nu(a_{**}) \) implies \( LR_{**}(g) > \tilde{L}(g) \), \( g < \nu(a_{**}) \) implies \( LR_{**}(g) < \tilde{L}(g) \), and \( LR_{**}(\nu(a_{**})) = \tilde{L}(\nu(a_{**})) \).

(ii) Suppose effort and type are substitutes at \( a_*,a_{**} \). Then \( g > \nu(a_{**}) \) implies \( LR_{**}(g) < \tilde{L}(g) \), \( g < \nu(a_{**}) \) implies \( LR_{**}(g) > \tilde{L}(g) \), and \( LR_{**}(\nu(a_{**})) = \tilde{L}(\nu(a_{**})) \).

Recall, the auxiliary function \( \tilde{L} \) meets the benchmark \( \beta \) at some level of public goods \( \tilde{g} > \hat{g}(a_*) \). If, at the benchmark \( \beta \), \( LR_{**} \) is to the right of \( \tilde{L} \), then we also have that \( \hat{g}(a_{**}) \geq \tilde{g} > \hat{g}(a_*) \). Refer
to Figures 4.2. When effort and type are complements at $a_s, a_{**}$, then $LR_{**}$ is to the right of $\tilde{L}$ at the benchmark $\beta$ provided the benchmark is less than 1. When effort and type are substitutes at $a_s, a_{**}$, then $LR_{**}$ is to the right of $\tilde{L}$ at the benchmark $\beta$ provided the benchmark is above 1. Using the fact that $1 \geq \beta$ if and only if $\pi_I \geq \pi_C$, this gives Theorem 4.1.

**Proof of Theorem 4.1.** By Lemma 4.1, it suffices to show that $\hat{g}(a_{**}) \geq \hat{g}$. To show this, it suffices to show that $LR_{**}(\hat{g}) \leq \tilde{L}(\hat{g}) = \beta$. If so, then using the MLRP and the fact that $LR_{**}(\hat{g}(a_{**})) = \beta$, it follows that $\hat{g}(a_{**}) \geq \hat{g}$.

Observe that, by symmetry of the PDF,

$$\tilde{L}(\nu(a_{**})) = \frac{\phi(-\nu(a_s))}{\phi(\nu(a_s))} = 1 = \frac{\phi(-\nu(a_{**}))}{\phi(\nu(a_{**}))} = LR_{**}(\nu(a_{**})).$$

So, if $\pi_I \geq p_C$ (resp. $\pi_C \geq p_I$) and effort and type are complements (resp. substitutes) at $a_{**}, a_s$, Lemma 4.2 gives that $LR_{**}(\hat{g}) \leq \tilde{L}(\hat{g}) = \beta$, as desired. ■

**Corollary 4.1.** Suppose $\phi$ is symmetric. Fix $a_{**} > a_s$, where $a_s$ and $a_{**}$ are each justifiable. If effort and type are neither strict substitutes nor strict complements at $a_s, a_{**}$, then $\hat{g}(a_{**}) > \hat{g}(a_s)$.

## 5 Variation in Effective Accountability

Up until now, we have attempted to understand the effect of higher levels of effective accountability on the Voter’s welfare and the Voter’s reelection strategy. In the examples we have seen, variation in effective accountability resulted from of variation in the institutions of formal accountability (reflected as variations in $(B, c)$). In this section, we will see that variation in effective accountability can arise even without variation in the institutions of formal accountability. Such variation can arise if, within a given model $\mathcal{M}(B, c)$, there are multiple equilibria that differ in their level of effective accountability—i.e., if, for a given model $\mathcal{M}(B, c)$, the set $\mathcal{E}(B, c)$ has at least two actions.

Section 2 pointed to why such multiplicity might arise: Distinct expectations about the Incumbent’s first-period behavior are associated with distinct cut-off rules for the Voter. This generates distinct first-period incentives for the Incumbent. But, while suggestive that this may lead to a situation of multiplicity, the multiplicity is not inevitable, as the next example illustrates.

**Example 5.1.** Take $A = \{0, 1, 2\}$, $\bar{\theta} = 2, \theta = 1$, and $f(a, \theta) = a + \theta$. Let $\pi_I = \pi_C = \frac{3}{4}$ and let $\Phi$ be the standard normal distribution. Then $\hat{g}(a) = a + 1.5$ (see Example 2.1) and IR $(a, a' | a)$ satisfies:

| IR $(0, 1 | 0)$ $\in (-.2771, -.2770)$ | IR $(0, 2 | 0) \in (-.3830, -.3829)$ |
|-------------------------------|-------------------------------|
| IR $(1, 0 | 1) \in (.3476, .3477)$ | IR $(1, 2 | 1) \in (.2771, -.2770)$ |
| IR $(2, 0 | 2) \in (.5440, .5441)$ | IR $(2, 1 | 2) \in (.3476, .3477)$ |

---

Appendix C shows that variations in the institutions of formal accountability can lead to changes in effective accountability, despite not leading to systematic changes in effective accountability.
Consider the model $\mathcal{M}(B,c)$, where $B = 1$ and $c(a) = .2a^2$. There is no equilibrium where the Incumbent chooses a first-period action $a \in \{0, 2\}$. (Observe that $c(0) - c(1) = -.2 > \text{IR}(0, 1|0) = \text{IR}(0, 1|0)(B - c(0))$ and $c(2) - c(0) = .8 > \text{IR}(2, 0|2) = \text{IR}(2, 0|2)(B - c(0))$.)

However, we can construct an equilibrium where the Incumbent chooses a first-period action $a_\ast = 1$. (Observe that $\text{IR}(1, 0|1)(B - c(0)) = \text{IR}(1, 0|1) > .2 = c(1) - c(0)$ and $\text{IR}(1, 2|1)(B - c(0)) = \text{IR}(1, 2|1) > -.6 = c(1) - c(2)$.)

We next point out that, by perturbing the cost of effort, we have multiple equilibria that differ in their levels of effective accountability.

**Example 5.2.** Again, consider the environment in Example 5.1. Now change only the cost function so that $c(a) = .3a^2$. We can still construct an equilibrium where the Incumbent chooses the first-period action $a_\ast = 1$, since $\text{IR}(1, 0|1)(B - c(0)) = \text{IR}(1, 0|1) > .3 = c(1) - c(0)$ and $\text{IR}(1, 2|1)(B - c(0)) = \text{IR}(1, 2|1) > -.9 = c(1) - c(2)$. But now we can also construct an equilibrium where the Incumbent chooses the first-period action $a_{\ast\ast} = 0$, since $\text{IR}(0, 1|0)(B - c(0)) = \text{IR}(0, 1|0) > -.3 = c(0) - c(1)$ and $\text{IR}(0, 2|0)(B - c(0)) = \text{IR}(0, 2|0) > -1.2 = c(0) - c(2)$.

Examples 5.1-5.2 hold fixed the production technology and beliefs (about type and noise). In Example 5.1, any equilibrium of the model induces the same first-period action. Example 5.2 perturbs the cost function and shows that there are now multiple equilibria that differ in their first-period action.

In what follows, we fix the production technology and the beliefs (about type and noise). We ask: Does there exist a benefit of reelection and a cost function so that there are multiple equilibria that differ in their first-period action? We provide a condition on the production technology and beliefs that is necessary and sufficient for an affirmative answer to this question.

Why is the approach we take of interest? We view both the Incumbent’s benefit from reelection and the cost of effort as fundamentally subjective. The set of institutions will influence 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 institutions cannot pin down the Politician’s utility from these material outcomes. On the other hand, for a given set of institutions, the analyst may have intuitions or empirical knowledge about the nature of the production function or the beliefs. For instance, 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.

**Choosing the Benefit of Reelection and Cost Function** Say the pair $(a_*, a_{\ast\ast})$ is **justifiable** if there exists some pair $(B, c)$ so that $\{a_*, a_{\ast\ast}\} \subseteq \mathcal{E}(B, c)$. (Observe that $a_*$ and $a_{\ast\ast}$ can each be justifiable even if the pair $(a_*, a_{\ast\ast})$ is not justifiable, i.e., we may have $a_* \in \mathcal{E}(B_*, c_*)$, $a_{\ast\ast} \in \mathcal{E}(B_{\ast\ast}, c_{\ast\ast})$, but no pair $(B, c)$ with $\{a_*, a_{\ast\ast}\} \subseteq \mathcal{E}(B, c)$.)
Begin by fixing a benefit of reelection and cost function, \((B, c)\). Observe that \((B, c)\) justifies \(a_\ast\) if and only if \(\text{IR}(a_\ast, a|a_\ast)(B - c(a)) \geq c(a_\ast) - c(a)\), for all \(a \in A\). So, if \(\{a_\ast, a_{\ast\ast}\} \subseteq \mathcal{E}(B, c)\) then

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

(Here we use the fact that \(-\text{IR}(a_\ast, a_{\ast\ast}|a_\ast) = \text{IR}(a_{\ast\ast}, a_\ast|a_\ast)\).) This trivial argument gives a necessary condition for the pair \((a_\ast, a_{\ast\ast})\) to be justifiable, namely

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

To interpret this necessary condition, take \(a_{\ast\ast} > a_\ast\). If the Voter uses the cutoff rule \(\hat{g}(a_{\ast\ast})\) (resp. \(\hat{g}(a_\ast)\)), then \(\text{IR}(a_{\ast\ast}, a_{\ast\ast}|a_{\ast\ast})\) (resp. \(\text{IR}(a_{\ast\ast}, a_{\ast\ast}|a_\ast)\)) is the incremental increase in the probability of reelection in moving from the less productive action to the more productive action. Equation (6) says that the incremental increase in the probability of reelection, in moving from the less productive action \(a_\ast\) to the more productive action \(a_{\ast\ast}\) is higher when the Voter uses the cutoff rule \(\hat{g}(a_{\ast\ast})\) vs. when the Voter uses the cutoff rule \(\hat{g}(a_\ast)\).

We are interested in providing a sufficient condition for a pair of distinct actions \((a_\ast, a_{\ast\ast})\) to be justifiable. Observe that, if Equation (6) is satisfied, then we can find a benefit of reelection and a cost function \((B, c)\) satisfying Equation (5). If \(a_\ast\) and \(a_{\ast\ast}\) are the only possible actions, Equation (5) suffices to conclude that \((B, c)\) justifies both \(a_\ast\) and \(a_{\ast\ast}\). But, when there are three or more actions, Equation (5) is no longer sufficient: For a given \((B, c)\), Equation (5) may be satisfied even if \((B, c)\) does not justify either of \(a_\ast, a_{\ast\ast}\), or both. Example 5.3 makes this point.

**Example 5.3.** Again consider the environment in Example 5.1, but let \(\pi_I = \pi_C = .85\). Then \(\text{IR}(a, a'|a)\) satisfies the following bounds:

| \(\text{IR}(0, 1|0) \in (-.2630, -.2629)\) | \(\text{IR}(0, 2|0) \in (-.3507, -.3506)\) |
| \(\text{IR}(1, 0|1) \in (.3617, .3618)\) | \(\text{IR}(1, 2|1) \in (-.2630, -.2629)\) |
| \(\text{IR}(2, 0|2) \in (.5763, .5764)\) | \(\text{IR}(2, 1|2) \in (.3617, .3618)\) |

Consider the model \(\mathcal{M}(B, c)\), where \(B = 1\) and \(c(a) = .1a^2\). Observe that

\[
\text{IR}(2, 0|2) \geq \frac{c(2) - c(0)}{B - c(0)} = .4 \geq -\text{IR}(0, 2|0) = \text{IR}(2, 0|0).
\]

So Equation (5) is satisfied for actions \(a_\ast = 0\) and \(a_{\ast\ast} = 2\). We might then conclude that \(\{0, 2\} \subseteq \mathcal{E}(B, c)\). But, in fact, \(0 \notin \mathcal{E}(B, c)\).

Observe that Equation (5) says that, if there is an equilibrium where \(a_\ast\) (resp. \(a_{\ast\ast}\)) is played, then the Incumbent does not have an incentive to deviate to \(a_{\ast\ast}\) (resp. \(a_\ast\)). It is silent about deviating to some action \(a \neq a_\ast, a_{\ast\ast}\) and, indeed, Example 5.3 illustrates that there may very well be an incentive to deviate to some alternate action.
In sum, if the local condition $\text{IR}(a_{ss}, a_s | a_{ss}) \geq \text{IR}(a_{ss}, a_s | a_s)$ is satisfied, it is trivial to construct a pair $(B, c)$ so that Equation (5) is satisfied. However, Equation (5) is not sufficient for $(B, c)$ to justify either of $a_s$ or $a_{ss}$. Nonetheless, a more subtle fact is true: Whenever the local condition $\text{IR}(a_{ss}, a_s | a_{ss}) \geq \text{IR}(a_{ss}, a_s | a_s)$ is satisfied, we can construct some pair $(B, c)$ that justifies both $a_s$ and $a_{ss}$. For instance, if we amend the cost function in Example 5.3 to $c(a) = .27d^{1.05}$, then $\{0, 2\} \subseteq \mathcal{E}(B, c)$.

**Theorem 5.1.** Fix $a_{ss}, a_s \in A$, so that $a_{ss} > a_s$. The following are equivalent:

(i) $\text{IR}(a_{ss}, a_s | a_{ss}) \geq \text{IR}(a_{ss}, a_s | a_s)$.

(ii) There exists a $(B, c)$ that justifies both $a_s$ and $a_{ss}$.

(iii) There is an uncountable set of pairs $(B, c)$, viz. $\mathcal{C}$, so that (a) if $(B, c) \in \mathcal{C}$ then $\{a_s, a_{ss}\} \subseteq \mathcal{E}(B, c)$, and (b) distinct elements of the set $\mathcal{C}$ represent different preferences.

Observe that the requirement of $\text{IR}(a_{ss}, a_s | a_{ss}) \geq \text{IR}(a_{ss}, a_s | a_s)$ is identified from the technological environment—i.e., the production technology and distributions of types and noise—alone. Thus, Parts (i)-(ii) provide a necessary and sufficient condition for the technological environment to be consistent with multiple equilibria that differ in their level of effective accountability: There are preferences of Politicians that justify (at least) two distinct actions if and only if there are actions $a_{ss} > a_s$ so that $\text{IR}(a_{ss}, a_s | a_{ss}) \geq \text{IR}(a_{ss}, a_s | a_s)$. Part (iii) says that such preferences are non-degenerate. If there is some preference profile $(B, c)$ that justifies both $a_s$ and $a_{ss}$, then there are an uncountable set of distinct preferences that each justify both $a_s$ and $a_{ss}$.

Focus on the equivalence between Parts (i) and (ii). One direction is trivial: It is immediate that Part (ii) implies Part (i). The converse is not obvious. When Part (i) holds, it is trivial to construct a pair $(B, c)$ so that Equation (5) holds. However, as we have seen, that is not sufficient for an equilibrium. In particular, if there is an equilibrium of $\mathcal{M}(B, c)$ with a first-period action $a_s$, then $\text{IR}(a_s, a | a_s) \geq \frac{c(a_s) - c(a)}{B}$ for all $a \in A$. Thus, there are $|A| - 1$ inequalities associated with any equilibrium action—for multiplicity there are at least $2(|A| - 1)$. This can result in uncountably many inequalities. Part (i), instead, verifies a single inequality.

**Technological Implications for Multiplicity**

We next show that any technology is consistent with the existence of multiple equilibria that differ in their first-period action.

**Proposition 5.1.** Suppose $\phi$ is symmetric. Fix actions $a_{ss} \neq a_s$. There exists a non-empty open set $\mathbb{P}[a_s, a_{ss}] \subseteq [0, 1] \times [0, 1]$ so that the pair $(a_s, a_{ss})$ is justifiable if $(\pi_I, \pi_C) \in \mathbb{P}[a_s, a_{ss}]$.

Proposition 5.1 says that, for any production technology, any two actions $a_s, a_{ss}$, and any distribution of noise, we can find some distribution of types so that both $a_s$ and $a_{ss}$ are justifiable. This distribution of types will typically depend on the productivity of the particular actions relative to the productivity of type. To clarify this point, focus on the case where the Incumbent and Challenger are drawn from the same pool of candidates, so that $\pi_I = \pi_C$.
Proposition 5.2. Suppose the model is symmetric, with \( \pi_I = \pi_C = \pi \) and fix \( a_{**} \neq a_* \). There exists a \( \hat{\pi}[a_*, a_{**}] \in (0,1) \) so that the pair \( (a_*, a_{**}) \) is justifiable if and only if \( \pi \geq \hat{\pi}[a_*, a_{**}] \).

Moreover:

(i) If effort and type are strict complements at \( a_*, a_{**} \), then \( \hat{\pi}[a_*, a_{**}] < \frac{1}{2} \).

(ii) If effort and type are both complements and substitutes at \( a_*, a_{**} \), then \( \hat{\pi}[a_*, a_{**}] = \frac{1}{2} \).

(iii) If effort and type are strict substitutes at \( a_*, a_{**} \), then \( \hat{\pi}[a_*, a_{**}] > \frac{1}{2} \).

Proposition 5.2 focuses on the case when the Incumbent and Challenger are drawn from the same pool of candidates. It says that the pair \( (a_*, a_{**}) \) is justifiable if the pool of candidates is sufficiently weighted toward high quality types. The phrase ‘sufficiently weighted’ is determined by \( \hat{\pi}[a_*, a_{**}] \), which in turn depends on whether effort and type are local substitutes or local complements.

The conclusions of Proposition 5.2 do not hinge on the Politicians being drawn from the same distribution. If they are drawn from ‘similar’ distributions, the conclusions obtain:

Proposition 5.3. Suppose \( \phi \) is symmetric. Fix actions \( a_{**} \neq a_* \) and let \( \pi \in (\hat{\pi}[a_*, a_{**}], 1) \). There exists some open neighborhood of \( (\pi_I, \pi) \in [0,1]^2 \), viz. \( N(\pi, \pi) \), so that, if \( (\pi_I, \pi_C) \in N(\pi, \pi) \), then \( (a_*, a_{**}) \) is justifiable.

Suppose the pair \( (a_*, a_{**}) \) is justifiable when \( \pi_I = \pi_C \). Proposition 5.3 says that, in all but a degenerate case, the pair \( (a_*, a_{**}) \) is also justifiable when \( \pi_C \) is close to \( \pi_I \).

Figures 5.1a–5.1c illustrate Proposition 5.2–5.3 in the context of an example. Specifically, \( \overline{\theta} = 2, \overline{\theta} = 1 \), and the production function is parameterized by \( f(a, \theta) = (.25a + 4\theta)^r \). (The PDF \( \phi \) is the standard normal distribution.) The figures illustrate the set of \( (\pi_I, \pi_C) \) so that the pair \( (a_*, a_{**}) = (1, 2) \) is justifiable. In the context of this example, the requirement that \( \pi_C \) is close to \( \pi_I \) does not appear particularly demanding.
Connection to Dewatripont, Jewitt and Tirole (1999b) In the context of wage-based incentives, DJT also show a multiplicity result. That result imposes assumptions on the distribution of types and the production technology. In particular, they consider a parameterized production technology that exhibits strict complementarities between the agent’s effort and ability. Theorem 5.1 says that, in our context, multiplicity (associated 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. This type of complementarity is generated endogenously through Voter learning. In particular, Proposition 5.1 says that strict complementarities between the Incumbent’s effort and ability are neither necessary nor sufficient for this type of complementarity: Effort and ability may be strict substitutes and still two actions can be justifiable.

DJT also provide a uniqueness result: In their setting, when the production function is additive between the agent’s effort and ability and the cost function is strictly convex, there is a unique equilibrium. We now produce an analogue of DJT’s result in our setting.  

Example 5.4. Let $A = \mathbb{R}_+$ and $f(a, \theta) = a + \theta$. Fix $(B, c)$ where $c$ is strictly convex and $c'(a) = 0$. Then $|\mathcal{E}(B, c)| \leq 1$.

To see this, choose $\bar{g}$ to solve

$$\frac{\phi(\bar{g} - \theta)}{\phi(\bar{g} - \theta)} = \frac{\pi_C(1 - \pi_I)}{\pi_I(1 - \pi_C)}.$$  

Then $\hat{g}(a) = \bar{g} + a$. With this, $\Pr(a|a_*) = \pi_I\Phi(\bar{g} + a_* - a - \theta) + (1 - \pi_I)\Phi(\bar{g} + a_* - a - \theta)$. So, if $a_* \in \mathcal{E}(B, c)$, then $a_*$ maximizes

$$(B - c(a)) \left[1 - \pi_I\Phi(\bar{g} + a_* - a - \theta) - (1 - \pi_I)\Phi(\bar{g} + a_* - a - \theta) \right] - c(a).$$  

(7)

Since $c(\cdot)$ is differentiable, $a_*$ solves

$$\pi_I\phi(\bar{g} - \theta) + (1 - \pi_I)\phi(\bar{g} - \theta) \leq \frac{c'(a_*)}{B - c(a)} \text{ with equality if } a_* > a.$$  

(8)

The left-hand side of Equation (8) is positive, so $c'(a) = 0$ implies $a_* \neq a$. Since $c(\cdot)$ is strictly convex, $c'(\cdot)$ is strictly increasing. Since the left-hand side of Equation (8) is constant in $a$, there can be at most one action $a_* > a$ that satisfies Equation (8) with equality.

Example 5.4 provides a production environment where, for any values of $(\pi_I, \pi_C)$, strict convexity of costs implies that equilibria do not differ in their level of effective accountability. On the other hand, Proposition 5.1 says that there exists $(\pi_I, \pi_C)$ and $(B, c)$ so that, in the same production environment, $|\mathcal{E}(B, c)| \geq 2$. The key is that the cost function $c$ need not be strictly convex.

That said, it would be a mistake to conclude that, in an additive production environment, multiplicity requires non-convex costs of effort. Example 5.2 showed that the production function may be additive, the cost function may be strictly convex, and yet there may be multiple equilibria.

---

7We thank the Associate Editor and Referee for pointing us to this example.
that differ in their level of effective accountability. (Example 5.5 to come will show an analogous conclusion for the case of strict substitutes.) The key difference between Example 5.2 and Example 5.4 is whether the action set is discrete vs. continuous.

These examples raise the question: In what circumstances can we find a pair \((B, c)\) where \(c\) is strictly convex, so that \((B, c)\) justifies distinct actions \(a_*\) and \(a_{**}\)? The answer appears to depend on (i) the cardinality of the action set, (ii) the production technology, (iii) the distributions of types, and (iv) the distribution of noise.

The examples also (at least, implicitly) raise the question of the extent to which multiplicity in the political agency setting coincides with or differs from multiplicity in the labor market setting. There is an important difference between the two settings: how agents are rewarded. In the labor market setting, the agent’s reward is constrained to be equal to the conditional expectation of his type. There, the benefit is a wage and the choice of wage is a strategic variable. As such, it makes sense that the wage is chosen after observing variables and is equal to the conditional expectation of the agent’s type. In the political setting, however, the benefit reflects a subjective utility from holding office, given the institutions of formal accountability. Consequently, the benefit is part of the description of the model and cannot vary with realized outcomes. On the one hand, this constrains the set of rewards we focus on. On the other hand, there is no reason for the benefit of reelection to equal the conditional expectation of the politician’s type. The fact that the choice of \(B\) is unconstrained appears to give us more freedom to get a multiplicity result. There may be pairs \((B, c)\) and \((B', c)\), so that \(\mathcal{E}(B, c)\) is a singleton and \(\mathcal{E}(B', c)\) contains at least two actions. (See Example C.1 in the Appendix for an illustration.)

**Welfare and Multiplicity** Example 3.1 illustrated that increasing effective accountability, by changing institutions of formal accountability, may decrease Voter welfare. We observe that the same effect can happen in a setting where the variation in effective accountability comes from equilibrium multiplicity.

**Example 5.5 (Example 3.1, Revisited).** Consider the model in Example 3.1, now parameterized by \((B, c)\) where \(B = 10\), \(c(a_{**}) = 5\), \(c(a_*) = 0.5\), and \(c(a_bar) = 0\). It can be verified that \(\{a_*, a_{**}\} \subseteq \mathcal{E}(B, c)\). And, as shown in Example 3.1, \(VW_1(a_*) + VW_2(a_*) > VW_1(a_{**}) + VW_2(a_{**})\).

When is it the case that, within a given model \(\mathcal{M}(B, c)\), higher levels of effective accountability are associated with higher levels of Voter welfare? Corollary 3.1 says that, when effort and type are (local) complements, higher levels of effective accountability are associated with higher levels of Voter welfare. Propositions 3.1 and 5.3 speak to the case when effort and type are (local) substitutes—at least when \(\pi_I = \pi_C = \pi\). Proposition 5.3 says that, in this case, the pair \((a_*, a_{**})\) is justifiable if and only if \(\pi\) meets some threshold \(\hat{\pi}[a_*, a_{**}]\), where \(\hat{\pi}[a_*, a_{**}] > \frac{1}{2}\). Proposition 3.1 adds that, in this case, \(a_{**}\) leads to higher Voter welfare if either \(\pi\) is sufficiently large, i.e., greater than some \(\overline{\pi}[a_*, a_{**}]\), or sufficiently small, i.e., less than some \(\underline{\pi}[a_*, a_{**}]\). Certainly then, if \(\pi\) is greater than \(\max\{\hat{\pi}[a_*, a_{**}], \underline{\pi}[a_*, a_{**}]\}\), we have that, within a given model, the equilibrium with a higher level of effective accountability will be the equilibrium with a higher level of Voter welfare.
However, we do not know if $\pi[a_*, a_{**}] > \hat{\pi}[a_*, a_{**}]$ and so do not know if the conclusion only holds when $\pi_I = \pi_C$ is sufficiently high.\footnote{The bound $\pi[a_*, a_{**}]$ is not determined constructively.}

Reelection Rules and Multiplicity  Examples 4.1-4.2 illustrated that increasing the level of effective accountability, by changing the institutions of formal accountability, may result in the Voter using a less stringent voting rule. In the case of complements, the same effect can arise in a setting where the variation in effective accountability comes from equilibrium multiplicity.

Example 5.6  (Example 4.1, Revisited). Consider the model in Example 4.1, now parameterized by $(B, c)$ where $B = 10$, $c(2) = 2$, $c(1) = 0$. It can be verified that $\{1, 2\} \subseteq \mathcal{E}(B, c)$. And, as shown in Example 4.1, $\hat{g}(1) > \hat{g}(2)$.

Recall Example 4.2 showed that, in the case of substitutes, variation in effective accountability can lead the Voter to adopt a less stringent voting rule. In the case of substitutes, we do not know if the same conclusion holds when the variation in effective accountability is due to equilibrium multiplicity. (We have neither been able to construct an analogous example nor prove that such an example does not exist.) It, thus, remains an open question.

6  Normative Implications

The results in Section 5 highlight that there can be variation in effective accountability, even if there is no variation in institutions. Variation in players’ expectations alone are sufficient to yield different levels of effective accountability. This fact has important normative implications, both for evaluating the performance of societies and institutions and for evaluating voter behavior.

6.1  Evaluating Society and Institutions: Accountability Traps

Our multiplicity and welfare results give rise to a concept that we will call an accountability trap. Fix a $\mathcal{M}(B, C)$. We say that the model is consistent with an accountability trap if there are (at least) two equilibria of the model which differ in terms of both their level of effective accountability and their level of voter welfare. A society is caught in an accountability trap if it is playing an equilibrium with a lower level of voter welfare than some other equilibrium.\footnote{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.}  Taken together, Corollary 3.1 and Proposition 5.2 show the possibility of an accountability trap.

The idea of accountability traps helps make sense of 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,
These observations are consistent with the idea that different societies, while sharing 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 possibility of accountability traps has important implications for improving effective accountability and voter welfare. The literature has presumed that improving voter welfare requires improving the level of effective accountability and, in turn, that increasing the level of effective accountability requires changing institutions. There are two difficulties with this statement. First, a society may be in an accountability trap that involves a high level of effective accountability relative to the level that would maximize voter welfare. In that case, improving voter welfare may require lowering the level of effective accountability. Second, 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 low voter welfare. 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 voter welfare in structurally similar societies.

How then can voter welfare 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. To escape such an accountability trap, a society must shift expectations. Communication and leadership can play an important role in shifting such expectations. (See, e.g., Dewan and Myatt, 2007, 2008, 2012 and Bidner and Francois, 2013.) 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.”

6.2 Evaluating Voter Behavior

The primary goal of electoral accountability is to ensure that politicians have incentives to take actions that benefit voters (Key, 1966; Pitkin, 1967; Fiorina, 1981; Achen and Bartels, 2004, 2006; Healy, Malhotra and Hyunjung, 2010). Voters are meant to achieve this goal within a given institutional setting. This raises the question: *How well are voters doing with respect to the goal of creating incentives for politicians to take costly actions on their behalf?* The question is of interest in political philosophy (Pitkin, 1967; Manin, Przeworski and Stokes, 1999; Mansbridge, 2003, 2009) and empirical political economy (Fiorina, 1981; Achen and Bartels, 2004, 2006; Healy, Malhotra and Hyunjung, 2010). A formal literature has developed to address this question.

The early formal literature addressed this by focusing on “pure moral hazard” models (Barro, 1973; Ferejohn, 1986; Austen-Smith and Banks, 1989). In those models, there is no candidate heterogeneity, so voters are indifferent between incumbents and challengers. As such, any reelection rule is a best response for the voter. Given this, evaluating voter behavior involves asking how close
voters come to using the rule that provides the politician with the most powerful incentives.

A critique of this literature, due to Fearon (1999), argues that this is the wrong benchmark. Choosing the reelection rule that provides the most powerful incentives is generally not credible: If the voter believes there is any difference in an incumbent’s and challenger’s expected future performance (i.e., if there is any payoff-relevant candidate heterogeneity), then sequential rationality requires him to elect the politician who is expected to perform better. Put differently, Fearon points out that the benchmark by which the voter is evaluated must incorporate the voter’s desire to select good types.\footnote{Notable exceptions to this view are Meirowitz (2007), Snyder and Ting (2008), and Schwabe (2009). In their models, equilibria exist in which the voter need not select good types. The key is that, in their models, the voter is indifferent between different types of politicians.} A subsequent literature has taken Fearon’s argument to suggest that, when politicians are heterogenous, the desire to select good types uniquely pins down voter behavior. The conclusion is that the (historically important) question of evaluating voter behavior is trivial.

Our multiplicity results show that this conclusion is unwarranted. When candidates are heterogeneous, there is a more subtle way to evaluate voter performance: Are voters using the reelection rule that creates the strongest possible incentives relative to the set of all reelection rules consistent with an equilibrium? Theorem 5.1 and Propositions 5.1–5.2 say that this set may be non-trivial and so there remains scope for evaluating voter performance.

That said, our results are suggestive that the standard evaluative procedure may be problematic. The literature presumes that increasing effective accountability is necessarily desirable from the perspective of voter welfare. But, as we showed in Examples 3.1 and 5.5, this need not be the case: When effort and type are substitutes, a higher level of effective accountability reduces the quality of 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. As such, our analysis suggests a rethinking of the evaluative criteria that underlies the existing literature—away from the level of effective accountability and towards a more comprehensive measure of voter welfare.

Theorems 3.1 and 4.1 also point to subtleties in how observed voter behavior may be mapped into evaluative measures. For instance, suppose we observe a voter adopting a more lenient benchmark by which to hold politicians accountable. The natural conclusion is that this reflects a decline in voter performance. However, our results show that this need not be the case. When $\pi_C > \pi_I$, it may well be welfare improving for the voter to adopt a more lenient benchmark. In the case of substitutes, this is because the more lenient benchmark corresponds to a lower level of effective accountability which, in turn, can improve voter welfare. In the case of complements, this is because the more lenient benchmark can correspond to a higher level of effective accountability which, in turn, improves voter welfare.
Appendix A  Proof for Sections 3

Fix some symmetric model \( \mathcal{M}(B,c) \) where \( \pi_I = \pi_C = \pi \). Consider an equilibrium associated with the first-period action \( a_* \). Write \( \Pr(\theta_{P2} = \bar{\theta}|a_*) \) for the probability that the Politician in office in the second period is type \( \bar{\theta} \). Note, the event “the Politician in office in the second period is type \( \bar{\theta} \)” can be realized in one of three ways: (i) the first-period Incumbent is type \( \bar{\theta} \) and is reelected, (ii) the first-period Incumbent and the Challenger are both type \( \bar{\theta} \) and the first-period Incumbent is not reelected, or (iii) the first-period Incumbent is type \( \bar{\theta} \), she is not reelected, and the Challenger is type \( \bar{\theta} \). Thus,

\[
\Pr(\theta_{P2} = \bar{\theta}|a_*) = \pi \Pr(f(a_*, \bar{\theta}) + \epsilon_1 \geq \hat{g}(a_*)) + \pi^2 \Pr(f(a_*, \bar{\theta}) + \epsilon_1 \leq \hat{g}(a_*)) + (1 - \pi) \pi \Pr(f(a_*, \bar{\theta}) + \epsilon_1 \leq \hat{g}(a_*)).
\]

(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 \( \Pr(f(a_*, \theta) + \epsilon_1 \geq \hat{g}(a_*)) = 1 - \Phi(\hat{g}(a_*) - f(a_*, \theta)) \),

\[
\Pr(\theta_{P2} = \bar{\theta}|a_*) = \pi + (1 - \pi)\pi[2\Phi\left(\frac{f(a_*, \bar{\theta}) - f(a_*, \bar{\theta})}{2}\right) - 1].
\]

The next lemma is an immediate consequence of this calculation.

Fix some \( a \in A \). It will be convenient to introduce a function \( \text{VW}_1[a] : [0,1] \rightarrow \mathbb{R} \) and \( \text{VW}_2[a] : [0,1] \rightarrow \mathbb{R} \) so that

\[
\text{VW}_1[a](p) = [f(a, \bar{\theta}) - f(a, \theta)]p + f(a, \theta)
\]

and

\[
\text{VW}_2[a](p) = (f(a, \bar{\theta}) - f(a, \theta)) \left[ p + (1 - p)p \left(2\Phi\left(\frac{f(a, \bar{\theta}) - f(a, \bar{\theta})}{2}\right) - 1\right)\right] + f(a, \theta).
\]

Then, define \( \text{VW}[a] = \text{VW}_1[a] + \text{VW}_2[a] \). Note, if there is an equilibrium where \( a \) is the first-period action and \( p \in (0,1) \) is \( p = \pi \), then \( \text{VW}_1[a](p) \) 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.

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, \( \rho_* = f(a_*, \bar{\theta}) - f(a_*, \theta) > 0 \), \( \rho_{**} = f(a_{**}, \bar{\theta}) - f(a_{**}, \theta) > 0 \), and \( \psi_{**} = f(a_{**}, \bar{\theta}) - f(a_*, \theta) > 0 \).

| \( a_{**} \) | \( \psi_{**} + \rho_{**} \) | \( \psi_{**} + \psi_{**} + \rho_{**} \) |
| \( a_* \) | \( \psi_{**} \) | \( \psi_{**} + \rho_* \) |

Figure A.1: The Function \( f \)

**Lemma A.1.** \( \psi_{**} + \rho_{**} - \rho_* > 0 \).
Proof. Note that \( f(a_{**}, \theta) - f(a_{*}, \theta) = \psi_{**} + \rho_{**} - \rho_{*} \), from which the claim follows. ■

Proof of Proposition 3.1. 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 \( \Delta[a_{**}, a_{*}] = VW[a_{**}] - VW[a_{*}] \). Note, this is a continuous function with \( \Delta[a_{**}, a_{*}](0) = \psi_{**} > 0 \) and (by Lemma A.1) \( \Delta[a_{**}, a_{*}](1) = \psi_{**} + \rho_{**} - \rho_{*} > 0 \). Moreover, the derivative of the difference function with respect to \( p \) is

\[
(p_{**} - \rho_{*}) + 2[f(a, \theta) - f(a_{*}, \theta)] \left[ \Phi \left( \frac{\rho_{**}}{2} \right) - \Phi \left( \frac{\rho_{*}}{2} \right) \right] (1 - 2p).
\]

Thus, if effort and type are strict substitutes (respectively, strict complements) at \( a_{*}, a_{**} \), the difference function is strictly decreasing (respectively, strictly increasing) on \((0, p_{\text{min}}) \neq \emptyset \) (respectively, \((0, p_{\text{max}}) \neq \emptyset \) and strictly increasing (respectively, strictly decreasing) on \((p_{\text{min}}, 1) \neq \emptyset \) (respectively, \((p_{\text{max}}, 1) \neq \emptyset \)). 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 \( \pi[a_{*}, a_{**}], \pi[a_{*}, a_{**}] \in (0, 1) \) so that \( p \in (0, \pi[a_{*}, a_{**}]) \cup (\pi[a_{*}, a_{**}], 1) \) if and only if \( \Delta[a_{**}, a_{*}](p) \geq 0 \). ■

Appendix B Proofs for Section 4

Prior to presenting the results in this appendix, we will establish a useful equivalence: Fix a real-valued random variable that is distributed according to some distribution full support differentiable PDF \( \psi \). Define \( \mathbb{L} : \mathbb{R}^3 \to \mathbb{R} \) so that

\[
\mathbb{L}(y, x, x') := \frac{\psi(y - x)}{\psi(y - x')}. 
\]

Say \( \psi \) satisfies the strict MLRP if, for each \( x > x' \), \( \mathbb{L}(\cdot, x, x') \) is strictly increasing. Observe that, if \( \psi \) satisfies the strict MLRP and each \( \mathbb{L}(\cdot, x, x') \) has non-zero derivatives, then the derivative of each \( \mathbb{L}(\cdot, x, x') \) with respect to \( y \) is strictly positive (at each point).

Lemma B.1.

(i) If \( \psi \) satisfies the strict MLRP and each likelihood ratio \( \mathbb{L}(\cdot, x, x') \) has non-zero derivatives, then \( \log \psi \) is strictly concave.

(ii) If \( \log \psi \) is strictly concave, then \( \psi \) satisfies the strict MLRP and each likelihood ratio \( \mathbb{L}(\cdot, x, x') \) has non-zero derivatives.

Remark B.1. For each \( \tilde{y} \) and \( x > x' \)

\[
\left. \frac{\partial \mathbb{L}(\cdot, x, x')}{\partial y} \right|_{\tilde{y}} = \frac{\psi'(\tilde{y} - x)\psi(\tilde{y} - x') - \psi(\tilde{y} - x)\psi'(\tilde{y} - x')}{\psi(\tilde{y} - x')^2} > 0
\]
if and only if

$$(\log \circ \psi)'(\tilde{y} - x) = \frac{\psi'(\tilde{y} - x)}{\psi(\tilde{y} - x)} > \frac{\psi'(\tilde{y} - x')}{\psi(\tilde{y} - x')} = (\log \circ \psi)'(\tilde{y} - x').$$

**Proof of Lemma B.1.** First suppose that $\psi$ satisfies the strict MLRP and each likelihood ratio $\mathbb{L}(:, x, x')$ has non-zero derivatives. Then, for each $\tilde{y}$, $\frac{\partial \mathbb{L}(:, x, x')}{\partial y} \bigg|_{\tilde{y}} > 0$. For each $z' > z$, choose $x > x'$ and $y'$ so that $z' = y' - x'$ and $z = y' - x$. It follows from Remark B.1 that $(\log \circ \psi)'(z') < (\log \circ \psi)'(z)$ and so $\log \circ \psi$ is concave.

Next suppose $\log \circ \psi$ is strictly concave. The fact that $\psi$ satisfies the strict MLRP is standard. (E.g., in Lehmann and Romano (2006, Example 8.2.1), replace weak inequality with strict inequality.) Remark B.1 completes the proof. ■

**Proof of Lemma 4.1.** Observe that, for each $g$, there exists some $g > g'$ so that $g = g' + (\nu(a^{**}) - \nu(a^*))$. Since

$$\hat{L}(g) = \frac{\phi(g' - \nu(a^*) - \Delta a^*)}{\phi(g' - \nu(a^*) + \Delta a^*)} = LR_*(g')$$

and $LR_*(g) > LR_*(g')$, it follows that $LR_*(g) > \hat{L}(g)$. Since $\hat{L}$ is strictly increasing, $\hat{L}(\hat{g}) = \beta$ implies that $\tilde{g} > \hat{g}(a^*)$. ■

**Lemma B.2.** Suppose $\phi$ is symmetric. Consider a function $\mathcal{L} : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}_*$ given by

$$\mathcal{L}(x, \delta) = \frac{\phi(x - \delta)}{\phi(x + \delta)}.$$

(i) $\mathcal{L}$ is increasing in $\delta$ if $x > 0$.

(ii) $\mathcal{L}$ is decreasing in $\delta$ if $x < 0$.

**Proof of Lemma B.2.** First, define $Q(x, \delta) = x - \delta$ and $\overline{Q}(x + \delta)$. Note that

$$\log \mathcal{L}(x, \delta) = \log \phi(Q(x, \delta)) - \log \phi(\overline{Q}(x, \delta)).$$

Now, differentiate with respect to $\delta$ to get:

$$\frac{\partial}{\partial \delta} \log \mathcal{L}(x, \delta) = \frac{\phi'(Q(x, \delta))}{\phi(Q(x, \delta))} \frac{\partial Q(x, \delta)}{\partial \delta} - \frac{\phi'(\overline{Q}(x, \delta))}{\phi(\overline{Q}(x, \delta))} \frac{\partial \overline{Q}(x, \delta)}{\partial \delta}$$

$$= -\frac{\phi'(x - \delta)}{\phi(x - \delta)} \frac{\phi'(x + \delta)}{\phi(x + \delta)}.$$

Since $\phi$ is symmetric, $\phi'(y) = -\phi'(-y)$. So, for all $\delta$, $\frac{\partial}{\partial \delta} \log \mathcal{L}(0, \delta) = 0$. With this, the claim will follow if $\frac{\partial}{\partial \delta} \log \mathcal{L}(x, \delta)$ is increasing in $x$. 27
Differentiating $\log \mathcal{L}(x, \delta)$,

$$\frac{\partial^2}{\partial \delta \partial x} \log \mathcal{L}(x, \delta) = - \frac{d}{dQ} \left( \frac{\phi'(Q(x, \delta))}{\phi(Q(x, \delta))} \right) \frac{\partial Q(x, \delta)}{\partial x} - \frac{d}{dQ} \left( \frac{\phi'(Q(x, \delta))}{\phi(Q(x, \delta))} \right) \frac{\partial Q(x, \delta)}{\partial x}$$

$$= - \frac{d}{dQ} \left( \frac{\phi'(Q(x, \delta))}{\phi(Q(x, \delta))} \right) - \frac{d}{dQ} \left( \frac{\phi'(Q(x, \delta))}{\phi(Q(x, \delta))} \right).$$

By Lemma B.1, $\phi$ is strictly log concave. From this, the last equivalence is strictly positive. ■

Proof of Lemma 4.2. Immediate from Lemma B.2. ■

Appendix C  Formal Accountability

Are there systematic changes in the institutions of formal accountability that serve to increase effective accountability? The main text provides examples (Examples 3.1-4.1) where higher levels of $B$ correspond to higher levels of effective accountability. But, this monotonicity need not hold more generally, as the next example shows.

Example C.1. Let $A = \{0, 1, 2\}$. The production function $f$ is described by the following table.

<table>
<thead>
<tr>
<th>$\theta$</th>
<th>$\bar{\theta}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>6.5</td>
</tr>
<tr>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 3: Production Function

Take $\Phi$ to be the CDF of the standard Normal distribution and let $\pi_I = \pi_C = .5$.

Using Example 2.1, the incremental returns satisfy the following bounds:

| IR (1, 0|1) $\in (0.0334, 0.0335)$ | IR (1, 0|1) $\in (0.4471, 0.4472)$ | IR (1, 0|2) $\in (0.4471, 0.4472)$ |
|-----------------------------|-----------------------------|-----------------------------|
| IR (2, 0|0) $\in (0.4999, 0.5001)$ | IR (2, 0|1) $\in (0.8943, 0.8944)$ | IR (2, 0|2) $\in (0.4999, 0.5001)$ |
| IR (2, 1|0) $\in (0.4665, 0.4666)$ | IR (2, 1|1) $\in (0.4471, 0.4472)$ | IR (2, 1|2) $\in (0.0528, 0.0529)$ |

Table 4: Incremental Returns in Reelection Probabilities

We will consider two models parameterized by $(B_{**}, c_{**}) = (1, c_{**})$ and $(B_*, c_*) = (5, c_*)$, where $c_{**}(2) = c_*(2) = 2.5$, $c_{**}(1) = c_*(1) = .4$, and $c_{**}(0) = c_*(0) = 0$. It can be verified that $\mathcal{E}(1, c_{**}) = \{0, 1\}$ and $\mathcal{E}(5, c_*) = \{0\}$. Thus, the model associated with $(B_{**}, c_{**}) = (1, c_{**})$ has an equilibrium with a higher level of effective accountability than any equilibrium of the model associated with $(B_*, c_*) = (5, c_*)$, even though $B_{**} < B_*$. 28
We can reparameterize Examples 3.1, 4.1, and C.1 to show that, holding B fixed and decreasing c also does not lead to systematic changes in effective accountability.

Appendix D  Proofs for Section 5

A strategy of the Incumbent can be written as \( s_I = (a, s_{I,2}) \), where \( a \) is a first-period action and \( s_{I,2} \), a second-period plan, maps each first-period level of public goods plus a decision to reelect the Incumbent to a second period action. 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 \( s_{I,2}(g) \). A strategy of the Voter \( s_V \) maps each level of public goods provided (in the first period) to a reelection decision. Thus, \( s_V(g) = 1 \) (\( s_V(g) = 0 \)) represents the fact that the Voter reelects (replaces) the Incumbent if \( g \) is the level of first-period public goods. A strategy for the Challenger \( s_C \) maps each level of first-period public goods provided and a decision to replace the Incumbent to a second period action. 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 \( s_C(g) \).

**Lemma D.1.** The pair \((B, c)\) justifies \( a_\ast \) if and only if \( c(a) \geq \text{IR}(a, a_\ast | a_\ast)(B - c(a)) + c(a_\ast) \) for each \( a \in A \).

**Proof.** If \((B, c)\) justifies \( a_\ast \), then

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

for each action \( a \in A \). From this the claim follows. Conversely, suppose that \( c(a) \geq \text{IR}(a, a_\ast | a_\ast)(B - c(a)) + c(a_\ast) \) for each \( a \in A \). Construct \((s_I, s_V, s_C)\) so that (i) \( s_I = (a_\ast, s_{I,2}) \), (ii) for each realization of public goods \( g \), \( s_{I,2}(g) = s_C(g) = a_\ast \), and (iii) \( s_V(g) = 1 \) if and only if \( g \geq \hat{g}(a_\ast) \). It is readily verified that \((s_I, s_V, s_C)\) is a perfect Bayesian equilibrium.

We begin with a Lemma that will be of use:

**Lemma D.2.** Fix some \( a_\ast \) and some \( a > a' \).

(i) \( \Pr(a | a_\ast) > \Pr(a' | a_\ast) \)

(ii) \( \text{IR}(a, a' | a_\ast) > 0 \).

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

**Proof.** Part (i) follows from the fact that \( f \) is increasing in \( a \). Part (ii) follows from Part (i). Part (iii) follows from Part (i), since \( \text{IR}(a_\ast, a' | a_\ast) = \Pr(a_\ast | a_\ast) - \Pr(a' | a_\ast) > \Pr(a_\ast | a_\ast) - \Pr(a | a_\ast) = \text{IR}(a_\ast, a | a_\ast) \).

Before coming to the proof of Theorem 5.1 we provide a sketch of the argument. Consider the case where \( a_\ast = a \). The idea will be to fix a benefit of reelection, viz. \( B \), and constants \( n_\ast \) and \( n_{\ast\ast} \).
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 $a_*$ and $a_{**}$, it will satisfy $c(a_*) = n_*$ and $c(a_{**}) = n_{**}$. As such, we fix $B > n_*> n_*$ so that

$$
IR(a_{**}, a_*|a_{**}) \geq n_{**} - n_* B - n_*$
$$

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

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

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

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

Figure D.1 illustrates the functions $N(\cdot, a_*)$ and $N(\cdot, a_{**})$. They are each strictly increasing. Moreover, by Equation (9), $N(a_*, a_*) = n_* \geq N(a_*, a_{**})$ and $N(a_{**}, a_{**}) = n_{**} \geq N(a_{**}, a_*)$. Thus, taking $c$ to be the upper envelope of the functions, $(B, c)$ justifies both $a_*$ and $a_{**}$.

**Proof of Theorem 5.1: Part (i) if and only if Part (ii).** The fact that part (ii) implies part (i) follows immediately from Lemma D.1. We show that part (i) implies part (ii). Suppose that $IR(a_{**}, a_*|a_{**}) \geq IR(a_{**}, a_*|a_*)$ and we will show that we can construct $(B, c)$ that justifies both $a_{**}$ and $a_*$.

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

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

\(iii_s \ n + [\text{IR}(a_{ss}, a|a_{ss}) - \text{IR}(a_{sss}, a\mid a_s)](B - n) > n_s.\)

The fact that requirements \(i_s - ii_s\) can be satisfied simultaneously follows from Lemma D.2(ii).

The fact that requirements \(i_s - iii_s\) can be satisfied simultaneously follows from \(\text{IR}(a_{ss}, a|a_{ss}) > \text{IR}(a_{ss}, a\mid a_s)\) (Lemma D.2(iii)) and \(\text{IR}(a_{ss}, a\mid a_{ss}) \geq \text{IR}(a_{ss}, a\mid a_s)\) (by assumption), so that \(\text{IR}(a_{ss}, a\mid a_{ss}) > \text{IR}(a_{ss}, a\mid a_s).\)

Now fix \(n_{ss}\)

\(i_{ss} \ n_{ss} \geq n_s + \text{IR}(a_{ss}, a\mid a_s)(B - n);\)

\(ii_{ss} \ n_s + \text{IR}(a_{ss}, a\mid a_s)(B - n) \geq n_{ss};\) and

\(iii_{ss} \ n + \text{IR}(a_{ss}, a\mid a_s)(B - n) \geq n_{ss}\) with strict inequality if \(a_s \neq a.\)

The fact that requirements \(i_{ss} - ii_{ss}\) can be satisfied simultaneously follows from the assumption that \(\text{IR}(a_{ss}, a\mid a_s) \geq \text{IR}(a_{ss}, a\mid a_s).\) Condition \(iii_{ss}\) follows from condition \(ii_{ss},\) if \(a_s = a.\) The fact that requirements \(i_{ss} - iii_{ss}\) can be satisfied simultaneously when \(a_s > a\) follows from condition \(iii_s\) above. Note, it follows from Lemma D.2 and \(i_{ss}\) that \(n_{ss} > n_s.\)

Construct a function \(N : A \times \{a_s, a_{ss}\} \to \mathbb{R}\) so that \(N(a, a_s) = \text{IR}(a, a\mid a_s)(B - n) + n_s\) and \(N(a, a_{ss}) = \text{IR}(a, a\mid a_{ss})(B - n) + n_{ss}.\) It follows from Lemma D.2 that \(N(\cdot, a_s)\) and \(N(\cdot, a_{ss})\) are strictly increasing in \(a.\) Moreover,

- \(n \geq \max\{N(a, a_s), N(a, a_{ss})\};\)
- \(n_s = N(a_s, a_s) \geq N(a_s, a_{ss});\)
- \(n_{ss} = N(a_{ss}, a_{ss}) \geq N(a_{ss}, a_s).\)

The first of these follows from requirement \(ii_s\) on \(n_s\) and requirement \(iii_{ss}\) on \(n_{ss}.\) The second of these follows from requirement \(ii_{ss}\) on \(n_{ss}.\) The third of these follows from requirement \(i_{ss}\) on \(n_{ss}.

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

- \(\hat{N}(a) \leq n;\)
- \(\hat{N}(a_s) = n_s;\) and
- \(\hat{N}(a_{ss}) = n_{ss}.\)

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

\(^{11}\)Of course, \(n = N(a, a_s)\) if \(a_s = a.\)
• $c(a) = n$,
• $c(a_*) = n_*$, and
• $c(a_{**}) = n_{**}$.

Applying Lemma D.1 we get that the $(B, c)$ constructed justifies both $a_{**}$ and $a_*$. ■

Preferences $(B, c)$ and $(B', c')$ represent different preferences over risk if there is no $(x, y) \in \mathbb{R} \times \mathbb{R}_+$ and so that $c'(a) = x + yc(a)$ for all $a \in A$.

Proof of Theorem 5.1: Part (iii) if and only if Part (i). It is immediate that part (iii) implies part (ii). Fix $a_{**} > a_*$, so that the pair $(a_*, a_{**})$ is justifiable and we will show that part (iii) is satisfied.

By the proof that part (i) implies part (ii), we can find a part $(B, c)$ that justifies both $a_*$ and $a_{**}$ and satisfies the following requirements:

i. if $a_* \neq a$, $c(a_*) > c(a)$;

ii. if $a_* \neq a$, $c(a) + \text{IR}(a_*, a|a_{**})(B - c(a)) > c(a_*)$;

iii. if $a_* \neq a$, $c(a) + [\text{IR}(a_{**}, a|a_{**}) - \text{IR}(a_{**}, a_*|a_*)](B - c(a)) > c(a_*)$;


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

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

iii. $c(a_*) + \text{IR}(a_{**}, a_*|a_*)(B - c(a)) \geq c(a_{**})$ with strict inequality if $a_* \neq a$.

Let $\alpha = \frac{c(a_{**}) - c(a_*)}{B - c(a)}$ and observe that

$$\text{IR}(a_{**}, a_*|a_*) \geq \alpha \geq \text{IR}(a_{**}, a_*|a_*) > 0.$$ 

For a given $\varepsilon > 0$, define

• $B' = B + 2\varepsilon$,
• $c'(a_{**}) = c(a_{**}) + \varepsilon(\alpha + 1)$,
• $c'(a_*) = c(a_*) + \varepsilon$, and
• $c'(a) = c(a) + \varepsilon$.

(Observe that $c'(a_{**}) > c'(a_*)$.) We will show that we can construct some pair $(\hat{B}, \hat{c})$ so that $\hat{B} = B'$, $\hat{c}(a_{**}) = c'(a_{**})$, $\hat{c}(a_*) = c'(a_*)$, $\hat{c}(a) = c'(a)$, and $(\hat{B}, \hat{c})$ justifies $(a_*, a_{**})$. Since $(\hat{B}, \hat{c})$ represents different preferences than $(B, c)$ and $\varepsilon > 0$ is chosen arbitrarily, this establishes the desired result.

First observe we have the following properties:
\[ i' \text{ if } a_s \neq a, \ c'(a_s) > c'(a); \]

\[ ii' \text{ if } a_s \neq a, \ c'(a) + \text{IR} (a_s, a|a_s)(B' - c'(a)) > c'(a); \]

\[ iii' \text{ if } a_s \neq a, \ c'(a) + [\text{IR} (a_s, a|a_s) - \text{IR} (a_{**}, a_s|a_s)](B' - c'(a)) > c'(a); \]

Property \( i' \) is immediate. Property \( ii' \) follows from \( ii \). Property \( iii' \) follows from \( iii \). Next observe that

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

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

\[ iii'_{**} \ c'(a) + \text{IR} (a_{**}, a|a_{**})(B' - c'(a)) \geq c'(a_{**}) \text{ with strict inequality if } a_s \neq a. \]

Property \( i'_{**} \) follows from \( i_{**} \) and the fact that \( \alpha \geq \text{IR} (a_{**}, a_s|a_s). \) Property \( ii'_{**} \) follows from \( ii_{**} \) and the fact that \( \text{IR} (a_{**}, a_s|a_s) \geq \alpha. \) Part \( iii'_{**} \) follows from \( iii_{**} \) and the fact that \( \text{IR} (a_{**}, a|a_{**}) \geq \text{IR} (a_{**}, a_s|a_s) \geq \alpha. \]

We next turn to prove Propositions 5.1-5.3. Define numbers

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

and

\[ Y \equiv \left[ \Phi \left( \hat{g} (a_{**}) - f(a_{**}, \emptyset) \right) - \Phi \left( \hat{g} (a_s) - f(a_{**}, \emptyset) \right) \right] - \left[ \Phi \left( \hat{g} (a_{**}) - f(a_{**}, \emptyset) \right) - \Phi \left( \hat{g} (a_s) - f(a_{**}, \emptyset) \right) \right]. \]

\textbf{Lemma D.3.} If \( \pi_I = \pi_C = \pi, \) then the following are equivalent:

\begin{enumerate}
  \item \( \text{IR} (a_{**}, a_s|a_s) \geq \text{IR} (a_{**}, a_s|a_s). \)
  \item \( \pi X \geq (1 - \pi) Y. \)
\end{enumerate}

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

\textbf{Lemma D.4.} Let \( \phi \) be symmetric.

\begin{enumerate}
  \item \( \Phi(x) = 1 - \Phi(-x). \)
  \item If \( y, z > 0 \) and \( x \in (-y + z), y, \) then \( \Phi(x + z) - \Phi(x) > \Phi(y + z) - \Phi(y). \)
\end{enumerate}

\textbf{Proof.} Part (i) follows symmetry, since

\[ 1 - \Phi(-x) = 1 - \int_{-\infty}^{-x} \phi(q) dq = 1 - \int_{-\infty}^{\infty} \phi(q) dq = \int_{-\infty}^{x} \phi(q) dq = \Phi(x). \]
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 \( \Phi(x + z) - \Phi(x) > \Phi(y + z) - \Phi(y) \). If \( x \in (-y - z, 0) \), then

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

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.

**Lemma D.5.** If \( \phi \) is symmetric, then \( X > 0 \).

**Proof.** Applying Lemma D.4(i), \( X \) can be written as

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

Write \( d = \rho_{**} + \psi_{**} - \rho_* \) and note that, by Lemma A.1, \( d > 0 \). We can then rewrite \( X \) as

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

Thus, by Lemma D.4(ii), \( X > 0 \) provided (i) \( \rho_* - \psi_{**} - \frac{1}{2} \rho_{**} > -\frac{1}{2} \rho_* - d \) and (ii) \( \frac{1}{2} \rho_* > \rho_\ast - \psi_{**} - \frac{1}{2} \rho_{**} \).

Condition (i) is immediate and Condition (ii) is by Lemma A.1.

**Lemma D.6.** If \( \phi \) is symmetric, then \( Y > 0 \).

**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.1, \( \frac{1}{2} \rho_{**} > \frac{1}{2} \rho_* - \psi_{**} \). Since \( \psi_{**} > 0 \) and \( \frac{1}{2} \rho_* - \psi_{**} > -\frac{1}{2} \rho_{**} - \psi_{**} \), we can apply Lemma D.4(ii) to conclude that \( Y > 0 \).

**Lemma D.7.** Suppose the model is symmetric with \( \pi_I = \pi_C = \pi \). If \( \pi' > \pi \) and \( \pi X \geq (1 - \pi) Y \), then \( \pi' X \geq (1 - \pi') Y \).

**Proof.** Immediate from Lemmata D.5-D.6.

The following is immediate from Lemmata D.3 and D.7.
Remark D.1. Suppose the model is symmetric with $\pi_I = \pi_C = \pi$. Then $\text{IR} \left( a_{**}, a_* | a_* \right)$ is strictly increasing in $\pi$.

Proof of Proposition 5.2. Take $\hat{\pi}[a_*, a_{**}] = \frac{Y}{\psi}$. By Lemmata D.5-D.6, $\hat{\pi}[a_*, a_{**}] \in (0, 1)$. Now, the pair $(a_*, a_{**})$ is justifiable if and only if $\pi \geq \hat{\pi}[a_*, a_{**}]$.

To conclude the proof, we 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 D.4(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.1, $\rho_{**} + \psi_{**} - \frac{1}{2} \rho_* > 0$. Moreover, $\frac{1}{2} \rho_* - \psi_{**} > -(\rho_{**} + \psi_{**} - \frac{1}{2} \rho_* - d) = -\frac{1}{2} \rho_{**} - \psi_{**}$. By Lemma D.4(ii), $X - Y < 0$.  

Proof of Proposition 5.3. Observe that $\hat{g}$ is a function, not only of $a$, but of $\pi_I$ and $\pi_C$. It will be convenient to write $\hat{g} : A \times [0, 1] \times [0, 1] \to \mathbb{R}$, so that, for each $(a, \pi_I, \pi_C)$, $\hat{g}(a, \pi_I, \pi_C)$

$$\text{LR} \left( \hat{g}(a, \pi_I, \pi_C), a \right) = \frac{\pi_C}{1 - \pi_C} \frac{1 - \pi_I}{\pi_I}. $$

Thus, for each $a, a' \in A$, $\text{IR} \left( a, a' | \cdot \right) : A \times [0, 1] \times [0, 1] \to \mathbb{R}$ is also a function of $(a, \pi_I, \pi_C)$.

It suffices to show that, for each $(a, a')$ and each $(a_*, \cdot, \cdot)$, $\text{IR} \left( a, a' | a_* \cdot, \cdot \right) : [0, 1] \times [0, 1] \to \mathbb{R}$ is continuous in $(\pi_I, \pi_C)$. If so, by Proposition 5.2, for each $(\hat{\pi}, \tilde{\pi}) \in (\hat{\pi}[a_*, a_{**}], 1)^2$,

$$\text{IR} \left( a_{**}, a_* | a_* \hat{\pi}, \hat{\pi} \right) - \text{IR} \left( a_{**}, a_* | a_* \tilde{\pi}, \tilde{\pi} \right) > 0.$$

By continuity, for each $(\hat{\pi}, \tilde{\pi}) \in (\hat{\pi}[a_*, a_{**}], 1)^2$, there is some open neighborhood $N$ of $(\hat{\pi}, \tilde{\pi})$ so that, for all $(\pi_I, \pi_C) \in N$

$$\text{IR} \left( a_{**}, a_* | a_* \pi_I, \pi_C \right) - \text{IR} \left( a_{**}, a_* | a_* \pi_I, \pi_C \right) > 0.$$

By Theorem 5.1, the claim follows.
To show that IR \((a, a', a_x, \cdot, \cdot, \cdot)\) it continuous in \((\pi_I, \pi_C)\), it suffices to show that, for each \(a\), \(\hat{g}(a, \cdot, \cdot)\) is continuous in \((\pi_I, \pi_C)\). If so, then by continuity of \(\Phi\), IR \((a, a'|a_x, \cdot, \cdot)\) it continuous in \((\pi_I, \pi_C)\).

Write \(LR : \mathbb{R} \times A \times [0, 1] \times [0, 1] \rightarrow \mathbb{R}\) so that

\[
LR(g, a, \pi_I, \pi_C) = \frac{\phi(g - f(a, \theta))}{\phi(g - f(a, \theta))} - \pi_I - 1 - \frac{\pi_C}{1 - \pi_I}.
\]

Observe that, for each \(a\), \(LR(\cdot, a, \cdot, \cdot) : \mathbb{R} \times [0, 1] \times [0, 1] \rightarrow \mathbb{R}\) is differentiable in \(g\) and continuous in \((g, \pi_I, \pi_C)\). (This follows from the fact that \(\phi\) is continuously differentiable.) Moreover, \(LR(\cdot, a)\) is differentiable with a non-zero derivatives (since \(LR\) is). Thus, it follows from the Implicit Function Theorem that each \(\hat{g}(a, \cdot, \cdot)\) is continuous in \((\pi_I, \pi_C)\).

**Proof of Proposition 5.1.** Immediate from Proposition 5.3.

**References**


Landa, Dimitri. 2010. “Selection Incentives and Accountability Traps: A Laboratory Experiment.” NYU Typescript.


